LITERATURE REVIEW ON GENERATIVE ARTIFICIAL INTELLIGENCE IN PHYSICS AND SCIENCE: CHALLENGES, POTENTIAL, AND IMPACTS ON THE LEARNING PROCESS[1]

https://doi.org/10.63330/aurumpub.057-011

Alex Linhares Bezerra[2] and Valdenildo Pedro da Silva[3]

Abstract

Generative Artificial Intelligence (GenAI) has become established as an emerging technology in education, particularly in physics and science education since 2022, prompting debate about its pedagogical potential and limitations. The relevance of this topic stems from the sensitive and complex nature of incorporating GenAI into fields that require conceptual rigor, critical thinking, and strong teacher mediation. The general objective of this study is to analyze how the academic literature addresses the use of GenAI in the physics and science teaching and learning process between 2022 and 2025, considering its impacts, challenges, and possibilities. Methodologically, a Systematic Literature Review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol. A search of the Scopus database resulted in the analysis of seven studies. The findings indicate that GenAI has the potential to support active learning, problem-solving, and methodological innovation, although it continues to face conceptual limitations, risks of superficial use, and technological dependence. The study concludes that the integration of GenAI requires critical pedagogical strategies, appropriate teacher training, and ethical use in order to enhance learning without compromising the quality of education.

---

[1] This article is the English-language version of a study originally published in the proceedings of the **2nd Seminar on Artificial Intelligence and Its Applications in Academic Research (II Seminário sobre Inteligência Artificial e suas Aplicações na Pesquisa Acadêmica—SIA)**, under the title **"Generative Artificial Intelligence in Physics and Science Education: Challenges, Potential, and Impacts on the Learning Process—A Systematic Literature Review,"** DOI **10.5281/zenodo.21269158**. This publication is a revised English translation adapted to the format of an international book chapter while preserving the authorship and scientific content of the original work.

[2] Bachelor's Degree in Physics Education – Instituto Federal do Rio Grande do Norte, Campus Santa Cruz-RN
E-mail: alexlinhares940@gmail.com
[3] Faculty Advisor– Instituto Federal do Rio Grande do Norte, Campus Santa Cruz-RN
E-mail: valdenildopedrosilva16@gmail.com

Keywords: Active Learning, Generative Artificial Intelligence, Physics Education, Science Education, Systematic Literature Review.

INTRODUCTION

The rise of Generative Artificial Intelligence (GenAI), driven by innovations that have emerged since 2022, has brought about significant changes across several sectors, particularly education. Its ability to generate texts, figures, images, sounds, voices, films, and animations has attracted growing interest in physics and science education, highlighting its potential to support innovative pedagogical practices, improve assessment processes, and strengthen human–machine collaboration. However, this integration also reveals important limitations, including conceptual errors, inadequate interpretations of visual representations, and the need for teachers and students to acquire new competencies. Accordingly, this study focuses on a critical analysis of the application of GenAI in physics and science education, considering the challenges associated with its adoption in a field characterized by traditional methodologies.

The relevance of this debate becomes even more evident when one considers that physics education remains largely grounded in transmission-based practices, which contribute to students' lack of interest and to limited methodological renewal in the teaching and learning process. The introduction of GenAI into this context is a complex matter involving technical, ethical, epistemological, and pedagogical issues. Recent studies indicate that, despite its potential, GenAI may foster superficial understanding, technological dependence, and difficulties in developing critical thinking. Discussing its incorporation into science and physics education is therefore essential to understanding both its contributions and its limitations.

In light of these considerations, the general objective of this study is to analyze how the academic literature has addressed the use of GenAI in the physics and science teaching process between 2022 and 2025, taking into account its impacts, limitations, and potential. To achieve this objective, the study seeks to identify the main conceptual and operational challenges associated with GenAI; analyze how the studies describe its use by students and teachers; examine its pedagogical potential; discuss its limitations and risks; and identify gaps and trends in recent scientific research.

METHODOLOGY

This study adopted a Systematic Literature Review (SLR), with clearly defined methodological stages and procedures, to analyze and synthesize the findings of existing research on Generative Artificial Intelligence in the teaching and learning of physics and science. This methodological rigor was intended to ensure the validity, reliability, and reproducibility of the SLR, as noted by Vilelas (2022). The review entails the development of a planned research method that enables a comprehensive and transparent analysis of the existing literature, as proposed by Wang, Jing, and Shen (2025).

The research was conducted according to predefined stages[4], in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), as established by Moher et al. (2009) and Page et al. (2021). The Scopus database was searched for relevant studies because, in addition to being one of the largest databases of abstracts and citations in the literature, with broad global and regional coverage of scientific journals, Scopus returned results for the search terms used. By contrast, the Web of Science and ERIC databases returned no records when the same search strategy was applied, as the topic remains highly restricted. Furthermore, because of the current limitations and narrow thematic scope of the literature, this research was not restricted exclusively to journal articles; it also included other relevant academic publications, such as conference papers, provided that they were aligned with the subject under investigation.

[4] The stages defined for this study according to the checklist were eligibility criteria, information sources, search strategy, selection process, data-collection process, data items, assessment of risk of bias in the studies, effect measures, synthesis methods, assessment of reporting bias, and assessment of certainty. These stages followed the PRISMA 2020 statement (Page et al., 2021; Abreu et al., 2024).

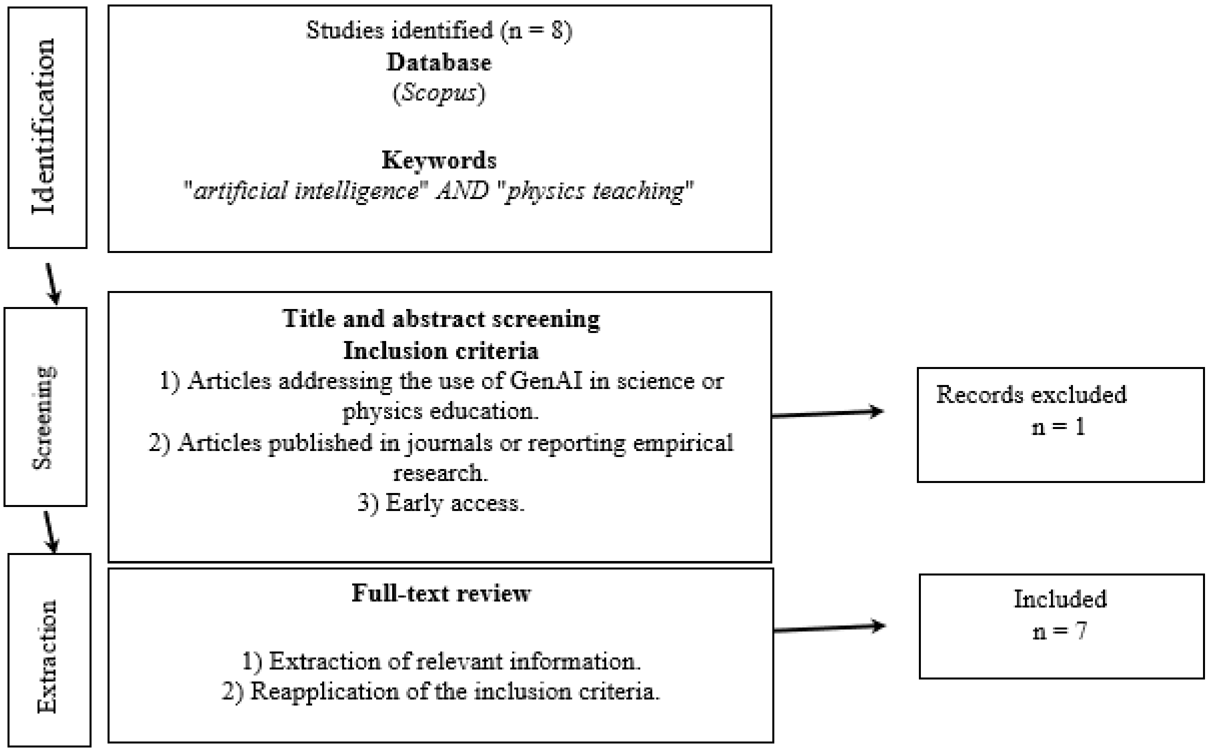
The search, conducted on November 5, 2025, used the following search string: TITLE-ABS-KEY (“artificial intelligence” AND “physics teaching”). It initially identified 18 documents in the Scopus database.

The search was not refined by document type and therefore included not only published journal articles but also conference papers. No language restrictions were imposed, and studies in both English and Portuguese were considered. No filters were applied for keywords, country, source type, source title, author name, publication stage, affiliation, funding sponsor, or open-access status. The only refinement concerned the period from 2022 to 2025, because 2022 marked the emergence of GenAI, particularly with the popularization of models such as ChatGPT. Thus, only publications from the preceding three years were considered, filtered by subject area to include Physics and Astronomy. After these criteria were applied, the search was narrowed to eight documents identified in the database.

Subsequently, the inclusion criteria encompassed only studies specifically addressing the application of GenAI to the teaching and learning of science and physics, articles published in journals or based on empirical research, and early-access articles. The following were excluded: reviews, editorials, or other theoretical articles; and articles that mentioned the terms “AI” or “GenAI” but in fact addressed either the application of GenAI to education in general or participants’ attitudes, intentions, or perceptions regarding GenAI education, without providing details about the implementation of educational practices in science and physics classrooms. At the end of the selection process, seven articles were chosen for inclusion in the analysis conducted in this systematic review, as shown in Figure 1, the systematic flow diagram.

Figure 1

Systematic Literature Review Flow Diagram



Source: Prepared by the authors, 2025.

BIBLIOMETRIC ANALYSIS

In this section, the publications are examined through a bibliometric analysis, which enables a better understanding of the studies according to year of publication, authorship, and the distribution of documents across countries and territories. This approach provides a clearer overview of the distribution of the research conducted over these three years, contributing to the replicability and greater rigor of the systematic review.

PUBLICATIONS BY YEAR

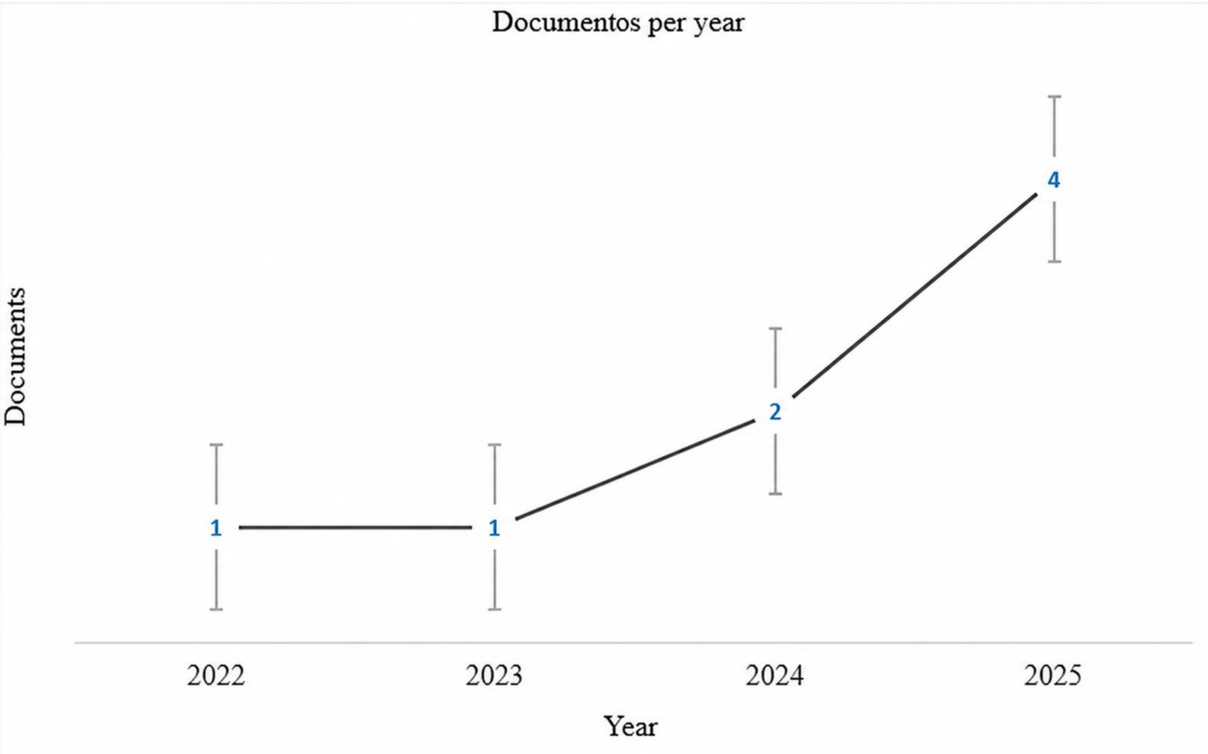
The first step consisted of categorizing the publications according to year of publication to provide a better understanding of the development of the research, as presented in Chart 1. The number of publications increased over the years, from one article in 2022 and one in 2023 to two in 2024 and four in

2025, indicating growing academic interest in the application of GenAI to physics education during this period.

Thus, although the topic has not yet been widely explored in the context of physics education, interest has steadily increased over time, contributing to greater visibility for the implementation of GenAI in this field.

Chart 1

Distribution of Documents by Year



Source: Prepared by the authors based on Scopus data, 2025.

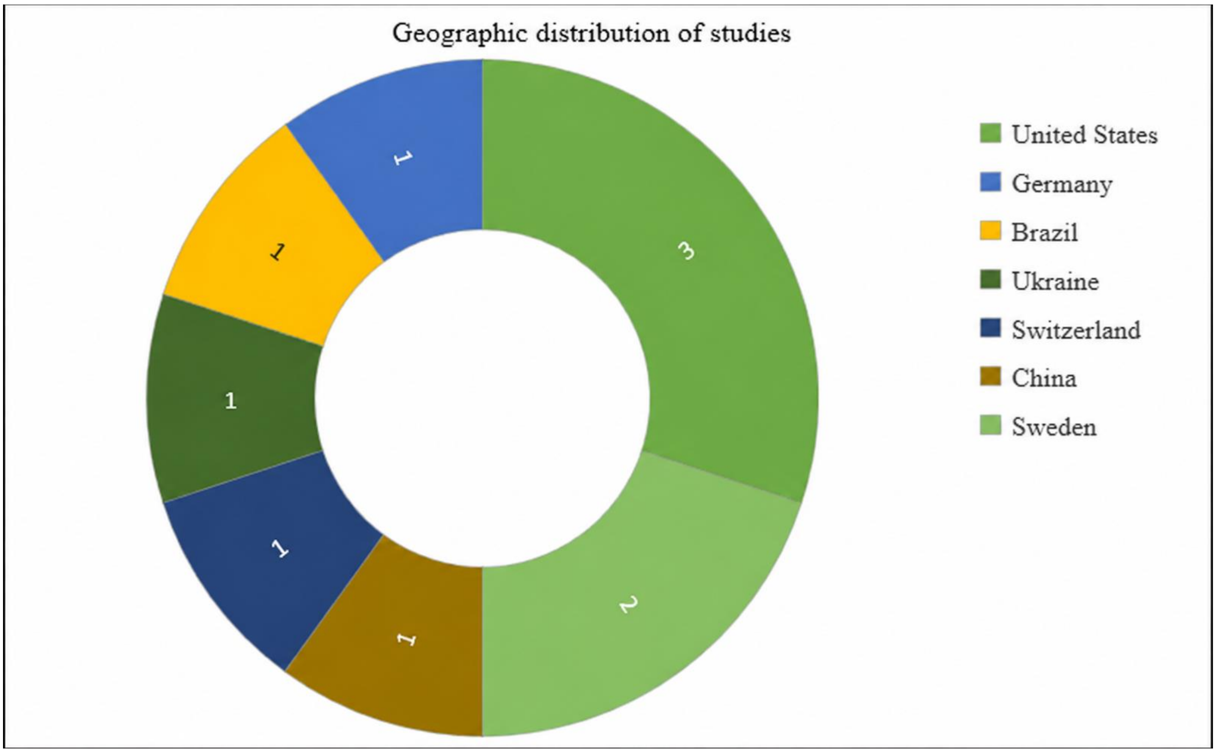
GEOGRAPHICAL DISTRIBUTION

In addition to the development of publications over time, the analysis also considered the geographical distribution of the studies, as illustrated in Chart 2. The highest concentration of publications was found in the United States, with three publications identified during the period analyzed, followed by Sweden, with two. The remaining countries—Brazil, China, Germany, Switzerland, and Ukraine—had one publication each.

The findings therefore indicate that the United States has shown greater interest in implementing GenAI in physics education and stands out as one of the leading centers for research development on the topic.

Chart 2

Geographical distribution of the studies



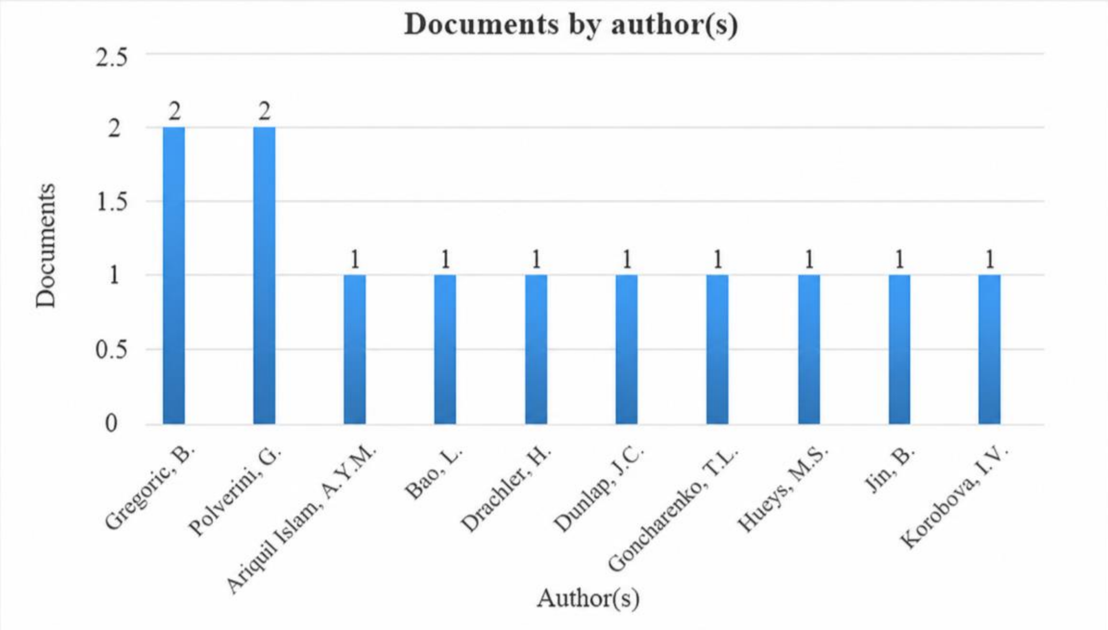
Source: Prepared by the authors based on *Scopus* data, 2025.

AUTHOR(S)

Finally, the documents were analyzed in terms of authorship, as presented in Chart 3. The authors with the highest number of publications were B. Gregorcic and G. Polverini, each with two publications. Their studies investigate the performance and limitations of GenAI in physics education, highlighting its potential as an educational tool despite its continuing conceptual and interpretive errors.

Chart 3

Distribution of documents by author(s)



Source: Prepared by the authors based on Scopus data, 2025.

RESULTS AND DISCUSSION

The findings support a more in-depth analysis of how the academic literature addresses the topic, revealing recurring patterns, knowledge gaps, and emerging trends identified through the systematic review.

A table was then prepared to present the distribution of the publications analyzed, including information such as title, authorship, journal, and number of citations, according to the data provided by the Scopus database, as illustrated in Table 1.

Table 1

Detailed distribution of each scientific publication

NO.	TITLE	AUTHOR(S)	JOURNAL NAME	NUMBER OF CITATIONS (SCOPUS)
1st	*Descendo um plano inclinado com um modelo de linguagem grande* [Descending an Inclined Plane with a Large Language Model]	Dunlap, Justin; Sissons Ryan; Widenhorn, Ralf.	Journals.aps.org.	0
2nd	*Desempenho do ChatGPT em tarefas envolvendo representações visuais de física: O caso da breve avaliação de eletricidade e magnetismo.* [Performance of ChatGPT on Physics Tasks Involving Visual Representations: The Case of the Brief Electricity and Magnetism Assessment]	Polverini, Giulia; Melin, Jakob; Önerud, Elias; Gregórcico, Bor.	Journals.aps.org.	4
3rd	*Explorando o papel da colaboração entre humanos e IA na resolução de problemas científicos.* [Exploring the Role of Human–AI Collaboration in Scientific Problem-Solving]	Tong, Dazhenum; Jin, Bangjianb;Tao, Yangc; Ren, Hong Meium; Atiquil Islam AYMd; Bao, Leif.	Journals.aps.org.	1
4th	*Auxílio na correção de uma prova de termodinâmica manuscrita utilizando inteligência artificial: um estudo exploratório.* [Assisting the Grading of a Handwritten Thermodynamics Exam Using Artificial Intelligence: An Exploratory Study]	Kortemeyer, Gerda d; Nöhl, Julianb; Onishchuk, Dariac.	Journals.aps.org.	9
5th	*Como a compreensão de grandes modelos de linguagem pode orientar o uso do ChatGPT no ensino de física.* [How Understanding Large Language Models Can Inform the Use of ChatGPT in Physics Education]	Polverini, Giulia; Gregórcico, Bor.	Iopscience.	66
6th	*Ensino de física mediado por aprendizado de máquina: o caso do pêndulo simples.* [Machine-Learning-Mediated Physics Education: The Case of the Simple Pendulum]	Prates, José Humberto de Souzada Costa; Camila Schneider Loureiro Pessoa; Elisa Fernanda Gurriti Mattedi; Guilherme do Prado Monteiro; Henrique Baleeiro de Jesus; José Victor Brito Namiok; Luiz Fernando Yano Huoya; Marçal Santos Matui; Yumi Alves.	Scielo.	1
7th	Experiência no desenvolvimento e implementação de um sistema de casos de ensino visualizados em Física utilizando um sistema de medição digital computadorizado Einstein. [Experience in Developing and Implementing a System of Visualized Teaching Cases in Physics Using the Einstein Computerized Digital Measurement System]	Yu Golovko N; Goncharenko TL; Korobova IV.	Iopscience.	3

Source: Prepared by the authors based on Scopus data, 2025.

The studies analyzed in Table 1 focus on the impacts of GenAI on the teaching and learning of physics and science, with an emphasis on identifying barriers through a systematic approach.

Accordingly, the selected studies are summarized below to ensure greater clarity, transparency, and replicability of the research.

In their article, Dunlap, Sissons, and Widenhorn (2025) seek to analyze how different permutations of the object, action verb, and properties of an inclined plane affect the model's responses. The research examines how different types of prompts can influence GenAI's final response, showing that minor variations in the wording of a problem can significantly alter how the model interprets and solves it. The authors emphasize that these variations reveal important shortcomings of GenAI in physics education, particularly its failure to make assumptions explicit and its difficulty in handling conflicting information.

Polverini et al. (2025) investigate ChatGPT-4's performance in solving visual physics problems, analyzing how the model interprets diagrams, vector fields, and other conceptual representations. The authors identify significant difficulties encountered by the chatbot, including incorrect interpretations, problems applying fundamental laws of physics, and limitations in spatial reasoning. Although they acknowledge that the implementation of GenAI has substantial potential and added value for education, particularly in supporting learning and accessibility, the study emphasizes that numerous barriers to its effective integration remain, including persistent errors in visual interpretation and conceptual inconsistencies that compromise the system's reliability.

Tong et al. (2025) analyze the solving of scientific problems by comparing human–AI collaboration with human–human collaboration. The findings show that many students used ChatGPT-4 only to obtain ready-made answers, without demonstrating an interest in developing critical thinking or engaging more deeply in the learning process. By contrast, human–human interactions produced more meaningful learning, as students discussed, argued, and reflected critically on the problems. The study also demonstrates that GenAI's limitations, such as inconsistencies in its responses and failures of visual and conceptual interpretation, represent a significant barrier to its effective implementation in physics and science education. Furthermore, it reinforces the need to instruct students on how to collaborate more productively with AI tools and avoid using these systems superficially and passively.

Kortemeyer, Onishchuk, and Nöhl (2024) evaluate four workflows for AI-assisted grading of thermodynamics examinations containing handwritten answers. The primary obstacle is converting handwriting into text that AI can read. Highly detailed rubrics produce grading errors because the model cannot process multiple criteria simultaneously; grading answers in separate parts is more reliable but loses nuances. Hand-drawn diagrams and graphs also increase the error rate. AI accurately identifies passing examinations, but failed examinations still require human verification.

Polverini and Gregorcic (2024) seek to explain how prompt engineering can improve the performance of GenAI in physics research and tasks, showing that how users formulate prompts directly influences the quality of the responses. In addition, the article provides a clear introduction to how large language models function and discusses their strengths and limitations in the context of conceptual problem-solving. Using practical examples, the authors demonstrate how specific prompt-engineering techniques can provide GenAI with better guidance, making its responses more accurate and useful for physics education. The study thus shows that understanding and applying effective prompting strategies are essential to fully harnessing the potential of these models in educational settings.

Prates et al. (2023) propose a new pedagogical approach that integrates machine-learning techniques into the study of the simple pendulum. The authors collect experimental data on the pendulum's motion and use a regression algorithm to model the system with a high degree of accuracy. The proposal demonstrates how machine-learning methods can complement traditional physics education, introducing students to data science and emerging technologies. The study emphasizes that the education of future physicists should include computational and data-analysis competencies.

Golovko et al. (2022) discuss the need to modernize the educational system, emphasizing the role of digital technologies, AI, and science, technology, engineering, and mathematics (STEM) education. They propose the development of a system of "visualized teaching cases" for physics using Google tools and the Einstein digital measurement system. The study reports the practical experience of implementing this approach at Kherson State University, showing how visualized cases integrated into real-world contexts can improve physics learning and increase students' interest in STEM fields.

Broadly speaking, the studies seek to investigate how GenAI can contribute to students' conceptual development, support assessment processes, foster active learning, and promote more meaningful interaction between humans and machines in educational settings.

Although the studies explore different aspects—including model performance on physics problems, human–AI collaboration, automated exam grading, the use of machine-learning algorithms, and the development of interactive digital resources—they all share a common focus on the relationship between technology and education, particularly the limitations and potential that these tools may bring to the physics teaching and learning process.

Although varied, the methodologies share a similar purpose: to evaluate the effectiveness and challenges of using GenAI as a pedagogical instrument. Whether through qualitative analyses of chatbot-generated responses, comparative experiments involving humans and AI, or practical applications of digital systems in the classroom, all the studies seek to understand how these tools influence scientific reasoning, the interpretation of visual representations, problem-solving, and student autonomy.

Furthermore, the studies share a concern with preparing teachers and students to meet the new demands of the digital age, as illustrated in Table 2. The authors indicate that integrating advanced technologies into education is not intended to replace traditional methods but rather to promote critical, reflective, and collaborative learning. In summary, these studies converge on the idea that GenAI can transform physics education, provided that its limitations are recognized and its use is accompanied by appropriate pedagogical strategies capable of enhancing the teaching and learning process.

Table 2

Aspects and Requirements of the Digital Age Identified in the Scientific Literature

	Aspects	Digital-Age Requirements and Transformations Identified in the Studies
1	Development of new competencies	Integrating artificial intelligence and digital technologies into physics education requires the development of new pedagogical, cognitive, and technological competencies. Teachers and students must understand how these tools function, recognize their conceptual limitations, and use them critically and meaningfully.
2	Teacher training focused on the pedagogical use of AI	Educators must receive specific training on the use of AI in the classroom, learning to formulate effective prompts, interpret model responses, and promote scientific inquiry activities that stimulate reasoning and reflection rather than merely encourage the search for automated answers.
3	Student preparation	Students must be prepared to interact with AI tools collaboratively, ethically, and critically, developing scientific thinking, intellectual autonomy, and the ability to validate information.
4	Digital and scientific literacy	The importance of digital and scientific literacy is emphasized. Such literacy entails understanding the foundations of artificial intelligence, machine learning, and data analysis and applying this knowledge to the study of physical phenomena.
5	Integrated teaching strategies	Pedagogical strategies combining practical experimentation, computational analysis, and interactive visual resources must be adopted to make the teaching and learning process more dynamic, contextualized, and meaningful.
6	Redefinition of the roles of teachers and students	Meeting these requirements entails reconsidering the teacher’s role. Rather than merely transmitting content, teachers must act as mediators and guides in technology-mediated learning processes. Students, in turn, must adopt a more active, investigative, and collaborative approach.
7	Science education for the twenty-first century	These transformations are essential to preparing teachers and students for science education that meets the cognitive and technological demands of the twenty-first century, promoting critical and reflective learning integrated into the digital context.

Source: Prepared by the authors based on the studies reviewed, 2025.

As indicated in Table 2, the requirements of the digital age demonstrate that integrating GenAI into physics education requires both teachers and students to develop new pedagogical, cognitive, and technological competencies. To use GenAI effectively, teachers must be proficient in formulating prompts, understand the technology’s limitations and educational potential, critically interpret the responses it generates, and propose activities that stimulate investigative interest and go beyond merely querying the model.

Furthermore, students must be prepared to interact with AI ethically and critically while developing scientific thinking, logical reasoning, and the ability to evaluate information. Another relevant aspect is digital and scientific literacy, which enables students to understand and interpret fundamental concepts in both AI and physics.

The studies therefore indicate that this transformation is essential in the context of the twenty-first century, demonstrating that teacher education plays a central role in enabling teachers to mediate technology-enhanced learning and foster students who are more active, investigative, and capable of taking a leading role in their own knowledge-construction process.

FINAL CONSIDERATIONS

The analysis of the selected studies indicates that the use of Generative Artificial Intelligence in physics and science education constitutes an emerging and rapidly expanding field characterized by growing academic interest. The findings indicate that GenAI has significant potential to support learning, particularly by fostering conceptual development, problem-solving, personalized instruction, and the adoption of more active and innovative pedagogical practices.

However, the findings also reveal relevant limitations that cannot be overlooked. Conceptual errors, the superficial nature of certain responses, and the tendency to use the tool uncritically are challenges that require conscious teacher intervention and mediation. In this context, integrating GenAI does not diminish the importance of pedagogical mediation. On the contrary, it requires the development of new competencies, such as proficiency in prompt engineering, the ability to critically evaluate generated responses, and the capacity to guide students toward an active, reflective, and investigative approach.

Furthermore, the studies analyzed indicate that the responsible and strategic adoption of GenAI can contribute to the development of more dynamic and inclusive learning environments aligned with contemporary demands. Its implementation must therefore be understood as a complement to existing methodologies rather than as a replacement for teachers' work or traditional practices. A balance between innovation and critical awareness is thus essential to ensuring that technology enhances rather than compromises the quality of education.

In summary, GenAI is a promising tool for physics and science education, provided that it is used ethically, critically, and with appropriate pedagogical guidance. Progress in this field will depend on continuing teacher education, the development of responsible educational practices, and the consolidation of a pedagogical culture that understands technology as a means of enhancing learning and strengthening students' agency in the educational process.

REFERENCES

Dunlap, Justin; Sissons, Ryan; Widenhorn, Ralf. Descending an inclined plane with a large language model. *Physical Review Physics Education Research*, 2025.

Golovko, N. Yu.; Goncharenko, T. L.; Korobova, I. V. Experience in developing and implementing a system of visualized teaching cases in physics using a computerized digital measurement system Einstein. *IOP Conference Series*: Materials Science and Engineering, 2022.

Kortemeyer, Gerda; Nöhl, Julian; Onishchuk, Daria. Assisting grading of handwritten thermodynamics exams using artificial intelligence: an exploratory study. *Physical Review Physics Education Research*, 2024.

Moreira, Marco Antônio. *Ensino de Física no Brasil*: desafios e perspectivas. [Physics Education in Brazil: Challenges and Perspectives] São Paulo: Editora Livraria da Física, 2018.

Page, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, v. 372, 2021.

Polverini, Giulia; Gregorcic, Bor. How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics*, 2024.

Polverini, Giulia et al. Performance of ChatGPT on physics tasks involving visual representations: the case of the Brief Electricity and Magnetism Assessment. *Physical Review Physics Education Research*, 2025.

Prates, José Humberto de Souza et al. Ensino de física mediado por aprendizado de máquina: o caso do pêndulo simples.[Machine-Learning-Mediated Physics Education: The Case of the Simple Pendulum] *Revista Brasileira de Ensino de Física*, 2023.

Tong, Dazhen et al. Exploring the role of human-AI collaboration in scientific problem solving. *Physical Review Physics Education Research*, 2025.

Vilelas, José. *Investigação*: o processo de construção do conhecimento. [Research: The Knowledge-Construction Process].Lisboa: Edições Sílabo, 2022.

Wang, Peijun; Jing, Yuhui; Shen, Shusheng. A systematic literature review on the application of generative artificial intelligence in higher education: instructional contexts, processes and strategies. *The Internet and Higher Education*, 2025.