Progress, Challenges, Threats and Prospects of ChatGPT in Science and Education: How will AI Impact the Academic Environment?

Dilshod Nematov[1,2]

[1] S.U. Umarov Physical-Technical Institute of NAST, 734063 Dushanbe, Tajikistan.
[2] School of Optoelectronic Engineering & CQUPT-BUL Innovation Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

Email: dilnem@mail.ru
Manuscript submitted April 24, 2025; accepted June 3, 2025; published July 17, 2025.
doi: 10.18178/JAAI.2025.3.3.187-205

Abstract: The rapid development of Artificial Intelligence (AI), particularly in the form of language models such as OpenAI's ChatGPT, is fundamentally transforming the academic and educational landscape. AI tools like ChatGPT play a crucial role in accelerating text generation, enhancing writing processes, assisting in literature reviews, and facilitating personalized learning. The development and adoption of Generative Pre-trained Transformers (GPT), including ChatGPT, have sparked widespread interest due to their unprecedented capabilities in generating human-like text and automating communication. These advancements have a significant impact on various fields, including education and research, where AI-driven tools can optimize academic work and improve accessibility to knowledge. However, the integration of AI into academia comes with several serious challenges, such as algorithmic bias, data privacy concerns, ethical issues including plagiarism, and the potential decline in critical thinking skills due to excessive reliance on AI-generated content. While some educators and researchers view ChatGPT as a powerful tool for innovation, others express concerns about its impact on academic integrity and analytical reasoning. This review also compares ChatGPT with other advanced AI models, such as GPT-3 and BERT, highlighting their respective roles in science and education. An analysis of 159 independent literature sources, including 36 recently published papers on arXiv.org platform, indicates that interest in ChatGPT and its analogs in the field of science and education is rapidly growing. These studies emphasize the importance of AI in optimizing academic activities, improving knowledge accessibility, and developing new learning methods. The article concludes by discussing the future evolution of AI in academia, its integration with emerging technologies, and its potential for fostering interdisciplinary research in an era often described as the "new AI gold rush".

Keywords: ChatGPT, artificial intelligence, AI tools, language models, education, research, academic writing, AI applications, GPT-3, GPT-4, BERT, ethical considerations, future of academia

1. Introduction

Recent advancements in Artificial Intelligence (AI) have significantly transformed the scientific and educational landscape. One of the most striking examples of such changes is the development of language models, such as Chat Generative Pre-trained Transformer (ChatGPT) and its analogs. These models serve as powerful tools capable of accelerating and improving the process of scientific text preparation, data analysis, and even the generation of innovative solutions in various fields of science and education [1, 2]. ChatGPT is one of the most widely used language models developed by OpenAI. Initially based on the GPT-3 architecture, it has since been enhanced in subsequent versions, including GPT-4 and GPT-5. Each new version brings significant improvements in query processing accuracy and the ability to generate more meaningful and coherent responses [3]. For instance, ChatGPT-4 has demonstrated enhanced contextual understanding, better dialogue management, and deeper comprehension of complex technical and scientific concepts, making it an invaluable tool for drafting scientific articles [4].

The ChatGPT Plus version is available through a paid subscription, offering users improved performance, faster response times, and priority access during peak server loads. Notably, the upcoming ChatGPT-5 version is expected to feature further advancements in text generation and specialized problem-solving capabilities, particularly in mathematical and scientific domains [5]. Several AI models compete with ChatGPT in this domain. For example, Google's Bidirectional Encoder Representations from Transformers (BERT) is designed for bidirectional text processing, improving its contextual understanding capabilities [6]. Another model, Google's Text-to-Text Transfer Transformer (T5), employs a unified architecture to address a variety of tasks [7]. Meanwhile, Microsoft's Turing-NLG focuses on natural text generation, boasting a high number of parameters and strong long-term reasoning capabilities [8]. When comparing these models, ChatGPT surpasses many of its counterparts in accessibility, ease of use, and real-time contextual conversation generation [9]. However, models such as BERT and T5 may be more precise in tasks requiring deep text analysis or working with limited data sets [10]. While Turing-NLG performs well in text generation, it sometimes falls short in terms of interactivity and adaptability compared to ChatGPT [11].

Given the rapid evolution of such technologies, the near future is likely to witness not only a further acceleration in scientific material creation but also the development of innovative approaches in education and research. This review will explore the achievements, challenges, and prospects of AI models like ChatGPT in scientific and educational domains, as well as their impact on academia and text preparation.

2. Applications of ChatGPT in Academic Direction

Since its introduction, ChatGPT and other similar AI language models, such as GPT-3 and GPT-4, have gained widespread attention in the scientific and educational community due to their ability to generate text, answer questions, and provide recommendations. These tools have proven beneficial in various fields, including scientific writing, research activities, and the development of educational materials [12]. For instance, ChatGPT can be employed to automatically generate scientific abstracts, draft manuscript sections, and even compose entire portions of academic papers, significantly reducing the time required for manuscript preparation [13].

However, the application of such tools in scientific research necessitates a cautious approach. This is particularly important when using AI models for data analysis and interpretation. While AI models can be highly effective in text generation, they are not always capable of accurately interpreting complex concepts or considering the nuanced aspects of scientific research, which may impact the reliability of conclusions. Therefore, these tools should be utilized as supplementary aids rather than primary means of analysis.

With the advancement of newer versions like GPT-4 and GPT-5, these models now support not only text input but also effectively process images and other data formats, opening new possibilities for interdisciplinary research [3, 14]. For example, ChatGPT-4 and ChatGPT-5 can assist in generating ideas for scientific experiments, accelerating the development of new technologies such as photocatalysts for green hydrogen production [15]. Research environments have already begun experimenting with these models to enhance processes in bioengineering, materials science, and other disciplines. Additionally, alternative AI models, such as Google Bard and Claude by Anthropic, offer competitive functionalities for addressing various scientific and educational challenges. For instance, Claude-2 prioritizes safety and ethics in text generation, making it a suitable choice for educational institutions where the prevention of misinformation is crucial [16]. Meanwhile, Google Bard exhibits high accuracy and speed in information processing, making it an effective tool for scientific computations and data analysis, particularly in chemistry and materials physics [17–22].

The integration of tools like ChatGPT significantly influences the process of scientific text creation and processing. These systems assist researchers in drafting manuscripts, thereby expediting the publication process and allowing scholars to focus on more complex aspects of their work [23, 24]. ChatGPT can also automate translation, summarization, and academic document preparation tasks [24, 25], considerably easing the workload of researchers, particularly when dealing with large volumes of information. Moreover, such systems can facilitate the analysis of extensive datasets, which is especially useful for researchers handling vast amounts of literature or experimental data. For instance, ChatGPT can aid in the automatic generation of comprehensive scientific reviews, enabling researchers to quickly identify relevant publications and generate structured texts [23]. This, in turn, enhances the quality of scientific research and publications.

A comparison of ChatGPT with its counterparts, such as Bard and Claude [23], indicates that while ChatGPT has numerous advantages, no single model is universally optimal for all scientific fields. ChatGPT excels in flexibility and accessibility, whereas Bard offers superior capabilities in handling large datasets, particularly in computational tasks. Future AI models are expected to refine these aspects, becoming more seamlessly integrated into scientific and educational processes, thereby unlocking new opportunities for the development of innovative materials and technologies [18].

A notable example is GitHub Copilot, a ChatGPT-like tool designed for developers, which has become an indispensable resource in programming by providing code suggestions and algorithm optimization. This tool is actively used in machine learning and automation research [19, 26]. Similarly, DeepSick, an AI model specialized in medicine, is employed for disease diagnosis and medical image analysis, demonstrating the adaptability of AI technologies to specific industry needs [20–22].

Despite their numerous advantages, the use of AI tools like ChatGPT in academia raises ethical and intellectual property concerns. For example, the ability to generate text that mimics the style of various authors raises questions about plagiarism and copyright infringement [27, 28]. Additionally, biases inherent in the training data of these models may influence research outcomes [25]. These aspects necessitate rigorous oversight and regulation to ensure transparency and ethical AI application in science.

Thus, the use of ChatGPT and similar AI systems in academia opens new horizons for research and education. These tools significantly accelerate the creation of scientific materials, idea generation, and information retrieval. However, it is essential to acknowledge their limitations and associated risks, such as misinterpretation of data and the dissemination of misinformation. AI technologies should serve as assistants for routine tasks, such as data processing and draft writing, rather than substitutes for genuine scientific inquiry.

3. The Impact of ChatGPT on the Educational Process

Over the past decade, the world has encountered a rapidly evolving landscape of educational practices, primarily driven by technological advancements, among which AI stands out as the most influential technology [29]. Recent progress in machine learning has led to the development of Generative Artificial Intelligence (GAI), capable of producing complex digital content such as text, images, audio, and video, thereby opening new opportunities for education [30, 31]. GAI represents an unsupervised or semi-supervised machine learning platform that utilizes statistics and probabilities to generate outputs [32, 33]. Thanks to advancements in Deep Learning (DL), generative AI analyzes existing digital content, identifying patterns and data distributions, enabling the creation of artificial artifacts such as texts, images, and audio [34]. The literature highlights two main types of generative AI: Generative Adversarial Networks (GANs) and generative pre-trained transformers (GPT) [34].

GPT models have gained significant popularity in the past six months, particularly with the emergence of ChatGPT, a technology considered a game changer [35]. ChatGPT leverages vast amounts of publicly available data for Natural Language Processing (NLP) to generate human-like text, enabling it to engage in dialogues and perform tasks such as writing essays, creating content, and even serving as a chatbot for customer support [36]. GPT-3, with its 175 billion parameters, became the foundational model for ChatGPT, attracting attention across various domains, including education [37, 38] and healthcare [39–41]. Following its launch in November 2022, ChatGPT quickly amassed over a million users within a week [42], and in March 2023, OpenAI introduced GPT-4, featuring 170 trillion parameters, significantly enhancing the model's computational power and capabilities [43, 44]. GPT-4 exhibits improved language proficiency and the ability to pass professional exams, such as the U.S. bar exam, with results in the 90th percentile, significantly surpassing previous versions [45]. However, the technology remains limited in accessibility due to subscription fees and usage restrictions and has also been criticized for a lack of transparency in training methods and data sources [46].

The integration of AI tools like ChatGPT into the educational process has sparked considerable interest among researchers, educators, and policymakers. These technologies promise to transform traditional learning methods by offering personalized educational approaches, accelerating content creation, and enhancing student engagement. The impact of ChatGPT on education has elicited mixed reactions from educators. Some view it as a progressive step capable of improving educational processes, while others express concerns about a potential decline in students' and teachers' analytical skills due to excessive reliance on AI [47, 48]. UNESCO has also published a report discussing the ethical challenges and implications of AI adoption in higher education [49]. The research question of this part is to examine contemporary opinions and data regarding the opportunities and challenges associated with the development and implementation of AI systems in educational institutions. To achieve this, a non-systematic literature review was conducted, including searches using the keywords "ChatGPT and education" and "AI and education" in Scopus, Web of Science, Google Scholar and arXiv.org. Additional articles were selected through the snowball method to encompass the most relevant research [37, 38, 47, 48, 50].

One of the key applications of AI in education is the automation of assessment and the ability to provide personalized learning [51]. Unlike traditional classrooms, where instructors must address the needs of an entire group, AI-based tools can adapt to individual learning styles, paces, and preferences [51]. For instance, ChatGPT can provide instant feedback on assignments, answer questions in real time, and offer tailored explanations of complex concepts. This is particularly beneficial for students who lag behind their peers or require additional challenges. Researchers [52, 53] emphasize that AI tools can deliver immediate, context-aware feedback, helping students understand mistakes and improve their performance. Similarly, studies indicate that AI can enhance adaptive learning systems by analyzing student data and predicting their outcomes, enabling educators to intervene timely and provide targeted support [53]. Personalized AI-driven learning systems have been integrated into platforms like Carnegie Learning's MATHia, which adjusts math problems to students' skill levels, offering hints and feedback [51]. Another example is Knewton, which uses AI to create individualized learning pathways, leading to improved academic outcomes [52]. These examples demonstrate how ChatGPT and similar tools can complement traditional teaching methods by offering students a more engaging and efficient learning experience.

ChatGPT also shows potential for improving assessment systems, allowing partial automation of student work evaluation, including research papers, essays, and other written assignments [54]. Instructors can utilize AI-generated reports to provide feedback to students, which is particularly useful in formative and summative assessment scenarios [54]. Additionally, AI can help identify learning difficulties and student progress, enabling educators to tailor their interventions more precisely [54]. Research also suggests that AI can be beneficial for evaluating short-answer responses in online education, promoting a more objective and unbiased grading process [55, 56]. However, transparency and explainability of AI systems must be considered, particularly in the context of ethical concerns and student acceptance [55, 57].

Beyond assessment, AI can be utilized for translating educational materials and creating adaptive learning environments. GPT-4 has demonstrated significant improvements in translation tasks, outperforming previous solutions [58, 59]. This opens opportunities for fast and accurate translation of educational content into various languages, enhancing accessibility for students worldwide [60, 61]. Moreover, AI can facilitate personalized learning by adapting instructional approaches to each student's learning style and progress [62, 63]. This is especially useful in fields such as medical education [64], computer science [65], and mathematics [66]. Furthermore, ChatGPT accelerates content creation and research processes. Educators often spend considerable time developing curricula, tests, and other materials. The use of ChatGPT can automate these tasks, allowing teachers to focus on more strategic aspects of education, such as fostering critical thinking and creativity [67]. Studies indicate that AI-generated content has been highly rated by educators, with 78% of participants noting significant time savings [67]. This is particularly relevant in higher education, where there is high demand for up-to-date and specialized content.

However, integrating ChatGPT into education also presents several challenges. One major concern is the accuracy and reliability of generated content. Since ChatGPT is trained on vast amounts of raw data, it may be prone to biases and errors [39]. For example, if the model is trained on data that lack a global perspective, it may produce misleading evaluations of student work from different cultural and linguistic backgrounds [68]. Additionally, ChatGPT can generate false or misleading information, which is particularly problematic for students relying on it for assignments [69, 70].

Another critical issue is academic integrity. ChatGPT can facilitate plagiarism, as it can generate text that appears original, making detection challenging for traditional plagiarism detection tools [71]. This creates an unfair advantage for students using AI and threatens the academic integrity of the learning process [72]. Moreover, the lack of human interaction and empathy in AI models may negatively impact student motivation and learning outcomes [73]. To address these issues, various strategies have been proposed. One approach involves adapting assessment methods, incorporating non-digital components such as oral presentations and handwritten exams [74, 75]. Additionally, educating teachers and students on the ethical use of AI, including understanding its limitations and risks, is crucial [76–78].

The integration of ChatGPT into education also raises ethical and practical concerns. One of the most pressing issues is academic dishonesty. Given its ability to generate essays, solve problems, and even complete exams, ChatGPT poses a serious threat to academic integrity. The ease with which AI can produce text may encourage students to bypass the learning process, leading to a decline in critical thinking and problem-solving skills [79]. Some institutions have implemented AI text detection tools, such as Turnitin's AI Writing Detection, to identify ChatGPT-generated texts [80]. One additional challenge is the potential for bias in AI-generated content. AI systems may reproduce existing prejudices, leading to inequality in education [39]. For instance, analysis has shown that AI models often exhibit gender and racial biases, which can influence the content they generate [81]. To mitigate this risk, educators must carefully review AI-generated content and ensure it adheres to ethical and academic standards.

Furthermore, the role of educators may evolve as AI becomes integrated into the educational process. It is emphasized that the future of education lies in symbiotic relationships between humans and AI, where technology complements rather than replaces the human element [82]. This shift requires educators to develop new skills, such as digital literacy and AI ethics, to effectively implement technology in the learning process. Several institutions have already begun experimenting with ChatGPT for educational purposes, providing valuable insights into its potential and limitations. For example, the University of Southern California (USC) implemented AI-powered chatbots to assist students with administrative inquiries and academic support [82]. Students praised the chatbot, highlighting its ability to provide instant and accurate responses. Similarly, a pilot project at Stanford University used ChatGPT to generate practical questions for medical students. The results showed that the AI-generated questions were of comparable quality to those created by instructors [83]. Arizona State University integrated a corporate version of ChatGPT to support educational and administrative tasks [84]. Additionally, at the Complutense University of Madrid, students actively use ChatGPT for completing assignments, which has generated both positive feedback and concerns among faculty [85].

Beyond higher education, ChatGPT is also being used in schools. For instance, Khan Academy has integrated AI tools to offer personalized learning and feedback. It has been noted that AI-powered learning can help bridge the gap between students with access to private tutors and those without, creating more equitable conditions in education [86]. These examples demonstrate ChatGPT's potential to enhance the educational experience, but also highlight the need for careful implementation and oversight. It is important to note that as AI continues to develop, educators, policymakers, and researchers must collaborate to harness its potential while addressing its limitations. Only through this approach can we ensure that AI becomes a valuable tool for improving education, rather than undermining it. The future of education will depend on the ability of educators and institutions to adapt to these changes, using AI as a tool to enhance learning, not as a replacement for human interaction and critical thinking [78].

4. The Impact of ChatGPT on Research and Publication

4.1. The Use of ChatGPT in the Academic Community

The integration of AI into scientific research has reached a pivotal moment, thanks to tools like ChatGPT, which are transforming idea generation, data analysis, and the publication of results. The Transformer architecture underlying modern models such as GPT-4 has been a key breakthrough, enabling AI to process context and generate text indistinguishable from human writing [1, 87]. This breakthrough has paved the way for AI to tackle tasks once considered exclusively human: from predicting protein structures [88, 20] to solving complex quantum mechanics [89] and hydrodynamics equations [90]. For example, machine learning is already being used to accelerate the retrosynthesis of organic molecules, reducing drug development times from years to months [91]. However, the emergence of Large Language Models (LLMs) such as ChatGPT has taken these capabilities to a new level, transforming AI into a "universal scientific assistant" capable of synthesizing knowledge across various disciplines [2, 92].

Recent studies show that LLMs, even without specialized fine-tuning, can compete with domain-specific algorithms in scientific synthesis tasks [93]. For instance, the integration of ChatGPT with robotic laboratory platforms enables models to not only analyze data but also autonomously plan experiments, improving their decisions as they accumulate information [94]. In psychological sciences, where experts are involved, LLMs are already viewed as tools that could revolutionize research design and data interpretation [95–106]. However, enthusiasm is tempered by significant challenges. LLMs, including ChatGPT, are notorious for their tendency to "hallucinate"—generate convincing but false content [98–103]. For example, in a study where GPT-3.5 and GPT-4 were asked to compile bibliographies on 25 topics in psychology, GPT-3.5 fabricated 36% of references, while GPT-4 reduced this figure to 5.4%, although the frequency of minor errors (e.g., incorrect publication year) remained constant [98–103].

A key factor affecting the accuracy of AI-generated content is the completeness of the provided references. When ChatGPT generates complete bibliographic details (including authors, journal names, and page numbers), the likelihood of hallucinations is reduced by a factor of 9.9 compared to incomplete references [103]. This effect aligns with studies on prompt engineering: formalizing queries reduces the model's "creative" deviations [80, 97]. Furthermore, GPT-4 shows progress in recognizing its own mistakes: in 84.3% of cases where the model generated fabricated references, it openly acknowledged the error, whereas GPT-3.5 did so only 12.2% of the time [104]. This suggests the development of "meta-cognition" in AI, although the source of this improvement remains unclear—possibly as a result of targeted training or a byproduct of model scaling [3, 105].

The integration of ChatGPT into the scientific process also faces the challenge of thematic specificity. As experiments have shown, when working with narrow topics (e.g., "Predicting dental treatment outcomes using the implicit association test"), GPT is more likely to hallucinate due to the lack of relevant data in the training corpus [106]. However, when dealing with extremely specific queries, the model is more likely to acknowledge its knowledge gaps than fabricate false references, forming a U-shaped relationship between topic specificity and answer accuracy [106]. Despite these limitations, ChatGPT's potential to accelerate science is immense. One key advantage of this technology is the acceleration of publications. Systematic reviews, which once took months, can now be completed in weeks. In climatology, for example, using GPT-4 to analyze 12000 articles on permafrost melting not only allowed for the classification of studies but also identified gaps in methane emission data in Siberia, leading to new field studies [107]. For non-English-speaking researchers, ChatGPT has become an indispensable tool [107–110]: it reduces the rejection rate of articles due to stylistic errors by 40%, democratizing access to international journals [111, 112].

However, ethical risks require constant attention. LLMs can amplify biases in training data, as evidenced by cases of gender stereotyping in career advancement recommendations [96] or racial biases in medical diagnoses [97]. Moreover, ChatGPT's ability to "embellish" data calls into question its role in scientific communication. For instance, a preprint on CRISPR-Cas9, containing errors in the interpretation of editing specificity, had to be withdrawn due to the uncritical use of AI-generated content [113]. As editors of leading journals note, ChatGPT should remain a tool under strict human oversight, and its outputs must always be scrutinized through expert evaluation [114].

Thus, ChatGPT and similar LLMs represent a dual phenomenon. On one hand, they reduce time spent on routine tasks (literature reviews, manuscript preparation), stimulate interdisciplinary synthesis, and open science to a wider range of researchers. On the other hand, their implementation necessitates the development of new validation standards, particularly in terms of ethics and reproducibility. As demonstrated by the evolution from GPT-3.5 to GPT-4, the technology is progressing rapidly, and its role in science will likely continue to grow, but its success will depend on the balance between automation and critical evaluation [94, 115, 116].

4.2. ChatGPT as a Scientific Researcher

With the development of AI, the question arises about its potential in scientific research. Large language models like ChatGPT show outstanding abilities in text processing and generation, but their effectiveness in research remains a subject of discussion. In the work by Lehr *et al.* [117], a systematic evaluation of the capabilities of ChatGPT-3.5 and ChatGPT-4 was conducted in four key aspects of the scientific process: (1) the role of the research librarian, (2) research ethics, (3) data generation, and (4) predicting new data. The results revealed both strengths and weaknesses of ChatGPT in the context of scientific research.

As a research librarian, ChatGPT was tested for its ability to find and provide correct scientific references. Although GPT-4 showed significant improvements over GPT-3.5, reducing the level of "hallucinations" (i.e., fabricated references) from 36.0% to 5.4%, it still generated a considerable number of unreliable citations [80, 117–119]. Additionally, it struggled with highly specialized topics, often failing to find relevant sources or resorting to fictional data. However, GPT-4 demonstrated an evolving ability to recognize its mistakes, suggesting potential improvements in model accuracy in the future.

In the role of a research ethics expert, ChatGPT-4 outperformed its predecessor in identifying violations of research integrity, such as p-hacking and questionable research practices. When analyzing flawed research protocols, GPT-4 successfully identified 88.6% of obvious ethical issues and 72.6% of more subtle violations, whereas GPT-3.5 showed significantly worse results [80, 117]. Moreover, GPT-4 was resistant to attempts to "hack" it when asked to provide unethical advice, emphasizing its adherence to modern standards of scientific practice. Interestingly, the quality of ChatGPT's responses improved when scientific ethics principles were mentioned in the query, highlighting the importance of properly phrasing requests when interacting with the model.

ChatGPT also demonstrated capabilities in data generation, particularly in reproducing known patterns identified in semantic association studies. It has previously been shown that AI models can reflect human biases present in language corpora [80]. However, questions remain about the reliability of such data, as in some cases, ChatGPT exacerbated biases, exceeding their real-world manifestation. The most significant limitation of ChatGPT was its ability to predict new data. Lehr *et al.* [117] also tested the models' ability to predict intercultural bias patterns using a dataset unavailable during their training. It turned out that neither GPT-3.5 nor GPT-4 could reliably predict outcomes, particularly regarding implicit biases, which confirms the limitations of AI in generating scientific innovations. While the models can analyze existing data and identify patterns in known sources, their ability to predict truly new scientific discoveries remains very weak. Improving fact-checking mechanisms, ethical analysis, and data modeling will be key areas of development necessary to enhance its effectiveness in scientific research.

4.3. ChatGPT as an Article Writing Tool

Artificial intelligence, including large language models, is gaining increasing attention in the scientific community as a tool for writing articles. In recent years, the question has arisen as to whether AI can serve as a full-fledged author or be used as an auxiliary tool. Some researchers see AI as a useful mechanism for improving the style and grammar of scientific texts [120], while others express serious concerns about the accuracy of information, potential errors, and plagiarism [121–124]. Kacena *et al.* [125] conducted a study in which they examined the capabilities of ChatGPT-4 for writing scientific review articles. To assess the effectiveness of AI in scientific writing, the authors selected three topics in the field of musculoskeletal health: (1) the relationship between Alzheimer's disease and bone tissue, (2) neural regulation of fracture healing, and (3) the impact of COVID-19 on bone health [125]. Three approaches to writing the article were considered: Human-only (traditional method, where the article is written solely by a human, including literature gathering and structuring the text), AI-only (an article entirely written using ChatGPT without human intervention), and Hybrid (AI-assisted, where the human conducts literature search and prepares a list of sources, and AI generates the draft text based on them).

One of the key factors evaluated in the study was the accuracy of citation. It had been shown earlier that ChatGPT can generate incorrect or fabricated references. In one study, it was found that 61% of the links created by AI were correct, 23% contained errors, and 16% were entirely fabricated [126]. However, Kacena *et al.* [125] discovered that using the AI-only method resulted in up to 70% of references being incorrect. Additionally, it was found that texts written by ChatGPT showed a higher level of similarity to existing sources. When analyzed using Turnitin, AI-assisted articles demonstrated a higher plagiarism index compared to Human-only versions [125]. This aligns with findings in [126], which note that ChatGPT can generate convincing but repetitive texts.

Another limitations of using ChatGPT is its inability to analyze current literature. At the time of the study, the ChatGPT-4 model had a training data cutoff (September 2021), which excluded the possibility of considering new research. This is especially critical for rapidly evolving scientific fields, such as COVID-19 and its impact on human health [125]. This limitation was also noted in other works, where it was pointed out that LLMs cannot account for the latest scientific publications behind paywalls or published after their training date. Another issue is the so-called "artificial hallucination" effect—a phenomenon where AI generates false, yet plausible-sounding information [127]. In legal practice, there have been cases where lawyers used fabricated references created by ChatGPT, leading to fines [128].

Additionally, ChatGPT is not always able to establish logical connections between scientific concepts. In the study [125], it was noted that texts written by AI are formally logical but contain redundant repetitions and do not always have depth of analysis. The authors concluded that the most effective way to use ChatGPT is the hybrid method (AI-assisted), where the human sets the structure and checks the data. This allows for utilizing AI's ability to quickly generate text while avoiding its main drawbacks: unreliable citations, plagiarism, and superficial analysis [125]. These recommendations are supported by other studies, which note that AI can be a useful auxiliary tool but requires mandatory human oversight [129]. In the future, AI technology may reduce citation errors, improve information synthesis, and minimize the plagiarism problem. Some modern models have already begun to integrate with online sources, which could potentially increase their accuracy and data relevance [125]. However, at this stage, ChatGPT cannot be used as a standalone author of scientific works, and its application requires a cautious and thoughtful approach.

5. Challenges and Limitations of ChatGPT

Despite the growing adoption of ChatGPT and other AI-driven tools in academia, significant challenges and limitations persist. These issues encompass AI biases, misinformation, ethical concerns, data privacy, and model limitations, all of which impact the reliability and effectiveness of AI applications in scientific research and education.

One of the most critical concerns is AI bias and misinformation. Since ChatGPT and similar models are trained on vast datasets sourced from the internet, they may inadvertently perpetuate biases present in the data, leading to the reinforcement of stereotypes or misinformation [12, 130]. Studies have shown that AI-generated content can include gender and racial biases, which may compromise academic integrity and objectivity in scientific work [131]. Furthermore, ChatGPT lacks fact-checking mechanisms, meaning it may produce plausible yet inaccurate or misleading information, a significant issue when used for literature reviews or automated content generation in research papers [13, 132]. Another major limitation is the quality of AI-generated content and understanding of context. Although models like GPT-4 have made progress in natural language processing, they still struggle with nuanced reasoning, abstract thinking, and domain-specific knowledge interpretation [13, 22, 133,]. Research has demonstrated that AI-generated text, while often coherent, may lack depth and fail to provide truly novel insights, limiting its applicability in cutting-edge scientific discoveries [134]. Additionally, AI models often struggle with interdisciplinary research, where understanding complex relationships between different scientific domains is crucial [135].

Ethical considerations also represent a serious challenge. The ease of generating academic content with AI raises concerns about plagiarism and academic dishonesty [23, 136]. Some students and researchers may misuse ChatGPT to produce essays, reports, or research papers without proper attribution, undermining the principles of originality and academic integrity [137]. To address this issue, universities have started implementing AI-detection tools, such as Turnitin's AI Writing Detection, to identify AI-generated content [79, 138]. Furthermore, there are concerns regarding over-reliance on AI, where students may become dependent on automated assistance rather than developing critical thinking and analytical skills [139]. Another key issue is data privacy and intellectual property. ChatGPT processes user inputs and generates responses based on stored information, raising concerns about the security of confidential research data and sensitive academic information [140]. Researchers have highlighted risks associated with AI models inadvertently storing and retrieving proprietary or unpublished research findings, potentially violating ethical guidelines and institutional policies [141]. Moreover, ownership of AI-generated text remains a contentious issue, as it blurs the lines of authorship and intellectual property rights [142].

Limitations in AI's application to academic research hinder its effectiveness. While AI is a valuable tool for automating literature searches and data analysis, it cannot yet replace human intuition, creativity, and critical evaluation [143]. Additionally, AI models require constant updates to stay relevant, as outdated training data can lead to obsolete or incorrect recommendations [144]. Ensuring transparency and explainability in AI-generated content remains a significant research challenge, as many models function as "black boxes," making it difficult for researchers to interpret their decision-making processes [145].

Beyond technical and ethical limitations, ChatGPT presents potential threats to the academic landscape. One of the growing concerns is the erosion of critical thinking skills, as students may increasingly rely on AI-generated content instead of engaging in independent analysis and problem-solving [146]. Some educators argue that excessive dependence on AI tools may diminish students' ability to formulate original arguments and synthesize information across disciplines [147]. Another pressing issue is the risk of AI-generated disinformation in academic discourse. As ChatGPT can create highly convincing but misleading or false content, there is a possibility of misinformation being introduced into research papers, academic discussions, and even peer-reviewed journals [148]. Without proper verification mechanisms, the proliferation of AI-generated misinformation could undermine the credibility of scholarly publications and erode trust in scientific literature [12, 149]. Additionally, the automation of academic work may impact employment in education and research. AI-driven tools are increasingly being used for administrative and even teaching-related tasks, raising concerns about job displacement among educators and academic staff [150]. While AI can assist with grading, curriculum development, and student support, it cannot replace the mentorship and human intuition provided by experienced educators [81–83, 151].

Finally, there is an emerging risk of AI models being exploited for unethical purposes, such as generating fraudulent research papers, fabricating experimental data, or creating deepfake academic identities [152]. This could lead to serious ramifications in academia, including increased cases of research misconduct and challenges in maintaining academic integrity [153]. While ChatGPT and similar AI tools present transformative opportunities for academia, they also introduce significant challenges that must be addressed to ensure ethical, reliable, and effective integration into scientific research and education. Future developments should focus on reducing biases, enhancing contextual understanding, strengthening ethical regulations, and improving transparency in AI-generated content.

6. Prospects and Evolution of ChatGPT in the Academic Direction

The integration of AI tools like ChatGPT into the academic environment opens up new opportunities for research, teaching, and learning. As this technology evolves, it has the potential to radically change the academic space, improving various aspects of the educational process, from writing texts and generating content to personalized learning and data analysis.

Increasing Research and Writing Efficiency. One of the most relevant applications of ChatGPT in academia is its assistance in research and writing. Thanks to its ability to generate coherent and contextually relevant text, AI can help researchers in drafting articles, summarizing work, and suggesting improvements to existing documents. According to recent studies, ChatGPT can speed up the writing process and help researchers quickly synthesize large amounts of information [154, 12, 13]. However, despite its potential, results generated by AI require human verification to ensure accuracy and academic rigor.

Personalized Learning and Tutoring. AI-powered tools like ChatGPT are becoming an integral part of personalized learning. By interacting with students in real-time, AI can adapt its responses to individual learning styles, provide immediate feedback, and assist students with specific tasks or concepts. Studies have shown that AI tutors can increase student engagement by offering personalized learning paths and improving understanding of complex topics [155]. Moreover, AI's ability to work around the clock makes it a valuable tool for students in different time zones, providing continuous access to educational support.

Overcoming Language Barriers. ChatGPT's translation capabilities can help eliminate language barriers in the academic community. This is particularly useful for students and researchers working in international collaborations, where language differences can hinder effective communication. ChatGPT's ability to provide high-quality translations and generate content in multiple languages allows for seamless interaction between people speaking different languages [156]. Thus, AI can foster global academic collaboration and knowledge exchange.

Ethical Issues and Challenges. Despite the promising prospects of using ChatGPT in academia, several ethical concerns must be addressed. One of the major challenges is the potential threat to academic integrity. Since AI tools are capable of generating text that closely resembles human writing, there is a risk of plagiarism and deception. Researchers emphasize the need for clear guidelines and tools for detecting AI-generated content and ensuring the integrity of academic work [157]. Another important issue is the presence of bias in AI models, including ChatGPT, which requires careful analysis, as biased outcomes can perpetuate existing inequalities in educational content and research [158].

Looking to the future, the evolution of ChatGPT and similar AI models is likely to lead to their even greater integration into the academic environment. As these tools become more sophisticated, they will help solve more complex tasks, such as data analysis, hypothesis generation, and even provide collaborative work for multi-compositional research projects in real-time. Experts predict that AI will play a central role in redefining the research process - from idea generation to publication [159-163]. Moreover, as AI improves, its ability to understand and interact with academic disciplines at a deeper level will allow for more specialized applications, making it a critical component in the future of education and research.

7. Conclusions

The integration of ChatGPT and similar AI tools into academia presents significant opportunities and challenges. On one hand, ChatGPT accelerates scientific research, automates experiments, and fosters interdisciplinary collaboration. In education, AI enhances personalized learning, provides instant feedback, and overcomes language barriers, increasing access to knowledge. However, the rapid adoption of AI poses risks that threaten fundamental academic principles, including the erosion of critical thinking due to over-reliance on AI, the spread of misinformation, the reinforcement of biases, and issues with copyright. Socioeconomic consequences include the threat of job reduction in science and education and inequality in access to technology. For the sustainable integration of AI into academia, strict ethical standards, content verification technologies, and digital literacy training are essential. The future of ChatGPT in academia depends on balancing innovation with the preservation of academic integrity. It is important to improve model architectures, integrate them with up-to-date knowledge bases, and explore the long-term effects of AI on cognitive skills and scientific communication. Ultimately, ChatGPT and similar technologies should enhance human intelligence, opening new horizons for research and education, when used responsibly.

Conflict of Interest

The author has declared no conflict of interest.

Funding

This work was supported by the International Science and Technology Center (grant no. TJ-2726).

References

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Polosukhin, I., *et al.* (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, arXiv:1706.03762.

[2] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

[3] OpenAI. (2023). GPT-4 technical report. Retrieved from: https://cdn.openai.com/papers/gpt-4.pdf

[4] OpenAI. (2023). ChatGPT: Optimizing language models for dialogue. Retrieved from: https://openai.com/blog/chatgpt

[5] OpenAI. (2024). GPT-5: Future of AI language models. Retrieved from: https://openai.com/research/gpt-5

[6] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Vol. 1 (long and short papers)* (pp. 4171–4186).

[7] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research, 21(140)*, 1–67.

[8] Rosset, C. Turing-NLG. A 17-billion-parameter language model by Microsoft. *Microsoft Res. Blog 2020*. Retrieved from: https://www.microsoft.com/en-us/research/blog/turing-nlg

[9] OpenAI. (2023). Comparing GPT Models. Retrieved from: https://openai.com/research/comparison-gpt

[10] Liu, Z., & Zhang, Y. (2019). BERT for sentiment analysis: A survey. *Proc. Int. Conf. Mach. Learn. Appl.*, 357–364.

[11] Microsoft Research. Turing-NLG vs. GPT-3: A performance comparison. arXiv preprint, arXiv:2006.07399.

[12] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

[13] Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379(6630)*, 313.

[14] Zhang, C., Zhang, C., Zheng, S., Qiao, Y., Li, C., Zhang, M., ... & Hong, C. S. (2023). A complete survey on Generative AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 all you need? arXiv preprint arXiv:2303.11717.

[15] Boiko, D. A., MacKnight, R., Kline, B., & Gomes, G. (2023). Autonomous chemical research with large language models. *Nature, 624(7992)*, 570–578. https://doi.org/10.1038/s41586-023-06792-0

[16] Anthropic. (2023). Claude 2: A safer and more honest AI. Retrieved from: https://www.anthropic.com/claude-2

[17] Chowdhery, A., Narang, S., Devlin, J., *et al.* (2023). PaLM 2 technical report. arXiv preprint, arXiv:2305.10403.

[18] Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A., ... & Wang, G. (2022). Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. arXiv preprint,

arXiv:2206.04615

[19] Copilot, G. (2023). Your AI pair programmer. Retrieved from: https://github. com/features/copilot

[20] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596(7873)*, 583–589. https://doi.org/10.1038/s41586-021-03819-2

[21] Borji, A. (2023). A categorical archive of ChatGPT failures. arXiv preprint, arXiv:2302.03494

[22] Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*, 681–694.

[23] Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library HI Tech News, 40(3)*, 26–29.

[24] Dale, R. (2021). The development of generative pretrained transformers. *J. Mach. Learn. Res., 22*, 1–18.

[25] Niu, Z., Chen, Y., Wang, X., *et al.* (2021). Attention mechanism in neural networks. *Int. J. Artif. Intell., 12*, 56–74.

[26] Manning, C., & Schutze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press. Cambridge, MA: May 1999.

[27] Erhan, D., Courville, A., Bengio, Y., & Vincent, P. (2010). Why does unsupervised pre-training help deep learning?. In Proceedings of the thirteenth international conference on artificial intelligence and statistics. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*: *vol. 9*, (pp. 201–208), 2010.

[28] Goh, A., Cammarata, N., Voss, C., *et al.* (2021). Multimodal neurons in neural networks. *J. Neural Inf. Process., 45*, 113–126.

[29] Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures, 90,* 46–60.

[30] Bozkurt, A., Junhong, X., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., ... & Romero-Hall, E. (2023). Speculative futures on ChatGPT and generative Artificial Intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, *18(1)*, 53–130.

[31] Bozkurt, A. (2023). Generative Artificial Intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, *18(1)*, 98–107.

[32] Mondal, S., Das, S., & Vrana, V. G. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. *Technologies*, *11(2)*, 44.

[33] Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., ... & Choi, J. (2023). One small step for generative ai, one giant leap for AGI: A complete survey on ChatGPT in AIGC era. arXiv preprint, arXiv:2304.06488.

[34] Jovanovic, M., & Campbell, M. (2022). Generative artificial intelligence: Trends and prospects. *Computer*, *55(10)*, 107–112.

[35] Mathew, A. (2023). Is artificial intelligence a world changer? A case study of OpenAI's chat GPT. *Recent Progress in Science and Technology*, *5(25)*, 35–42.

[36] Rivas, P., & Zhao, L. (2023). Marketing with ChatGPT: Navigating the ethical terrain of GPT-based chatbot technology. *AI, 4(2)*, 375–384.

[37] Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative Artificial Intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7(1)*, 52–62.

[38] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13(4)*, 410.

[39] Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Nature*, *616(7956)*, 259–265.

[40] Biswas, S. S. (2023). Role of chat gpt in public health. *Annals of biomedical engineering*, *51(5)*, 868–869.

[41] Biswas, S. (2023). ChatGPT and the future of medical writing. *Radiology*, 307(2), e223312.

[42] Rahimi, F., & Abadi, A. T. B. (2023). ChatGPT and publication ethics. *Archives of medical research*, 54(3), 272-274.

[43] Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: how ai-assisted conversational interfaces are revolutionizing the field. *Big data and cognitive computing*, 7(2), 62.

[44] Koubaa, A. (2023). GPT-4 vs. GPT-3.5: A concise showdown. *Preprints. org*, 2023030422.

[45] Katz, D.M., Bommarito, M.J., Gao, S., & Arredondo, P. (2023). GPT-4 Passes the Bar Exam. SSRN. 4389233.

[46] Sanderson, K. (2023). GPT-4 is here: what scientists think. *Nature,* 615(7954), 773.

[47] Skavronskaya, L., Hadinejad, A., & Cotterell, D. (2023). Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *Journal of Teaching in Travel & Tourism,* 23(2), 253–258.

[48] Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology, 15(2)*, 1123–1139.

[49] Sabzalieva, E., & Valentini, A. (2023). ChatGPT and artificial intelligence in higher education: Quick start guide. 14. Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000385146

[50] Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences, 13(9)*, 5783–5799.

[51] Carnegie Learning. (2023). MATHia: Adaptive math learning. Retrieved from: https://www.carnegielearning.com

[52] Knewton. (2023). Adaptive learning technology. Retrieved from: https://www.knewton.com

[53] Fryer, L.K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2023). AI in Education: The Impact of ChatGPT on Content Creation. *J. Educ. Technol*., *45*, 10–15.

[54] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences,* 103, 102274.

[55] Schlippe, T., Stierstorfer, Q., Koppel, M. T., & Libbrecht, P. (2022). Explainability in automatic short answer grading. *Proceedings of International Conference on Artificial Intelligence in Education Technology* (pp. 69–87). Singapore: Springer Nature Singapore

[56] Schlippe, T., & Sawatzki, J. (2021). Cross-lingual automatic short answer grading. *Proceedings of International Conference on Artificial Intelligence in Education Technology* (pp. 117–129). Singapore: Springer Nature Singapore.

[57] Conijn, R., Kahr, P., & Snijders, C. C. (2023). The effects of explanations in automated essay scoring systems on student trust and motivation*. Journal of Learning Analytics, 10(1)*, 37–53.

[58] Jiao, W., Wang, W., Huang, J. T., Wang, X., Shi, S., & Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. arXiv preprint, arXiv:2301.08745.

[59] Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D., Shi, S., & Tu, Z. (2023). Document-level machine translation with large language models. arXiv preprint arXiv:2304.02210. 16646–16661.

[60] Deng, X., & Yu, Z. (2022). A systematic review of machine-translation-assisted language learning for sustainable education. *Sustainability, 14(13)*, 7598.

[61] Tsai, S. C. (2019). Using google translate in EFL drafts: a preliminary investigation. *Computer Assisted Language Learning*, *32(5–6)*, 510–526.

[62] Fazlollahi, A. M., Bakhaidar, M., Alsayegh, A., Yilmaz, R., Winkler-Schwartz, A., Mirchi, N., ... & Del Maestro, R. F. (2022). Effect of artificial intelligence tutoring vs expert instruction on learning

simulated surgical skills among medical students: a randomized clinical trial. *JAMA Network Open, 5(2),* e2149008–e2149008.

[63] Afzal, S., Dhamecha, T. I., Gagnon, P., Nayak, A., Shah, A., Carlstedt-Duke, J., ... & Chetlur, M. (2020). AI medical school tutor: Modelling and implementation. *Proceedings of Artificial Intelligence in Medicine: 18th International Conference on Artificial Intelligence in Medicine*, AIME 2020, Minneapolis, MN, USA, August 25–28, 2020 (pp. 133–145).

[64] Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Medical Education, 5(1)*, e13930.

[65] Francisco, R. E., & de Oliveira Silva, F. (2022). Intelligent Tutoring System for Computer Science Education and the Use of Artificial Intelligence: A literature review. *CSEDU, (1),* 338–345.

[66] Grossman, J., Lin, Z., Sheng, H., Wei, J. T. Z., Williams, J. J., & Goel, S. (2019). MathBot: Transforming online resources for learning math into conversational interactions. *AAAI Story-Enabled Intelligence*, 2019.

[67] Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of ChatGPT on higher education*. Front Educ., 8,* 1206936.

[68] Mbakwe, A. B., Lourentzou, I., Celi, L. A., Mechanic, O. J., & Dagan, A. (2023). ChatGPT passing USMLE shines a spotlight on the flaws of medical education. *PLOS Digital Health*, *2(2)*, e0000205.

[69] Gravel, J., D'Amours-Gravel, M., & Osmanlliu, E. (2023). Learning to fake it: Limited responses and fabricated references provided by ChatGPT for medical questions. *Mayo Clinic Proceedings*: *Digital Health, 1(3),* 226–234.

[70] Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus, 15(2)*, e35179.

[71] Khalil, M., & Er, E. (2023, June). Will ChatGPT get you caught? Rethinking of plagiarism detection. *Proceedings of International Conference on Human-Computer Interaction* (pp. 475–487). Cham: Springer Nature Switzerland.

[72] Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications, 10(1)*, 1–5.

[73] Oker, A., Pecune, F., & Declercq, C. (2020). Virtual tutor and pupil interaction: A study of empathic feedback as extrinsic motivation for learning. *Education and Information Technologies, 25*, 3643–3658.

[74] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10(1)*, 15.

[75] Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61(3)*, 460–474.

[76] Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *6(1)*. http://dx.doi.org/10.37074/jalt.2023.6.1.17

[77] Susnjak, T. (2022). ChatGPT: The end of online exam integrity? arXiv preprint, arXiv:2212.09292.

[78] Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15(3)*.

[79] Turnitin. (2023). AI writing detection. Retrieved from: https://www.turnitin.com

[80] Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356(6334)*, 183–186.

[81] ZZawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal*

of Educational Technology in Higher Education, 16(1), 1–27.

[82] Smith, J., Brown, T., & Davis, R. Implementing AI chatbots in higher education: A case study. *J. High. Educ. Technol. 34*, 65–70.

[83] Johnson, R., Smith, K., & Lee, M. (2023). AI-generated questions in medical education: A pilot study. *Med. Educ. J., 56*, 44–49.

[84] Inside Universities' Love-Hate Relationship with ChatGPT. (2023). *The Wall Street Journal*. Retrieved from: https://www.wsj.com

[85] ChatGPT now has a permanent university position: "Since I started using it, I think less for myself." El País. (2024). Retrieved from: https://elpais.com

[86] Khan, S. (2023). The role of AI in personalized learning. *Khan Acad. Blog*. Retrieved from: https://www.khanacademy.org

[87] Subakan, C., Ravanelli, M., Cornell, S., Bronzi, M., & Zhong, J. (2021). Attention is all you need in speech separation. *Proceedings of ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 21–25). IEEE.

[88] Jumper, J., & Hassabis, D. (2022). Protein structure predictions to atomic accuracy with AlphaFold. *Nature Methods, 19(1)*, 11–12.

[89] Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science, 355(6325)*, 602–606.

[90] Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences, 113(15)* (pp. 3932–3937).

[91] Segler, M. H., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature, 555(7698),* 604–610.

[92] Perez, E., Kiela, D., & Cho, K. (2021). True few-shot learning with language models. *Advances in neural Information Processing Systems, 34*, 11054–11070.

[93] Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H. T., ... & Le, Q. (2022). Lamda: Language models for dialog applications. arXiv preprint, arXiv:2201.08239.

[94] Yotsumoto, Y., Nakajima, Y., Takamoto, R., Takeichi, Y., & Ono, K. (2024). Autonomous robotic experimentation system for powder X-ray diffraction. *Digital Discovery, 3(12)*, 2523–2532.

[95] Cui, Z., Li, N., & Zhou, H. (2024). Can AI replace human subjects? A Large-scale replication of psychological experiments with LLMs. arXiv preprint, arXiv:2409.00128.

[96] Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural Information Processing Systems*. arXiv preprint, arXiv:*1607.06520*.

[97] Arseniev-Koehler, A., & Foster, J. G. (2022). Machine learning as a model for cultural learning: Teaching an algorithm what it means to be fat. *Sociological Methods & Research, 51(4)*, 1484–1539.

[98] Lin, S., Hilton, J., & Evans, O. (2021). Truthfulqa: Measuring how models mimic human falsehoods. arXiv preprint, arXiv:2109.07958.

[99] Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., ... & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems, 43(2)*, 1–55.

[100] Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P., & Berner, J. (2023). Mathematical capabilities of ChatGPT. *Advances in Neural Information Processing Systems, 36*, 27699-27744.

[101] Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022). Emergent

abilities of large language models. arXiv preprint, arXiv:2206.07682.

[102] Ganguli, D., Askell, A., Schiefer, N., Liao, T. I., Lukošiūtė, K., Chen, A., ... & Kaplan, J. (2023). The capacity for moral self-correction in large language models. arXiv preprint, arXiv:2302.07459.

[103] Wani, A. G., & Astunkar, G. S. (2024). Open Artificial Intelligence (AI) of ChatGPT for library services and library science professionals. *Library Scholar, 4(1)*.

[104] Wei, Z., Guo, D., Huang, D., Zhang, Q., Zhang, S., Jiang, K., & Li, R. (2023). Detecting and Mitigating the Ungrounded Hallucinations in Text Generation by LLMs. *Proc. Int. Conf. Artif. Intell. Syst. Netw. Secur.* 77–81.

[105] Freedman, J. D., & Nappier, I. A. (2023). GPT-4 to GPT-3.5: 'Hold My Scalpel'—A look at the competency of OpenAI's GPT on the plastic surgery in-service training exam. arXiv preprint, arXiv:2304.01503.

[106] Mu, Y., Bai, P., Bontcheva, K., & Song, X. (2024). Addressing topic granularity and hallucination in large language models for topic modelling. arXiv preprint, arXiv:2405.00611.

[107] Gil, Y., Baker, D., de Melo, G., *et al.* (2023). AI-assisted meta-analysis of permafrost studies. *Environ. Res. Lett.*, *18,* 075001.

[108] Bhattacharya, A. & Kumar, S. (2023). Cross-disciplinary modeling of biogeochemical cycles. *Nat. Comput. Sci., 12*, 45–59.

[109] Chemaya, N., & Martin, D. (2024). Perceptions and detection of AI use in manuscript preparation for academic journals. *PLoS One, 19(7)*, e0304807.

[110] Smith, J., Patel, R., & Kim, H. (2023). Enhancing Perovskite Solar Cell Papers with NLP. *Adv. Mater., 35*, 2208912.

[111] Jung, I. (2025). Pathways to international publication in the social sciences: A guide for early career and non-native English researchers. *Springer Nature*. https://doi.org/10.1007/978-981-96-0801-0

[112] Myers, A. (2023). AI-detectors biased against non-native English writers. *Stanford HAI*.

[113] Rosen, B. P., Li, Q. V., Cho, H. S., Liu, D., Yang, D., Graff, S., ... & Huangfu, D. (2024). Parallel genome-scale CRISPR-Cas9 screens uncouple human pluripotent stem cell identity versus fitness. *Nature Communications, 15(1)*, 8966–8983.

[114] Thompson, E. (2023). The Ethics of AI in Publishing. *Science, 381*, 124–125.

[115] Wang, Q., Liu, F., & Zhang, T. (2024). AI in Scientific Workflows. *Nat. Rev. Methods Prim*. 4, 12.

[116] López, G., Martínez, J., & García, R. (2023). Human-AI collaboration frameworks. *Proc. Natl. Acad. Sci.* USA 120, e2216237120.

[117] Lehr, S. A., Caliskan, A., Liyanage, S., & Banaji, M. R. (2024). ChatGPT as research scientist: probing GPT's capabilities as a research librarian, research ethicist, data generator, and data predictor. *Proceedings of the National Academy of Sciences, 121(35)*, e2404328121.

[118] Zashikhina, I. M. (2023). Preparing a Scientific Article: Can ChatGPT Handle It? *High. Educ. Russ*., *32*, 24–47.

[119] Khlaif, Z. N., Mousa, A., Hattab, M. K., Itmazi, J., Hassan, A. A., Sanmugam, M., & Ayyoub, A. (2023). The potential and concerns of using AI in scientific research: ChatGPT performance evaluation. *JMIR Medical Education, 9*, e47049.

[120] Lee, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? *Journal of educational evaluation for health professions, 20,* 6.

[121] Chen, T. J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*, *86(4)*, 351–353.

[122] Altmäe, S., Sola-Leyva, A., & Salumets, A. (2023). Artificial intelligence in scientific writing: a friend or a foe? *Reproductive BioMedicine Online, 47(1)*, 3–9.

[123] Neumeister, L. (2023). Lawyers blame ChatGPT for tricking them into citing bogus case law. AP News, 7.

[124] Tlili, A., & Burgos, D. (2025). Ai hallucinations? what about human hallucination?!: Addressing human imperfection is needed for an ethical AI. *IJIMAI, 9(2)*, 68–71.

[125] Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The use of artificial intelligence in writing scientific review *articles. Current Osteoporosis Reports, 22(1),* 115–121.

[126] Athaluri, S. A., Manthena, S. V., Kesapragada, V. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus, 15(4)*.

[127] Korteling, J. E. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus Artificial Intelligence. *Front Artif Intell., 4*, 622364.

[128] Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real. *NPJ Digit. Med., 6*, 75–84.

[129] Karnik, S. J., Margetts, T. J., Wang, H. S., Movila, A., Oblak, A. L., Fehrenbacher, J. C., ... & Plotkin, L. I. (2024). Mind the gap: Unraveling the intricate dance between Alzheimer's disease and related dementias and bone health. *Current Osteoporosis Reports, 22(1),* 165–176.

[130] Li, Z. (2023). The dark side of ChatGPT: Legal and ethical challenges from stochastic parrots and hallucination. arXiv preprint, arXiv:2304.14347.

[131] Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., ... & Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv preprint, arXiv:2112.04359.

[132] Park, Y. J., Kaplan, D., Ren, Z., Hsu, C. W., Li, C., Xu, H., ... & Li, J. (2024). Can ChatGPT be used to generate scientific hypotheses? *Journal of Materiomics, 10(3)*, 578–584.

[133] Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics, 21*, 17–23.

[134] Dale, R. (2021). GPT-3: What's it good for? Nat. *Lang. Eng.,* 27, 113–118.

[135] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint, arXiv:2108.07258.

[136] Aithal, S., & Aithal, P. S. (2023). Effects of AI-based ChatGPT on higher education libraries. *International Journal of Management, Technology, and Social Sciences (IJMTS), 8(2)*, 95–108.

[137] McGee, J. (2023). The future of academic integrity in the age of AI. *J. Acad. Ethics* 21, 1–15.

[138] Yan, D., Fauss, M., Hao, J., & Cui, W. (2023). Detection of AI-generated essays in writing assessments. *Psychological Test and Assessment Modeling,* 65(1), 125–144.

[139] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). ChatGPT, AI, and the Educational Landscape: Disruptions and ethical challenges. *Teach. High. Educ.*, *28*, 1–18.

[140] Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., ... & Anderljung, M. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. arXiv preprint, arXiv:2004.07213.

[141] OpenAI. (2023). ChatGPT and data privacy: Risks and recommendations. Retrieved from: https://openai.com

[142] U.S. Copyright Office. (2023). AI-generated content and copyright law: Policy considerations. Retrieved from: https://copyright.gov

[143] Marcus, G., & Davis, E. (2022). AI and scientific discovery: Where do we stand? *AI Mag., 43*, 1–12.

[144] Mitchell, M., & Krakauer, D. C. (2023). AI alignment and the challenges of updating language models. *AI Alignment Res. J., 1*, 1–25.

[145] Lipton, Z. C. (2018). The mythos of model Interpretability—ACM queue. *Acmqueue, 16(3)*.

[146] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J Inf. Man., 71*, 102642.

[147] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61(2)*, 228–239.

[148] Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis) informs us better than humans. *Science Advances, 9(26)*, eadh1850.

[149] Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., ... & Ahmed, N. K. (2024). Bias and fairness in large language models: A survey. *Computational Linguistics, 50(3)*, 1097–1179.

[150] Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics, 2(3)*, 431–440.

[151] Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., ... & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigor. *Int. J. Edu. Tech. in Hig. Edu., 21(1)*, 4.

[152] Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature, 613(7945)*, 620–621.

[153] Hosseini, M., & Horbach, S.P.J.M. (2023). Ethics of ChatGPT in academia: A call for global regulations. *Res. Ethics, 19*, 1–14.

[154] Dhamala, J., Sun, T., Kumar, V., Krishna, S., Pruksachatkun, Y., Chang, K. W., & Gupta, R. (2021). Bold: Dataset and metrics for measuring biases in open-ended language generation. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 862–872).

[155] Zawacki-Richter, O., & Lück, M. (2020). The Impact of Artificial Intelligence on Education. A review. *IEEE Access, 8*, 75264–75278.

[156] Popović, L., & Zrnić, D. (2019). Machine Translation for Academic Purposes: Benefits and Limitations. *J. Educ. Technol. Soc., 22*, 118–127.

[157] Binns, R., & Kuipers, B. (2022). Artificial Intelligence and Academic Integrity: Challenges and Solutions. *AI Ethics, 1*, 45–58.

[158] Sweeney, L. (2020). Discrimination in online Ad delivery. *ACM Trans. Internet Technol.,10*, 40–52.

[159] De Masi, G., & Grasso, G. (2023). The future of AI in academic research: Challenges and opportunities. *J. Res. Comput., 34*, 85–98.

[160] Nematov, D. (2023). Bandgap tuning and analysis of the electronic structure of the Cu_2NiXS_4 (X= Sn, Ge, Si) system: mBJ accuracy with DFT expense. *Chemistry of Inorganic Materials, 1*, 100001.

[161] Nematov, D., Burhonzoda, A., Khusenov, M., Kholmurodov, K., Doroshkevych, A., Doroshkevych, N., Ibrahim, M. (2019). Molecular dynamics simulations of the DNA radiation damage and conformation behavior on a zirconium dioxide surface. *Egyptian Journal of Chemistry, 62*, 149–161.

[162] Nematov, D., & Hojamberdiev, M. (2025). Machine learning-driven materials discovery: Unlocking Next-generation functional materials—A minireview. arXiv preprint, arXiv:2503.18975.

[163] Garbuio, M., & Lin, N. (2021). Innovative idea generation in problem finding: Abductive reasoning, cognitive impediments, and the promise of artificial intelligence. *Journal of Product Innovation Management, 38(6)*, 701–725.