Computer Applications in Engineering Education

WILEY

RESEARCH ARTICLE OPEN ACCESS

Chatbots and Generative AI in Engineering Education: A Structured Literature Review (2020–2024)

Evelina Staneviciene[1] | Daina Gudoniene[1] | Jurgita Cerneckiene[2] | Rokas Valancius[2] | Kestutis Jankauskas[1] | Dylan Bulseco[3] | Charles Xie[3] | Andriy Kashyrskyy[3] | Hongling Deng[3]

[1]Faculty of Informatics, Kaunas University of Technology, Kaunas, Lithuania | [2]Faculty of Civil Engineering and Architecture, Kaunas University of Technology, Kaunas, Lithuania | [3]Institute for Future Intelligence, Natick, Massachusetts, USA

Correspondence: Evelina Staneviciene (evelina.staneviciene@ktu.lt)

Received: 2 July 2025 | **Revised:** 19 June 2026 | **Accepted:** 25 June 2026

Keywords: chatbots | engineering education | generative AI | generative technologies | higher education

ABSTRACT

The integration of chatbots and generative artificial intelligence (AI) tools into engineering education is rapidly changing the way students learn, interact, and solve complex problems. These technologies offer new opportunities for personalized learning, real-time feedback, and enhanced student engagement. However, a comprehensive understanding of their implementation, pedagogical value, and limitations in engineering education remains limited. This structured literature review examines how chatbots and generative AI tools are integrated into engineering education and evaluates their educational impact and associated challenges. The review focuses on their effects on student learning outcomes, engagement, and skills development, as well as the challenges associated with their implementation. Following PRISMA guidelines, literature was identified through Scopus, Google Scholar, Taylor & Francis Online, and ScienceDirect. After applying predefined inclusion criteria, 16 studies were included in the final review and were analyzed thematically. The findings show that chatbots and generative AI tools can improve learning outcomes in engineering education by promoting student engagement, supporting conceptual understanding, encouraging self-directed learning, and contributing to the development of problem-solving and creative thinking skills. However, the review also identifies challenges related to ethical concerns, accuracy limitations, and the risk of student overreliance, which can impact the development of deeper learning and collaboration.

1 | Introduction

In recent years, the integration of artificial intelligence into education has accelerated rapidly [1], transforming the way learners acquire knowledge, interact with content, and receive feedback [2]. Engineering students are expected to acquire not only theoretical knowledge but also practical competencies in problem solving, system thinking, and design thinking, which are essential for addressing complex engineering challenges [3, 4]. The complexity and interdisciplinary nature of engineering learning environments require educational technologies that can support iterative experimentation and collaboration, where generative AI tools can indicate high potential [5].

The use of AI in engineering education has emerged as a promising development, and generative AI technologies such as ChatGPT or other intelligent chatbots have opened new opportunities to improve the quality of education and learning outcomes [6, 7]. Such AI-based systems have the potential to provide personalized and effective learning experiences by offering customized feedback and explanations [7, 8]. Intelligent chatbots can transform pedagogical practices by interactively guiding students and teachers, addressing student engagement challenges, and providing real-time support [6, 9]. Chatbots can transform knowledge acquisition and skill development [10], facilitate learning in conditions similar to those of a human teacher [11], and offer unique communication and information dissemination opportunities in digital learning environments [2]. Furthermore, these tools go beyond simple text generation to help in tasks such as LaTeX code generation and debugging programming exercises [9]. Among these intelligent chatbots, ChatGPT stands out as a conversational AI interface that has received significant attention in the educational settings [12]. However, other definitions of chatbots exist that rely on the inclusion of additional knowledge or content through approaches such as retrieval-augmented generation (RAG) [13]. A common example is chatbots used on consumer websites that integrate a company- or domain-specific knowledge base.

However, the implementation of AI in engineering education is not without challenges. There are ethical concerns about the potential unethical use by students and issues of academic integrity [7, 14]. There are pedagogical, technological, and socio-cultural risks associated with the widespread use of AI chatbots in education that need to be carefully considered and managed [15]. Additionally, AI systems can retain bias or generate false information because they are limited by the quality of their training data [7] and may be subject to misuse, including plagiarism, academic integrity, and privacy violations [1]. Engineering educators, while recognizing the benefits of AI tools in facilitating academic tasks, have also expressed concerns about the quality of AI-generated output and the risk that students will become too dependent on technology [16]. To fully exploit the benefits of AI in engineering education, curricula need to be adapted to the needs of the user. This includes developing future-ready curricula that incorporate AI literacy, new teaching and learning approaches, analytical and critical thinking, reflection, collaboration, and complex problem solving [14]. Despite these challenges, a growing body of research on educational chatbots shows their positive impact on student learning and service improvement in various educational settings [11].

While interest in AI-enhanced learning is growing, there is still a lack of reviews on how chatbots and generative AI tools are used in specific engineering contexts and what impact they have on learning outcomes, engagement, and skill development. Most existing reviews [17–20] take a broad approach to AI in education, in general, often without considering the different cognitive and practical needs of engineering learners. To address this gap, this review aims to summarize current research on the application of chatbots and generative AI technologies in engineering education by answering the main research questions (RQ):

RQ1. *How do chatbots and generative AI technologies affect students' learning outcomes, engagement, and skills development in engineering courses?*

RQ2. *What are the challenges of using chatbots and generative AI in higher engineering education?*

The rest of this paper is structured as follows: Section 2 describes the methodological approach used in the study. Section 3 presents the results and discusses the key findings. Finally, Section 4 concludes the paper by summarizing the main findings, limitations, and contributions.

2 | Background

In the emerging landscape of engineering education, while student outcomes remain a central emphasis, the experiences of educators and their adaptive responses merit equal consideration. Recent evidence indicates that instructors in engineering and related STEM fields perceive both opportunity and disruption when integrating advanced AI tools such as chatbots and generative AI. For example, in a survey of 98 engineering educators, most reported familiarity with generative AI but showed highly variable use; those who actively engaged students in discussions of AI tended to hold more optimistic views, yet also recognized ethical, reliability, and pedagogical risks [21]. This finding underscores a key adaptation-challenge: educators must shift from the role of sole knowledge provider toward that of facilitator and curator of AI-mediated learning environments [22].

One prominent challenge is the need for instructors to develop new technological literacies and pedagogical strategies. Bravo and Cruz-Bohorquez [23] found that although AI chatbots in engineering courses provided time savings in lesson planning and materials creation, staff reported operational burdens in configuring, maintaining, and validating such systems. These burdens often resulted in increased workload or role ambiguity for educators, particularly when alignment of tool outputs with course outcomes was not well defined. Indeed, the realities of engineering pedagogy with its hands-on labs, design projects, and problem-solving focus mean that replacing traditional modes of feedback and interaction with AI systems requires careful scaffolding, not mere substitution [24, 25].

Furthermore, adaptation requires institutional and cultural change. Educators expressed uncertainty about assessment design, academic integrity, and student–faculty relationships when AI becomes more embedded in the learning process [21]. There is a risk that increased reliance on AI for routine queries or administrative tasks may reduce opportunities for instructors to monitor student progress intimately or to intervene holistically, thus weakening mentor–student relationships [26]. Consequently, institutions may need to provide professional development, clearly delineated roles, policy frameworks, and technological support to facilitate successful adaptation.

The integration of chatbots and large language models (LLMs) in engineering education has demonstrated both promising advantages and notable challenges. Recent studies reveal that educational chatbots enhance learning engagement and alleviate instructional workload by automating routine academic and administrative tasks [27, 28]. However, despite these benefits, the widespread adoption of chatbots in traditional educational contexts remains limited. This limitation is primarily attributed to the high developmental costs, maintenance requirements, and a general lack of awareness about chatbot potential in pedagogy [29]. In education, chatbots operate on the premise that teaching is an act of communication and interaction. Chatbots can address repetitive questions with straightforward solutions or assist students in finding the correct answers to simple assignments. However, it is crucial to emphasize that teachers remain behind the teaching process. Technology, including chatbots, should be employed as an assistant rather than a replacement [29].

The advantages of chatbots in open learning for schools, universities, and other learning scenarios in which repetitive practice is important are well known. In addition to answering frequently asked questions (FAQs), there are cases of chatbots whose main purpose is to serve as educational agents with the aim of relieving the workload of human teachers in their subjects or reinforcing learning in different areas. It is especially on the main advantages and disadvantages of using chatbots as teaching assistants that this research intends to focus [11].

Educators must negotiate pedagogical redesign and role-redefinition to ensure that the human dimension of teaching remains central [24, 25]. Without attending to proper pedagogical integration, the shift in engineering education risks becoming substitute-driven rather than supplement-focused [26].

Students' interaction with AI technologies directly shapes learning outcomes, engagement, and skill development [24, 25, 27]. Moreover, understanding student experiences provides essential feedback for improving pedagogical design and developing adaptive frameworks for AI integration [25].

Extending this focus to practical implementation, ChatGPT serves as a representative example of AI integration in education. ChatGPT interacts with students and collects data, ensuring the privacy and security of sensitive information becomes paramount [12]. Educational institutions and developers must implement robust data protection measures to safeguard student data and comply with relevant privacy regulations. Moreover, despite some limitations, the 5E model proved to be an effective framework for student engagement and learning. One of the limitations observed was that ChatGPT occasionally provided incorrect evaluations of the students' answers during the evaluation phase [30].

3 | Methodological Approach

This literature review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency, replicability, and methodological rigor [31]. PRISMA offers a structured approach with clearly defined steps, including formulating research questions, systematically identifying relevant literature, screening for eligibility, and synthesizing and presenting findings. The methodological process of this review followed five main steps: (1) defining the research objectives and questions, (2) identifying and retrieving relevant studies, (3) applying the PRISMA protocol for screening and selection, (4) extracting and organizing key information from the included articles, and (5) conducting a thematic synthesis to summarize the findings.

The research process began with the formulation of two central research questions: (1) How do chatbots and generative AI tools impact student learning outcomes, engagement, and skill development in engineering education? (2) What are the primary challenges and limitations associated with their integration?

These two research questions are selected to discuss in a review because they address both the potential benefits and the challenges of integrating chatbots and generative AI tools into engineering education. By examining the impact on student learning outcomes, engagement, and skill development, the review highlights how these technologies can enhance educational experiences and foster essential competencies such as problem-solving and critical thinking. Simultaneously, exploring the challenges and limitations—such as ethical concerns, accuracy issues, and overreliance—ensures a balanced understanding of the risks and constraints. Together, these questions provide a comprehensive view that can guide educators, policymakers, and researchers in making informed decisions about the effective and responsible use of AI in engineering education.

3.1 | Literature Identification and Source Selection

To fully understand the research landscape related to chatbots and generative artificial intelligence in engineering studies, a comprehensive literature search was conducted. This search focused on peer-reviewed scientific publications written in English and published between 2020 and 2024. This 5-year period covers a critical period that marks the emergence and rapid development of modern LLMs, from the initial breakthrough of GPT-2 in 2019 to the widespread adoption of ChatGPT in 2022. The focus on this most recent phase ensures that the review highlights the latest innovations, pedagogical applications, and challenges.

The literature search was conducted using widely recognized academic sources that provide comprehensive and relevant coverage of education and technology research. Specifically, literature was identified through Scopus (a bibliographic database), ScienceDirect (a publisher-specific full-text platform), Taylor & Francis Online (a publisher platform), and Google Scholar (an academic search engine). These sources were selected because they provide access to a wide range of peer-reviewed publications relevant to the scope of the review. Web of Science was initially explored as a potential source. However, after applying the predefined search strategy, only a limited number of records matched the search criteria. Therefore, it was not included among the literature sources used in the final review. Boolean operators and filters were applied to refine the search. Only English-language articles that provided empirical data or case-based evidence on AI integration in educational settings were included. This process ensured a focused and reliable dataset for the review.

3.2 | Search Strategy and Keyword Selection

The search strategy was carefully designed using thematic keyword clusters to ensure complete coverage and accuracy. Keywords were grouped into three distinct categories: (1) chatbot terminology: chatbot, virtual assistant, AI tutor, AI-assisted; (2) generative AI terminology: generative technologies, generative AI; (3) terminology of the educational context: engineering teaching, engineering learning, engineering education.

Within each thematic category, individual keywords were combined using the logical operator OR to ensure comprehensive coverage of each conceptual area. Different thematic categories were combined using the logical operator AND, thus refining the search to find studies that explicitly address the interaction of chatbots or generative artificial intelligence in engineering studies. The search strategy was designed to capture a comprehensive range of studies related to the integration of chatbots and generative AI tools in engineering education. A combination of keywords and Boolean operators was used to ensure that relevant literature was identified. The key terms included, and these were combined using operators such as AND, OR, and NOT to refine the search results. The search string used to identify the relevant literature was as follows:

("chatbot" OR "virtual assistant" OR "AI tutor" OR "AI-assisted") AND ("generative technologies" OR "generative AI") AND ("student support" OR "personalized learning") AND ("engineering teaching" OR "engineering learning" OR "engineering education").

This search string was used in Google Scholar, Scopus, and Taylor & Francis Online. However, due to platform limitations, a modified version of the search string was used in ScienceDirect, where the number of Boolean operators is limited to eight. The simplified search string used in ScienceDirect was:

("chatbot" OR "virtual assistant") AND ("generative technologies" OR "generative AI") AND ("student support" OR "personalized learning") AND ("engineering teaching" OR "engineering learning" OR "engineering education").

This adapted query maintained the main thematic focus while adhering to technical limitations. Filters were applied to include only publications published in English, with a preference for peer-reviewed journal articles and research. The chosen search string was designed to capture a broad, yet focused range of studies related to AI technologies in engineering education. Terms like "chatbot," "virtual assistant," "AI tutor," and "AI-assisted" reflect different forms of conversational or instructional AI tools. Including "generative technologies" and "generative AI" ensured coverage of studies involving content-creating systems such as ChatGPT. Phrases like "student support" and "personalized learning" targeted the pedagogical impact of these tools, while "engineering teaching," "engineering learning," and "engineering education" narrowed the scope to the relevant academic field. This combination enabled the retrieval of diverse but context-specific literature. Due to ScienceDirect's Boolean operator limit, a simplified version was used, maintaining core concepts while adapting to platform constraints. This strategic keyword selection enhanced both the precision and breadth of the literature search across the selected academic sources.

3.3 | Screening Process and Criteria for Inclusion and Exclusion

After the initial search, all retrieved records were further screened and organized. Duplicate records were identified and manually removed. The remaining studies were then screened in three stages: first, by reviewing the titles to exclude apparently unrelated publications; second, by evaluating the abstracts to assess thematic relevance and alignment with the inclusion criteria; and third, by reviewing the full text of the remaining articles to confirm eligibility based on content and scope.

To ensure consistency and relevance, predefined inclusion and exclusion criteria were applied (see Table 1).

Initially, a total of 1018 possible articles were retrieved from: Google Scholar (952), Scopus (55), ScienceDirect (6), and Taylor & Francis Online (5). After removing 47 duplicates and 7 non-English articles, 964 unique records remained for selection. In the first stage of screening, the relevance of the titles was reviewed, resulting in the exclusion of 848 records. In the second stage, the abstracts of the remaining 116 articles were evaluated, and 70 were excluded for not meeting the thematic scope of the review. In the third and final stage, 46 full-text articles were evaluated based on inclusion and exclusion criteria. Of these, 30 were excluded for reasons such as lack of peer review, focus on non-engineering domains, or insufficient relevance to generative AI or chatbots. As a result, 16 studies were included in the final set. The entire screening and selection process is illustrated in the flow diagram (see Figure 1), which shows the number of records identified, selected, excluded, and finally included.

This filtering process ensured that only high-quality, thematically relevant studies were included for detailed analysis, enabling a focused, evidence-based synthesis of the research landscape.

3.4 | Data Analysis

After the screening phase, a detailed content analysis was conducted on the selected articles. A thematic synthesis approach was employed in this review instead of a quantitative meta-analysis. Statistical generalization of effect sizes was not

TABLE 1 | Criteria for inclusion and exclusion of articles.

Inclusion criteria	Exclusion criteria
Studies that focus primarily on the application or impact of chatbots or generative artificial intelligence in engineering studies.	Studies that focus primarily on contexts other than engineering studies.
Peer-reviewed journal articles.	Conference proceedings, academic books, and opinion pieces.
Articles published between 2020 and 2024.	Articles published outside of the stated publication date.
Articles written in English.	Articles written in a language other than English.

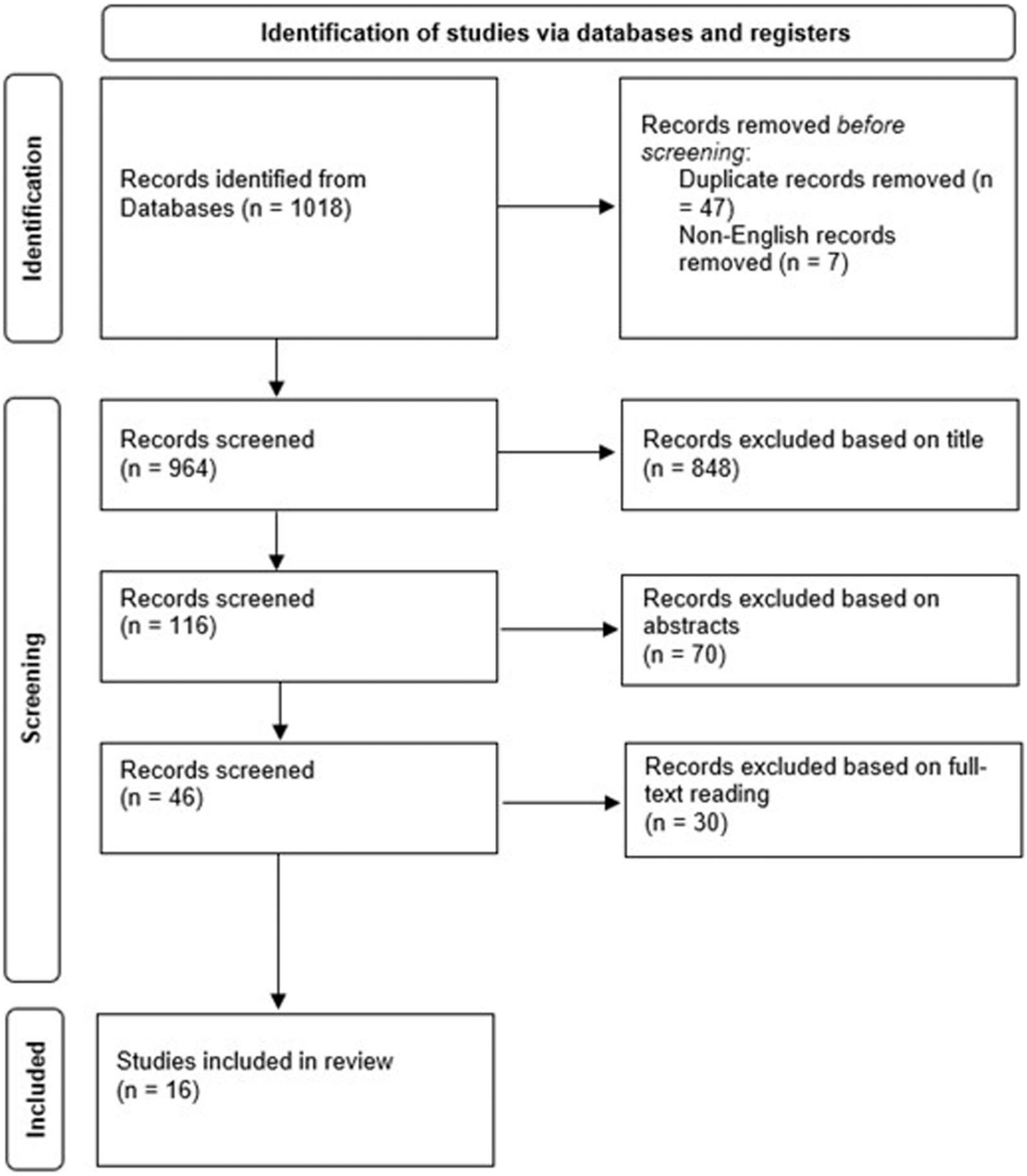

FIGURE 1 | Flowchart of the study selection process in the review.

conducted because of variations in study design, measurement instruments, and reported outcomes across the included studies. Nevertheless, the qualitative synthesis yielded a comprehensive understanding of the overall patterns and findings. Although this review followed the PRISMA guidelines, no formal quality assessment system was applied to the included studies to ensure methodological rigor. Instead, methodological quality was determined based on study design and clarity of reporting. The aim of this analysis was to identify, categorize, and synthesize recurring themes and patterns relevant to the research objectives. In order to better understand the thematic focus of the selected articles, a diagram of the most frequently used keywords was created (see Figure 2). Keywords that appeared only once were excluded from the visualization.

This diagram represents the most frequently occurring keywords extracted from a selected set of publications included in this literature review. The keyword chart visualizes key concepts such as ChatGPT, higher education, generative artificial intelligence, and engineering education, with the importance of each term reflecting its frequency and relevance in the dataset.

The analysis focused on several key aspects, including the educational context and target population (e.g., engineering studies or specific modules), the type of technology used (e.g., GPT-based chatbots or other generative models), and the pedagogical functions performed by these tools, including personalized learning, student support, content creation, or assessment assistance. Furthermore, attention was paid to the reported outcomes, such as student engagement, academic performance, as well as challenges and limitations.

4 | Results

After completing the selection and eligibility assessment process, a total of 16 studies were included in the final review. These studies varied in geographical location, study design, and focus areas, but all met pre-specified inclusion criteria related to

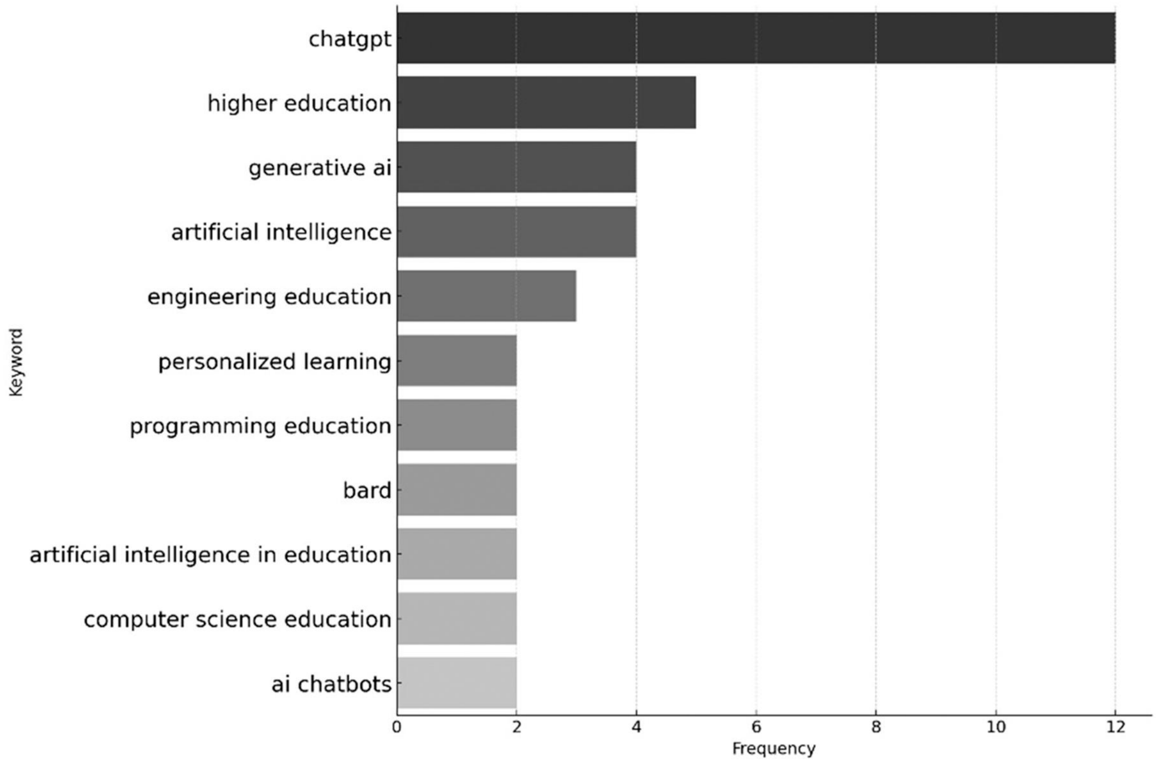

FIGURE 2 | Keywords for selected publications.

the use of chatbots or generative artificial intelligence in engineering studies. Table 2 provides an overview of the studies included in this literature review, highlighting key characteristics relevant to the use of generative AI in different research contexts. For each article, the table indicates the research area, the main role of generative AI in the study, the participant groups, and the research instruments used. Participants' educational level was only determined when it was clearly stated in the original studies, and no assumptions were made when this information was unclear. This categorization provides a clearer understanding of how generative AI is applied across disciplines and serves as a basis for further thematic analysis.

The reviewed studies reveal that generative artificial intelligence has been applied in a variety of support roles for both technical and cognitive aspects of student learning. The analyzed literature identified four main AI functions: (1) virtual tutoring and problem solving assistance, including answering questions, explaining code, and solving exercises [23, 30, 32, 34–36, 38–40, 42–45]; (2) assistance with writing and performing tasks, such as idea generation, editing and report writing [23, 33, 34, 39, 41, 43, 45]; (3) assistance with coding and debugging, including code generation, optimization, and error checking [23, 33, 34, 36, 38, 43, 44]; (4) personalizing and supporting education, especially through pre-lesson preparation, explaining concepts, and providing personalized resources [33, 34, 37, 40, 44]. According to the reviewed studies, ChatGPT was the most widely used generative AI tool, appearing in the majority of research projects (at least 14 out of 16). Several authors, including Albdrani and Al-Shargabi [30], Bravo and Cruz-Bohorquez [23], and others, relied solely on ChatGPT in their methodologies. Some studies used multiple tools; for example, Mustapha et al. [39] experimented with ChatGPT, Bard, DALLE, and 3DGPT; Randall et al. [44] used ChatGPT-3.5, ChatGPT-4, Bing, and Bard; Neyem et al. [40] investigated GPT-4, GPT-3.5, and large language model meta-AI (LLaMa). While ChatGPT clearly dominated, the inclusion of other tools in several studies indicates a growing interest in exploring a broader range of generative AI systems. Most studies indicate that these tools primarily help facilitate communication, provide feedback, and support active problem-solving tasks within educational education.

Methodologically, most studies used qualitative or mixed methods, combining instruments such as surveys, semi-structured interviews, and thematic analysis [23, 32, 43, 45]. Several studies also used experimental or quasi-experimental designs with control and experimental groups, assessing learning outcomes through pre- and post-tests [30, 35, 37]. The number of participants varied greatly from small qualitative studies with fewer than 10 participants [38, 42] to large-scale quantitative surveys with more than 200 participants [37, 41]. Most studies focused on undergraduate engineering studies, with fewer studies examining postgraduate studies or mixed participant groups.

Several empirical studies report quantitative indicators related to usage and performance that go beyond participant characteristics. Elkefi et al. [34] report that 58.46% of students used ChatGPT frequently and 5.75% rarely, while qualitative responses highlighted perceived time savings during learning activities. Randall et al. [44] provide comparative performance results showing that GPT-4 passed 33% of unit tests compared to 6.6% for GPT-3.5, along with usage statistics showing that 80% of students used AI tools for coding and 84% for time savings. Pavlenko and Syzenko [41] report discipline-specific

TABLE 2 | Main characteristics of the reviewed studies.

Authors	Focus area (subject/field)	Primary generative AI role	Participants	Educational level	Research method	Limitations
Albdrani and Al-Shargabi (2023) [30]	Data science, specifically deep learning	ChatGPT acted as a virtual assistant, helping students by answering questions, providing examples, and offering learning support.	A total of 20 participants, 10 in the experimental group and 10 in the control group.	Undergraduate	Case study approach with two groups (experimental and control). Use of ChatGPT with the 5Es model in the experimental group. Mixed-methods approach (quantitative and qualitative data). Quantitative analysis of quiz results. Qualitative observations of engagement and interaction.	The sample size is relatively small. Focuses on short-term effects.
Bravo and Cruz-Bohorquez (2024) [23]	Engineering education	AI chatbots played multiple roles in the learning process, including tutor, peer, and supportive assistant.	A total of 38 participants from five different electronic and mechatronic engineering courses.	Undergraduate	Surveys (with closed-ended and open-ended questions). Thematic analysis. Coding process (two rounds).	The study has limitations due to the qualitative approach of the study. The sample is not representative of all students in similar courses. The findings cannot be generalized to all students.
Caccavale et al. (2024) [32]	Chemical engineering	The primary role of the AI was to act as a virtual tutor to perform audit exercises, replacing physical teachers in repetitive tasks.	A total of 22 students were enrolled in the GMP course (Good Manufacturing Practice) and 12 participants from the Spring Meeting of the European Network for Business and Industrial Statistics (ENBIS).	Postgraduate (Master's)	Surveys conducted among students and industry/academia practitioners.	Preliminary study with limited scope. Initial theoretical framework only. Lack of detailed quantitative results.

(Continues)

TABLE 2 | (Continued)

Authors	Focus area (subject/field)	Primary generative AI role	Participants	Educational level	Research method	Limitations
Duhaylungsod and Chavez (2023) [33]	Engineering and computer science	The primary role was to help students with academic tasks such as writing, programming, and conceptualizing ideas.	A total of 16 participants.	Not specified (college students)	One-on-one interviews. Thematic analysis.	N/A
Elkefi et al. (2024) [34]	Engineering education	Educational support tool for tasks such as writing, homework assistance, and coding.	A total of 195 participants.	Not specified (engineering students)	Descriptive statistics. Logistic regression. Thematic analysis. Semi-structured surveys with open-ended and multiple-choice questions.	Lack of comparative data from developed countries. Focus only on students' perceptions. Limited data collection period. Insufficient exploration of ethical implications and data privacy concerns.
Gasaymeh and AlMohtadi (2024) [35]	Programming	ChatGPT was used to enhance interactivity in both in-class and out-of-class activities. Students used it to solve exercises, prepare for class, clarify understanding, and deepen their knowledge of educational topics.	A total of 74 participants, 36 in the control group, and 38 in the experimental group.	Not specified (university students)	Pre- and post-tests.	Small sample size limits generalizability. Focus on immediate impact only. Limited to ChatGPT AI technology. External variables not considered.
Groothuijsen et al. (2024) [36]	Programming education in higher engineering education	Students used ChatGPT primarily for code explanation, error checking and debugging, code generation, increasing conceptual understanding, code optimization, and solving mathematical problems.	A total of 30 participants, 29 students, and 1 teacher.	Postgraduate (Master's)	Questionnaire (including the Technology Readiness Index Scale [TRI2]). Semi-structured group interviews with students. Semi-structured interviews with the teacher. Student performance data (grades for	Small sample size affects understanding of pre-existing skills and attitudes. Uncertainty about data saturation in qualitative data.

(Continues)

TABLE 2 | (Continued)

Authors	Focus area (subject/field)	Primary generative AI role	Participants	Educational level	Research method	Limitations
					individual and group assignments).	
Huesca et al. (2024) [37]	Engineering, specifically programming	To provide personalized and interactive learning resources for students in pre-class activities.	A total of 356 participants, 214 in the focus group and 128 in the control group.	Undergraduate	Pre-test-post-test analysis.	Limited to engineering and programming. Special cognitive skills required. Access to a computer and the internet is required.
Humble et al. (2024) [38]	Computer science education, specifically programming education	The primary role of ChatGPT in this study was to answer questions related to programming and CS education.	6 participants (the authors themselves).	Higher education (not specified); researchers	Thematic analysis of collected data.	Limited data collection. Limited number of interactions analyzed. Findings not generalizable.
Mustapha et al. (2024) [39]	Mechanical engineering	The primary generative role of AI is to enhance mechanical engineering and professional practices through applications in material selection, manufacturing troubleshooting, technical documentation, and product positioning.	The paper does not provide information on participants or units.	N/A	The researchers conducted experiments with selected GenAI tools for mechanical engineering-related tasks. They used specific prompts to evaluate the tools' performance.	The rapid pace of publication limited the coverage of bibliometric analysis. GenAI tools are constantly updated, potentially resolving identified issues. GenAI tools are stochastic, yielding variable outputs. No use of advanced prompt engineering techniques for reproducibility.
Neyem et al. (2024) [40]	Software engineering education	The primary role of generative AI was to provide context-aware recommendations and answers to students' queries related to their software development projects.	A total of 150 participants, 37 females and 113 males.	Undergraduate	Survey with evaluation rubric.	The platform provides generic recommendations, rather than tailoring them to the user's profile.

(Continues)

TABLE 2 | (Continued)

Authors	Focus area (subject/field)	Primary generative AI role	Participants	Educational level	Research method	Limitations
Pavlenko and Syzenko (2024) [41]	Business, engineering, IT	ChatGPT was primarily used by students for information retrieval, speeding up assignment completion, brainstorming ideas, improving assignment quality, and text editing.	A total of 247 participants.	Mixed higher education (Bachelor's, Master's, PhD)	Questionnaire. 5-point Likert scale. Online survey through Google Form.	Relatively small sample size. Focus on only three academic domains.
Pham et al. (2023) [42]	Embedded systems education	ChatGPT was used to answer questions and exercises from the course materials.	1 participant (the lecturer).	Undergraduate	Evaluation of ChatGPT as an AI-assisted learning tool. Development of an AI-assisted learning flow. Use of 15 selected questions and exercises. Use of rating systems to evaluate answers.	Findings are based on the lecturer's perspective. The study involved only the lecturer. Qualitative insights from observational methods.
Rajabi et al. (2024) [43]	Computer science and engineering	The study explores various potential roles for ChatGPT in higher education, including initiating ideas for assignments, finding and fixing code mistakes, and completing tasks where accuracy can be verified.	A total of 40 participants, 13 undergraduate students, 21 graduate students, and 6 faculty members.	Undergraduate; postgraduate; faculty members	Focus group sessions. Thematic analysis.	The participant sample may not represent diversity. Time constraints impacted discussion depth. Reliance on a single notetaker may omit points.
Randall et al. (2023) [44]	Software engineering education	The primary roles included writing code, explaining existing code, and supporting students in completing software engineering tasks.	Bachelor's students with varying coding proficiencies and two highly skilled students studying SE courses.	Undergraduate	Survey. Interviews. Experiments using ChatGPT. Analysis using machine learning techniques.	The study had a limited scope, as it analyzed only a small group of students.
Remoto (2024) [45]	Mathematics	Assist students in learning mathematics by providing solutions, explanations,	A total of 15 participants.	Higher education (not specified)	Semi-structured one-on-one interviews. Thematic analysis.	The generalizability of results is limited by research design.

(Continues)

TABLE 2 | (Continued)

Authors	Focus area (subject/field)	Primary generative AI role	Participants	Educational level	Research method	Limitations
		and proofs, as well as helping with tasks like essay writing.			Collection of chat prompt responses from AI models. Purposive sampling.	Design is not sufficient to expand results to other fields. Need for experimental and longitudinal studies to capture long-term data. Certain concepts, such as awareness, attitude, motivation, and their effects on academic success, were not covered.

Abbreviation: N/A, data not provided in the article.

usage intensity: 43% of IT students report regular use of ChatGPT, as well as task-related usage patterns such as writing assignments and literature review assistance. Rajabi et al. [43] refer to time-constrained focus group sessions used to verify task accuracy without providing quantitative indicators of system performance.

A consistent pattern across all studies is that generative AI was not used as a stand-alone teaching method, but rather as an augmentation tool to existing pedagogical strategies. In several studies, it replaced or supplemented repetitive or technical tasks, allowing educators to focus on higher-level teaching activities [32, 39].

In addition to general study designs and participant characteristics, several studies also described specific methods used to assess the correctness and reliability of responses generated by chatbots. Caccavale et al. [32] describe small-scale pilot tests and qualitative peer review to assess the appropriateness of model responses before implementation. In technical and programming contexts, correctness was often assessed through task-based verification. For example, Humble et al. [38] compared responses generated by chatbots with deterministically verifiable programming solutions, while Mustapha et al. [39] assessed compliance with basic engineering decision procedures and systematically identified errors. Furthermore, Neyem et al. [40] evaluated the results using a rubric-based recommendation quality assessment supplemented by readability analysis and similarity measures. Additional structural validation methods were described in several studies. In the context of mathematics and software engineering, the correctness of chatbots was assessed by comparing generated responses to verified solutions or teacher-provided reference responses, as demonstrated by Remoto [45]. In programming education, automated unit testing was used to evaluate the functionality and correctness of generated code and to quantify performance in terms of test pass rates [44]. Meanwhile, responses were validated using multi-stage verification procedures that combine manual code review with hardware testing on the Arduino platform using a structured evaluation framework in embedded systems education [42]. Overall, the reviewed studies show that chatbots are most often evaluated based on learning outcomes and user perceptions, while systematic technical validation of the accuracy and consistency of responses is less frequently provided.

To better understand how generative AI was applied across all included studies, Table 3 summarizes the specific objectives and key findings of each study. This focused synthesis highlights not only what the researchers set out to investigate but also the key findings and insights from their research.

Most studies have focused on investigating the effectiveness of AI chatbots, especially ChatGPT, in enhancing learning processes in a variety of engineering educational contexts, including programming, data science, and creative disciplines [23, 30, 33, 35, 37, 39, 42, 45]. A primary focus across the reviewed studies was enhancing personalized learning experiences [30, 41], improving learning outcomes, and supporting students in a range of academic tasks [33, 40, 42, 43] and integrating AI chatbots like ChatGPT into flipped learning strategies to

TABLE 3 | Overview of study aims and results.

Authors	Study aim	Key findings
Albdrani and Al-Shargabi (2023) [30]	To investigate the effectiveness of ChatGPT in providing personalized learning experiences for data science education, specifically focusing on deep learning.	ChatGPT with the 5E model increased student engagement by providing personalized support. The results showed improved learning outcomes with ChatGPT and the 5E model. The students felt supported and encouraged by the personalized learning experience with ChatGPT.
Bravo and Cruz-Bohorquez (2024) [23]	To examine the influence of using AI chatbots on learning within the context of engineering education across five different courses.	AI chatbots help students recall previous knowledge and learn new concepts. They act as tutors, explaining concepts and adapting to individual needs. Students had difficulty assessing the reliability of the information provided by chatbots, despite some awareness.
Caccavale et al. (2024) [32]	To investigate the use of AI models, specifically Large Language Models, as virtual tutors in chemical engineering education.	Most students are unsure about replacing teachers with Virtual Tutors (VTs) in specific courses. The majority of students worry about the quality of VT responses. Both students and industry/academia practitioners generally believe VTs should support teachers rather than replace them, and they could reduce teacher workload and improve information quality. Implementing AI in education requires careful attention to ethics, including data privacy and fairness.
Duhaylungsod and Chavez (2023) [33]	To analyze how AI models can support students in making ideas and innovations and developing creative ideas.	Conversational AI models help students with tasks such as writing, programming, and idea development. They provide instant feedback and real-time information, saving research time. However, overreliance on AI can reduce creativity, critical thinking, and problem-solving skills. The convenience of AI can be addictive, reducing students' motivation to think independently.
Elkefi et al. (2024) [34]	To examine the opportunities and challenges associated with ChatGPT utilization among engineering students in developing nations, using both open-ended interview data and a mixed-methods approach.	The internet access and ease of use of technology are important factors that enable the widespread use of ChatGPT among students. The main concerns that may limit the continued use of ChatGPT relate to accuracy, privacy, dependency, and potential negative academic consequences.
Gasaymeh and AlMohtadi (2024) [35]	To examine the impact of the Flipped Interactive Learning (FIL) instructional model based on ChatGPT on students' programming skills in a large undergraduate programming class.	Both the conventional Flipped Learning (FL) and the Flipped Interactive Learning instructional models significantly improved students' programming skills in Visual Basic 6. The FIL instructional model, which integrated ChatGPT to enhance interactivity, showed a slightly higher mean posttest score compared to the conventional FL model, suggesting the FIL model may be more effective.

(Continues)

TABLE 3 | (Continued)

Authors	Study aim	Key findings
Groothuijsen et al. (2024) [36]	To identify how to adapt programming education in higher engineering education following students' use of AI chatbots.	Students used ChatGPT in six ways, including error checking, conceptual understanding, code optimization, and code generation. Students' use of ChatGPT did not fully align with the teacher's expectations, as they used it more for explanations and debugging rather than code generation. The teacher perceived a decrease in students' learning, particularly in the pragmatic aspects of programming, despite their learning of syntax and semantics.
Huesca et al. (2024) [37]	To evaluate how ChatGPT complements and enriches the traditional flipped learning strategy in higher education, particularly in engineering courses.	ChatGPT-enhanced flipped learning outperformed traditional video-based approaches. ChatGPT provides interactive, tailored tutoring support. Real-time explanations and examples deepen conceptual understanding.
Humble et al. (2024) [38]	To identify potential opportunities and threats of ChatGPT for programming education.	ChatGPT could have significant implications for programming and higher education. ChatGPT could help teach and learn, potentially opening the way to new educational practices. ChatGPT poses a threat because it allows cheating and conceals a lack of understanding. The potential of ChatGPT requires teachers to rethink how they assess basic skills.
Mustapha et al. (2024) [39]	To track the landscape of issues surrounding GenAI tools and elucidate opportunities and limitations for enhancing mechanical engineering education and professional practices.	The study presents the first comprehensive assessment of GenAI in mechanical engineering, focusing on the implications for various aspects of the field. Experiments with GenAI tools have revealed their capabilities and limitations, indicating their imminent integration into mechanical engineering education and practice.
Neyem et al. (2024) [40]	To introduce and evaluate an AI Knowledge Assistant designed to enhance software engineering students' ability to gather and utilize knowledge effectively in capstone project courses.	Students preferred large language models for queries, either directly or enriched with insights from local repositories. Using LLMs significantly improved student comprehension compared to traditional lessons learned (LL) repositories. Integrating LLMs into project tracking platforms improved readability and aligned language styles with the student experience.
Pavlenko and Syzenko (2024) [41]	To explore students' experiences using ChatGPT for learning purposes and investigate their satisfaction levels with the support provided by this tool.	Students mostly use ChatGPT for information retrieval and language-related tasks such as text editing. The study highlights the need for individually tailored strategies to meet the different needs of academic disciplines. Business, engineering, and IT students report high satisfaction with using ChatGPT for learning purposes.

(Continues)

TABLE 3 | (Continued)

Authors	Study aim	Key findings
Pham et al. (2023) [42]	To explore the potential of using ChatGPT as an AI-assisted learning tool in an introductory course on embedded systems.	ChatGPT has significant potential to assist students in an introductory course on embedded systems by providing valuable insights and personalized support. Students need to be trained in how to use ChatGPT effectively. Lecturers should adapt courses to utilize ChatGPT while maintaining academic integrity.
Rajabi et al. (2024) [43]	To explore student and faculty perspectives on the use of ChatGPT in postsecondary education settings, focusing on how it could be integrated and how faculty might respond to this technology.	Clear guidelines and reporting are needed for students to use ChatGPT to protect academic integrity. Opinions about ChatGPT in higher education are mixed, with both advantages and disadvantages. Educators need to align the use of ChatGPT with academic standards.
Randall et al. (2023) [44]	To investigate the impact of AI on software engineering education and explore how AI tools should be integrated into curricula.	Detecting code written by AI is not always possible, so there is a shift towards tolerating the use of AI. AI should be integrated into student workflows to increase efficiency and productivity, and curricula should fully implement AI.
Remoto (2024) [45]	To assess the advantages and disadvantages of using AI chat models in learning mathematics based on the experiences of mathematics-oriented students.	AI chat models can accurately respond to chat prompts in problem-solving and explanations, reinforcing mathematical learning. These models sometimes provide inaccurate results or incorrect methods, which can be corrected upon re-running the prompts. AI models are more useful for essay tasks than mathematical solutions.

increase interactivity and conceptual understanding [35, 37]. Key findings included improved learning outcomes [30, 35, 37], increased student engagement and motivation [30], and the development of higher-order skills such as critical thinking and conceptual understanding [33, 37]. The research consistently highlighted the potential of AI chatbots to provide immediate, personalized feedback and support [30, 33, 40, 42]. At the same time, multiple studies emphasized the need for proper implementation, student guidance, and pedagogical alignment to ensure responsible and effective use [37, 40, 41, 43, 45]. Several studies suggested that educators may need to adjust their teaching methods and assessment strategies to effectively integrate AI while maintaining academic standards [43, 44]. However, some studies raised concerns about the potential for students to over-rely on AI tools, which could undermine the development of critical thinking and problem-solving skills [23, 33]. Although artificial intelligence has educators' potential to reduce the workload and offer personalized assistance, skepticism still remains about its reliability and broader ethical implications, including data privacy, fairness, and academic integrity [23, 32, 43]. Furthermore, a misalignment is evident between how students and educators use or perceive AI tools. For example, students tend to use ChatGPT primarily for explanations and debugging rather than for code generation, which some instructors interpret as a reduction in the depth of applied learning [36]. Additionally, while students who used ChatGPT often reported learning more efficiently and produced higher-quality solutions in less time, instructors often noted a decline in code quality and overall learning outcomes [36].

Although the reviewed articles examined various aspects of integrating generative artificial intelligence into engineering studies, the results are presented below to directly answer the research questions raised in this study.

RQ1. *How do chatbots and generative AI technologies affect students' learning outcomes, engagement, and skills development in engineering courses?*

Table 4 provides a comprehensive overview of the observed learning outcomes, engagement methods, and skill development in the disciplines. By comparing the results of several studies, this summary highlights both the potential benefits and challenges of integrating AI tools into educational practice.

In their study, Albdrani and Al-Shargabi [30] found that the implementation of ChatGPT in conjunction with the 5Es model had a positive impact on student learning outcomes. The

TABLE 4 | The impact of AI technologies on student learning outcomes, engagement, and skill development.

Paper	Learning outcomes	Engagement method	Skills development
Albdrani and Al-Shargabi (2023) [30]	Improved outcomes through tailored, active learning support	Increased participation, active engagement	N/A
Bravo and Cruz-Bohorquez (2024) [23]	Improved programming mastery and efficiency	Increased engagement through personalized learning and efficiency	Coding, critical thinking
Caccavale et al. (2024) [32]	Improved understanding and performance	Increased participation, easier question-asking	Critical thinking
Duhaylungsod and Chavez (2023) [33]	Enhanced task completion and innovation	Increased engagement through personalized learning and time efficiency	Innovation and creativity
Elkefi et al. (2024) [34]	Improved performance and continuous learning	Increased engagement through accessible AI support	N/A
Gasaymeh and AlMohtadi (2024) [35]	Improved programming skills	Increased participation through personalized attention	Programming skills enhanced
Groothuijsen et al. (2024) [36]	Syntax and semantics understanding; Pragmatics insufficient	Decreased collaboration	Mixed impact on programming skills; improved syntax and semantics, insufficient pragmatics
Huesca et al. (2024) [37]	Normalized learning gain improvement	Increased participation, more independent learning, stimulated interest	Problem-solving, conceptual understanding
Humble et al. (2024) [38]	N/A	Enhanced engagement through critical thinking	N/A
Mustapha et al. (2024) [39]	Correctness of answers in analytical reasoning tasks, errors in unit conversions and multiplication	N/A	N/A
Rajabi et al. (2024) [43]	Enhanced research support but potential reduction in learning	N/A	N/A
Randall et al. (2023) [44]	N/A	Increased engagement through improved perceptions and productivity	Improved coding efficiency
Remoto (2024) [45]	Enhanced mathematical understanding and problem-solving support	Increased engagement through accessibility and practicality	N/A

Abbreviation: N/A, data not provided in the article.

personalized ChatGPT guidance contributed to improved learning by addressing individual needs and providing opportunities for active application of knowledge. Regarding engagement, the authors noted that the integration of ChatGPT had a positive impact on student participation because its personalized nature encouraged a sense of responsibility and agency in the learning process. A study by Caccavale et al. [32] aimed to enhance the learning experience and increase the engagement of chemical engineering students by providing stimulating materials and games. They hypothesized that students would be more engaged by the gamified nature of the virtual tutor and would be less afraid to ask more questions. Their study also emphasizes the need for students to develop critical thinking skills and a comprehensive understanding of the principles of LLMs. Furthermore, the use of AI models can promote students' creativity and innovative thinking by providing new ideas and expanding their broadening concepts [33]. Elkefi et al. [34] provided insights into the usage patterns of ChatGPT, reporting that 58.46% of users used it frequently, while only 5.75% used it rarely. The high percentage of users with constant access to the internet (97.70%) and comfortable using technology (87.36%) indicates greater engagement due to the available AI assistance.

The studies show that integrating AI tools into programming education can significantly enhance student learning outcomes. Huesca et al. [37] reported improved normalized learning achievement, increased engagement, more independent learning, and increased student interest when using ChatGPT with flipped learning strategies. This led to improved problem-solving skills and conceptual understanding. Gasaymeh and AlMohtadi [35] studied the effectiveness of flipped learning and flipped interactive learning models in programming instruction. They found that both models significantly improved learning outcomes, with FIL showing superior results. The FIL group achieved a higher average normalized learning gain (52.85%) compared to the FL group (41.29%). Bravo et al. [23] reported that students found AI chatbots useful in acquiring prior programming knowledge, dispelling doubts raised during previous lessons and facilitating the acquisition of new concepts. They observed that students used AI chatbots to optimize learning time, answer specific questions more effectively, and save time searching for information. The study also found that AI chatbots contributed to the development of critical thinking skills as students learned to detect errors and carefully check their answers. However, Groothuijsen et al. [36] observed mixed effects on the development of programming skills. Although the students improved their understanding of syntax and semantics, their pragmatic skills were insufficient. The authors also observed a decrease in student collaboration when using ChatGPT, as students tended to rely on the AI tool rather than interact with their peers. Similarly, Randall et al. [44] emphasized the importance of teaching students to develop general application strategies rather than relying solely on the AI-generated code. Research revealed that a large majority of students in their institution's software engineering courses had incorporated AI tools into their programming practices. The researchers recommended designing curricula that encouraged or potentially required students to independently formulate the basic structure and approach of their programs.

Although the integration of AI tools has shown potential in programming education, similar benefits and challenges have been observed in mathematics education. Remoto [45] investigated the use of AI models in mathematics education and found that students found these tools accessible and practical for understanding mathematical concepts. The study reported improved mathematical understanding and problem-solving support, suggesting that AI models may offer students additional resources to explore mathematical ideas. However, the study also highlighted potential shortcomings, noting cases where AI models provided inaccurate or incomplete answers to mathematical queries.

Rajabi et al. [43] presented mixed findings on the impact of ChatGPT on learning, highlighting its ability to enhance learning experiences and provide research support. However, they also warned that inappropriate use of the tool could potentially hinder learning progress. Complementing these findings, Mustapha et al. [39] examined the accuracy of AI tools in analytical reasoning tasks, specifically noting errors in unit conversions and multiplication.

RQ2. *What are the challenges of using chatbots and generative AI in higher engineering education?*

Studies identified a range of challenges related to the use of chatbots and generative AI in higher education, spanning technical, pedagogical, and ethical areas. Tools such as ChatGPT often struggled to cope with complex, subject-specific content, leading to oversimplified or inaccurate results. From a pedagogical perspective, overreliance on these tools has been associated with reduced critical thinking, creativity, and collaboration among students. Ethical issues such as plagiarism, cheating, and data privacy risks are also commonly reported. Table 5 presents the main challenges identified in the reviewed studies.

The integration of AI tools into educational contexts poses a number of interrelated and multifaceted challenges. Quality and accuracy issues are often raised, with studies reporting incorrect assessments [30], errors in the generated code [23], and general accuracy issues [34]. Generative AI has occasionally provided incorrect evaluations of student responses and sometimes repeated or omitted steps, requiring teacher intervention [30]. Generative AI effectiveness is highly dependent on the quality of user input, which can be disadvantageous for students who struggle to formulate clear queries. The complexity of Generative AI responses can also be difficult for students to understand, sometimes requiring additional explanations [23].

There are some concerns about the impact of ChatGPT on critical thinking and creativity, with Duhaylungsod & Chavez [33] warning that overreliance on generative AI may discourage these skills. This concern is also mentioned by Huesca et al. [37], who stress the need to ensure that students do not become overly dependent on ChatGPT. Overreliance on ChatGPT is a recurring challenge [34, 36], which can lead to a reduced investment of time in learning and skill development. The use of the tool appears to have a negative impact on student collaboration and peer learning [36], as students rely on ChatGPT rather than working with peers, which can replace valuable peer interactions. Groothuijsen et al. [36] also observed insufficient learning of pragmatic aspects of programming and a decrease in code quality.

Humble et al. [38] highlight that ChatGPT can hinder learning by facilitating cheating and masking a lack of understanding and highlights the need to reconsider the assessment of basic skills in terms of ChatGPT abilities. Academic integrity issues are prominent, including difficulties in detecting plagiarism [43], challenges in preventing the use of AI in unsupervised situations, and ensuring equal access to the most powerful AI tools [44]. There are also important ethical and privacy issues [32, 43] related to data privacy, bias in AI responses, and the need for fact checking. In addition, potential challenges include how to avoid encouraging laziness [34], and copyright issues related to image-based generative AI tools [39].

Several authors identified technological challenges related to ChatGPT's limited depth of understanding, simplification of complex topics, and inconsistent quality of responses. Mustapha et al. [39] noted that ChatGPT does not understand specific technical details well enough and tends to oversimplify complex engineering concepts. Neyem et al. [40] identified challenges related to improving the efficiency of the LLM API, improving the clarity and relevance of responses, and addressing issues such as misleading information, insufficient detail, and lack of contextual information. Pham et al. [42] noted that complex questions often require multiple follow-up steps and that some responses initially provide misleading information.

TABLE 5 | Challenges of generative AI in engineering education.

Paper	Challenges
Albdrani and Al-Shargabi [30]	Incorrect evaluations by ChatGPT
Bravo and Cruz-Bohorquez [23]	Dependency on query clarity Complexity of responses Difficulty in critical evaluation Errors in generated code and inaccurate answers
Caccavale et al. [32]	Ethical and fairness concerns Privacy risks Quality of responses
Duhaylungsod and Chavez [33]	Overreliance on the tool discourages creativity and critical thinking
Elkefi et al. [34]	Accuracy issues Privacy concerns Laziness promotion Dependence on the tool
Groothuijsen et al. [36]	Decrease in code quality Negative influence on pair programming Decrease in student collaboration Insufficient learning of pragmatics Student overreliance on ChatGPT
Huesca et al. [37]	Ensuring students do not become overly reliant on ChatGPT Cognitive skill necessity
Humble et al. [38]	Language quality differences Inaccurate responses Cheating potential Interferes with learning
Mustapha et al. [39]	Errors in computation Challenges in developing informal learning skills Symmetry and geometric errors in image generation
Neyem et al. [40]	Response time issues Inconsistent accuracy and relevance of responses Lessons learned are not described in sufficient detail No error messages Lack of contextual information in code examples PLM performance differences
Pham et al. [42]	Inconsistent AI responses results Challenges in maintaining assessment integrity Difficulty evaluating AI-generated content
Rajabi et al. [43]	Academic integrity challenges Plagiarism detection Ethical concerns Bias in AI responses Accuracy issues in grading Data privacy concerns
Randall et al. [44]	Monitoring AI use in practical assignments Detecting AI-written code Identifying mixed human-AI code Handling incorrect AI information Ensuring good coding practices Accessibility issues with AI tools Preventing educational inequality
Remoto [45]	Inability to provide sources and handle complex mathematical symbols Limited applicability for mathematical tasks Reduced critical thinking and research skills Inaccuracy and inconsistency in responses

Furthermore, Remoto [45] highlighted limitations in handling complex mathematical symbols and the inability of AI models to provide sources for their solutions.

5 | Discussion

An examination of the reviewed studies shows that none of them explicitly report whether supervised learning, unsupervised learning, or reinforcement learning was employed. All empirical studies treat chatbots and generative AI systems as existing tools, focusing mainly on experimental design, data collection, and educational impact evaluation rather than on the technical training mechanisms of the models.

The reviewed studies also indicate that chatbots are increasingly used to support learning activities in engineering education, including programming, problem solving, and feedback provision [23, 35, 37, 44].

Despite these educational benefits, AI chatbots also introduce a range of pedagogical and practical challenges. AI chatbots may face limitations in understanding and appropriately responding to emotions and contexts, which can result in inappropriate responses. Likewise, there is a risk of creating a cold and mechanical environment if interaction with chatbots becomes exclusive, which could negatively affect the warmth and human dimension in the class environment [12]. If students become overly dependent on AI chatbots for consultation and academic support, their engagement in direct interactions with professors may decrease [23]. As a result, professors may have fewer opportunities to closely monitor students' progress and guide them in their learning process [23]. Moreover, educators may find it difficult to accurately assess student performance when using AI chatbots for assignments and assessments, thus preventing effective tracking of student learning issues [23].

Beyond pedagogical concerns, a cross-analysis of the reviewed literature indicates that LLM-based chatbots exhibit recurring limitations when applied to multi-step engineering problem solving. Several studies report numerical inaccuracies, incorrect calculations, and procedural errors in tasks involving mathematics, programming, and analytical reasoning, including unit conversion mistakes, flawed arithmetic operations, and incomplete solution steps [32, 34, 39, 45]. In programming and software engineering contexts, chatbots were found to support syntax and basic semantics but struggle with pragmatic understanding, complex task decomposition, and maintaining solution quality across multi-step workflows [36, 38, 42, 44]. Multiple studies further indicate that chatbots may generate incorrect or misleading solutions with high confidence and without internal verification, requiring students to critically evaluate outputs, particularly in engineering tasks that demand sequential reasoning and procedural accuracy [34, 45].

Preventing plagiarism plays a critical role in protecting academic integrity and the ethical foundation of education [12]. ChatGPT can generate plausible data, but it does not reference the source of the data, which leads to concerns about the integrity of the information [32]. Beyond ethical and integrity concerns, ChatGPT also faces human-interaction challenges that highlight its limitations as a teaching tool. While efforts are still underway to advance NLP technology to detect human emotions through tones and expressions, technology still has limitations and cannot serve as a replacement for humans. Though ChatGPT can offer consistent support and guidance, it cannot surpass the level of interaction through traditional teaching methods provided by humans [32].

In addition to the studies included in this review, several recent publications compared the performance of different generations of GPT-based chatbot systems in educational contexts. ChatGPT-3.5 demonstrated robust functionality on structured learning tasks, successfully completing the tasks in 45 out of 51 cases and significantly improving the quality of students' summaries ($t = -9.220$, $p < 0.05$) [46]. The more advanced models performed better in all areas: GPT-4 achieved 92.1% accuracy on advanced cardiovascular life support assessments and 96% accuracy on basic life support tests, and also outperformed engineering students on physics concept tasks [47]. Customized GPT systems based on GPT-4.0 were found to enhance positive emotions, situational interest, and self-efficacy while reducing cognitive load, although no significant differences were found in learning outcomes ($p = 0.750$) [48]. In direct model comparisons, ChatGPT-4o achieved significantly better results than ChatGPT-4 and outperformed other tested models on graph analysis tasks, while Falcon 7B demonstrated competitive results on university FAQ applications (BERTScore = 0.771; ROUGE-1 = 0.671) [49]. These studies collectively found that GPT-4 and ChatGPT-4o are the best-performing systems among currently used educational chatbots, with consistent improvements across model generations in terms of accuracy and task efficiency.

Beyond the studies included in this review, recent examples also illustrate the potential of AI-supported learning in domain-specific engineering environments. For example, in the Aladdin modeling environment [50], generative AI tools such as chatbots can be used to support learning in energy-efficient building design by helping engineering students define design constraints, explore alternative strategies, and interpret modeling results. In this context, AI acts as a thinking partner, helping with reasoning, reflection, and outcome-oriented decision-making during the design process.

The additional literature also draws attention to the challenges educators face when implementing AI-supported learning environments. As the implementing agents in direct interaction with students, teachers hold a critical role in achieving the goal of fostering AI literacy among students within their teaching subjects. However, teachers face several practical difficulties in designing, managing, and implementing AI-integrated lessons. Teachers reported to perceive that they needed to be supported with more time and appropriate resources or activities for AI-integrated lessons, even if they were eager to implement AI lessons [51].

Additionally, teachers may worry about integrating AI into their practices, as well as ethical considerations like data privacy, algorithmic bias, academic integrity, and teacher autonomy. Moreover, without sufficient institutional support and alignment with teaching goals, these concerns can create implementation challenges [12, 22].

While institutions may provide support in terms of policy, infrastructure, or general guidelines, this support often falls short of addressing the specific, actionable needs of teachers who adopt AI tools, making AI less relevant to day-to-day teaching contexts. Furthermore, teachers' autonomy in selecting technologies also plays a significant role, as they prioritize tools that reflect their unique teaching styles, relying on personal experiences or peer recommendations rather than institutional directives [52, 53].

For instance, although AI tools may in certain regards raise teachers' confidence by providing supporting justifications for teaching plans, this feeling was juxtaposed by some teachers experiencing uncertainty due to the occasionally unreliable AI outputs and unresolved evidential questions. Moreover, generative AI tools were used by teachers to find recent research and pedagogical techniques, in contradistinction, some teachers encountered the problem that the technologies sometimes produced outdated or inaccurate information [54].

Educators face comparable challenges, with anxieties often stemming from uncertainties about how to effectively integrate AI into the classroom and feeling overwhelmed with its technical demands. Excessive reliance on AI or perceptions of constant monitoring in classrooms may contribute to educator anxiety and stress [55].

Additionally, some studies suggest that teachers often lack the necessary understanding of AI technologies and may feel disempowered by their use, which contributes to a decline in self-efficacy and reluctance to adopt these tools. For instance, despite positive perceptions of AI's potential benefits, only a minority of primary teachers have effectively implemented AI and ChatGPT in their classrooms, often due to a lack of readiness and knowledge [56].

Overall, the reviewed studies demonstrate a consistent pattern across engineering education contexts. While the specific applications of chatbots and generative AI varied across disciplines, most studies reported positive effects on student engagement, personalized learning, programming support, and problem-solving activities. At the same time, studies highlight concerns related to academic integrity, overreliance on AI-generated content, reduced student collaboration, and limitations in handling complex engineering tasks. While educational benefits were reported fairly consistently across studies, findings regarding the impact of AI on skills varied, with some studies reporting enhanced critical thinking and conceptual understanding, while others raised concerns about reduced independent thinking and overreliance on AI assistance.

6 | Conclusions

The integration of generative AI tools into higher education, particularly in engineering disciplines, has shown significant potential to improve learning outcomes, support skill development, and promote student engagement. These tools can help students understand complex technical content, promote self-directed learning, and foster creativity and problem-solving, competencies that are important in engineering studies. Students have also benefited from the direct and personalized impact of AI-generated feedback, which can make learning more accessible and interactive. Furthermore, AI has shown the potential to optimize learning time and improve understanding in both theoretical and practical learning contexts.

The identified challenges highlight the need for a cautious and contextual approach to integrating chatbots and generative AI into higher education. Rather than simply relying on the capabilities of these tools, institutions need to critically assess how their use aligns with pedagogical goals, contributes to skill development, and supports academic integrity. This is particularly important in engineering education, where learning is based on technical accuracy, problem solving, and practical application. Particular attention should be paid to reducing overreliance on generative AI, ensuring that students are active participants in the learning process, and preserving opportunities for collaborative and experiential learning. In addition, addressing technical limitations, clarifying ethical boundaries, and providing guidance on responsible use will be essential to maximizing educational value and minimizing risks. In the context of engineering programs, it is also crucial to recognize the limitations of AI in processing content specific to the subject area and to adapt assessment practices accordingly.

This review contributes to the growing body of research on AI-enhanced engineering education by synthesizing evidence on the use of chatbots and generative AI tools in engineering learning environments. The findings provide a consolidated overview of their educational benefits, implementation challenges, and implications for future research and educational practice.

One limitation of this study is the relatively narrow availability of studies due to the chosen time period and the specific focus on engineering studies. Since the widespread use of generative AI tools in higher education, especially in the face-to-face classroom, is relatively new, the number of peer-reviewed studies that met the inclusion criteria was limited. This limitation may have excluded relevant but new work that is still in preprint form or is not published in selected academic sources, and the findings should be interpreted in light of the rapid evolution of the field and the expanding research base. In addition, although all studies focused on engineering, they varied in design, sample size, and assessment methods, which may affect the consistency and comparability of the findings. As this review employed a qualitative rather than a quantitative approach, no statistical summary of study results was conducted. This limitation may reduce the ability to compare the strength of effects across studies. Furthermore, the lack of a formal quality assessment system (e.g., GRADE or MMAT) limits the ability to systematically assess the reliability of the evidence included. Therefore, the overall strength of the findings should be interpreted in consideration of these methodological differences. As the use of generative AI in education continues to grow, future work should examine its long-term impact on learning outcomes, skill development, and student engagement in engineering. It would be useful to analyze more research to explore its role in specific engineering disciplines and to better understand how it impacts critical thinking and collaboration.

Acknowledgments

The scientific paper is funded by the international project "Renewable Energy for Bolstering Ukraine's Infrastructure by Learning and Design," REBUILD, No.: IMPRESS246.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. I. J. Akpan, Y. M. Kobara, J. Owolabi, A. A. Akpan, and O. F. Offodile, "Conversational and Generative Artificial Intelligence and Human–Chatbot Interaction in Education and Research," *International Transactions in Operational Research* 32, no. 3 (2025): 1251–1281.

2. S. Sowa, C. Brown, T. H. Choi, and R. Newman, "Examining How Generative AI Tools Benefit and Challenge Teachers' Research-Informed Practice," *European Journal of Teacher Education* 48, no. 5 (2025): 1–20.

3. Z. Huang, A. Peng, T. Yang, S. Deng, and Y. He, "A Design-Based Learning Approach for Fostering Sustainability Competency in Engineering Education," *Sustainability* 12, no. 7 (2020): 2958.

4. A. Kolmos, J. E. Holgaard, H. W. Routhe, M. Winther, and L. Bertel, "Interdisciplinary Project Types in Engineering Education," *European Journal of Engineering Education* 49, no. 2 (2024): 257–282.

5. A. Johri, A. S. Katz, J. Qadir, and A. Hingle, "Generative Artificial Intelligence and Engineering Education," *Journal of Engineering Education* 112, no. 3 (2023): 572–577.

6. G. Ilieva, T. Yankova, S. Klisarova-Belcheva, A. Dimitrov, M. Bratkov, and D. Angelov, "Effects of Generative Chatbots in Higher Education," *Information* 14, no. 9 (2023): 492.

7. J. Qadir, "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education," in *2023 IEEE Global Engineering Education Conference (EDUCON)* (2023), 1–9.

8. S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are We There Yet?—A Systematic Literature Review on Chatbots in Education," *Frontiers in Artificial Intelligence* 4 (2021): 654924.

9. M. M. Chan, H. R. Amado-Salvatierra, R. Hernandez-Rizzardini, and M. De La Roca, "The Potential Role of AI-Based Chatbots in Engineering Education. Experiences From a Teaching Perspective." in *2023 IEEE Frontiers in Education Conference (FIE)* (2023), 1–5.

10. N. H. Mthombeni, R. Maladzhi, K. Moloi, T. Mashifana, M. Tsoeu, and F. J. Nemavhola, "AI Chatbots as Inherent Tools for Comprehensive Learning of Engineering Design: A Systematic Review," in *2023 World Engineering Education Forum-Global Engineering Deans Council (WEEF-GEDC)* (2023), 1–6.

11. J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the Use of Chatbots in Education: A Systematic Literature Review," *Computer Applications in Engineering Education* 28, no. 6 (2020): 1549–1565.

12. A. Tlili, B. Shehata, M. A. Adarkwah, et al., "What If the Devil Is My Guardian Angel: Chatgpt as a Case Study of Using Chatbots in Education," *Smart Learning Environments* 10, no. 1 (2023): 15.

13. J. Swacha and M. Gracel, "Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications," *Applied Sciences* 15, no. 8 (2025): 4234.

14. M. I. Ciolacu, C. Marghescu, B. Mihailescu, and P. Svasta, "Does Industry 5.0 Need an Engineering Education 5.0? Exploring Potentials and Challenges in the Age of Generative AI," in *2024 IEEE Global Engineering Education Conference (EDUCON)* (2024), 1–10.

15. K. Moloi, R. W. Maladzhi, F. J. Nemavhola, N. H. Mthombeni, M. S. Tsoeu, and T. Mashifana, "The Risks Associated With AI Chatbots in Teaching Future Engineering Graduates: A Systematic Review," in *2023 World Engineering Education Forum-Global Engineering Deans Council (WEEF-GEDC)* (2023), 1–6.

16. P. M. Simelane and J. Kittur, "Use of Generative Artificial Intelligence in Teaching and Learning: Engineering Instructors' Perspectives," *Computer Applications in Engineering Education* 33, no. 1 (2025): e22813.

17. Z. Boussouf, H. Amrani, M. Z. Khal, and F. Daidai, "Artificial Intelligence in Education: A Systematic Literature Review," *Data and Metadata* 3 (2024): 288.

18. A. O. Ifelebuegu, P. Kulume, and P. Cherukut:, "Chatbots and AI in Education (AIEd) tools: The Good, the Bad, and the Ugly," *Journal of Applied Learning and Teaching* 6, no. 2 (2023): 332–345.

19. C. Kooli, "Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions," *Sustainability* 15, no. 7 (2023): 5614.

20. L. Labadze, M. Grigolia, and L. Machaidze, "Role of AI Chatbots in Education: Systematic Literature Review," *International Journal of Educational Technology in Higher Education* 20, no. 1 (2023): 56.

21. U. Dewan, A. Hingle, N. McDonald, and A. Johri, "Engineering Educators' Perspectives on the Impact of Generative AI in Higher Education," in *2025 IEEE Global Engineering Education Conference (EDUCON)*, April (2025), 1–10.

22. P. Lata, "Beyond Algorithms: Humanizing Artificial Intelligence for Personalized and Adaptive Learning," *International Journal of Innovative Research in Engineering and Management* 11, no. 5 (2024): 40–47.

23. F. A. Bravo and J. M. Cruz-Bohorquez, "Engineering Education in the Age of AI: Analysis of the Impact of Chatbots on Learning in Engineering," *Education Sciences* 14, no. 5 (2024): 484.

24. J. Liu, A. R. Rao, and M. Liu, "Establishing the Effectiveness and Acceptance of AI-Based Feedback in an Engineering Course," in *2025 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (2025), December, 1–8.

25. K. W. M. Siu, J. Zou, Y. Jiang, Z. Yang, K. Zhang, and T. Zhao, "Dynamic Scaffolding: Exploring the Role of Artificial Intelligence in Urban Design Education," *Frontiers of Urban and Rural Planning* 3, no. 1 (2025): 8.

26. M. Menekse, "Envisioning the Future of Learning and Teaching Engineering in the Artificial Intelligence Era: Opportunities and Challenges," *Journal of Engineering Education* 112 (2023): 578–582.

27. M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting With Educational Chatbots: A Systematic Review," *Education and Information Technologies* 28, no. 1 (2023): 973–1018.

28. C. D. B. Majorana, R. B. Gonçalves, F. L. A. Neto, and R. Z. Camargo, "Enhancing Administrative Efficiency in Higher Education With AI: A Chatbot Solution," *Review of Artificial Intelligence in Education* 3 (2022): e23.

29. M. D. L. Roca, M. M. Chan, A. Garcia-Cabot, E. Garcia-Lopez, and H. Amado-Salvatierra, "The Impact of a Chatbot Working as an Assistant in a Course for Supporting Student Learning and Engagement," *Computer Applications in Engineering Education* 32, no. 5 (2024): e22750.

30. R. N. Albdrani and A. A. Al-Shargabi, "Investigating the Effectiveness of ChatGPT for Providing Personalized Learning Experience: A Case Study," *International Journal of Advanced Computer Science & Applications* 14 (2023): 11.

31. "PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses," accessed June 29, 2025, https://www.prisma-statement.org.

32. F. Caccavale, C. L. Gargalo, K. V. Gernaey, and U. Krühne, "Towards Education 4.0: The Role of Large Language Models as Virtual Tutors in Chemical Engineering," *Education for Chemical Engineers* 49 (2024): 1–11.

33. A. V. Duhaylungsod and J. V. Chavez, "ChatGPT and Other AI Users: Innovative and Creative Utilitarian Value and Mindset Shift," *Journal of Namibian Studies* 33 (2023): 4367–4378.

34. S. Elkefi, A. Tounsi, and M. A. Kefi, "Use of ChatGPT for Education by Engineering Students in Developing Countries: A Mixed-Methods Study," *Behaviour & Information Technology* (2024): 1–17, https://doi.org/10.1080/0144929X.2024.2354428.

35. A. M. M. Gasaymeh and R. M. AlMohtadi, "The Effect of Flipped Interactive Learning (FIL) Based on ChatGPT on Students' Skills in a Large Programming Class," *International Journal of Information and Education Technology* 14, no. 11 (2024): 1516–1522.

36. S. Groothuijsen, A. van den Beemt, J. C. Remmers, and L. W. van Meeuwen, "Ai Chatbots in Programming Education: Students' Use in a Scientific Computing Course and Consequences for Learning," *Computers and Education: Artificial Intelligence* 7 (2024): 100290.

37. G. Huesca, Y. Martínez-Treviño, J. M. Molina-Espinosa, et al., "Effectiveness of Using ChatGPT as a Tool to Strengthen Benefits of the Flipped Learning Strategy," *Education Sciences* 14, no. 6 (2024): 660.

38. N. Humble, J. Boustedt, H. Holmgren, G. Milutinovic, S. Seipel, and A. S. Östberg, "Cheaters or AI-Enhanced Learners: Consequences of ChatGPT for Programming Education," *Electronic Journal of e-Learning* 22, no. 2 (2024): 16–29.

39. K. B. Mustapha, E. H. Yap, and Y. A. Abakr, "Bard, ChatGPT and 3DGPT: A Scientometric Analysis of Generative AI Tools and Assessment of Implications for Mechanical Engineering Education," *Interactive Technology and Smart Education* 21, no. 4 (2024): 588–624.

40. A. Neyem, L. A. González, M. Mendoza, J. P. S. Alcocer, L. Centellas, and C. Paredes, "Towards an AI Knowledge Assistant for Context-Aware Learning Experiences in Software Capstone Project Development," *IEEE Transactions on Learning Technologies* 17 (2024): 1599–1614.

41. O. Pavlenko and A. Syzenko, "Using ChatGPT as a Learning Tool: A Study of Ukrainian Students' Perceptions," *Arab World English Journal Special Issue on ChatGPT* (2024): 252–264.

42. T. Pham, T. B. Nguyen, S. Ha, and N. T. N. Ngoc, "Digital Transformation in Engineering Education: Exploring the Potential of AI-Assisted Learning," *Australasian Journal of Educational Technology* 39, no. 5 (2023): 1–19.

43. P. Rajabi, P. Taghipour, D. Cukierman, and T. Doleck, "Unleashing ChatGPT's Impact in Higher Education: Student and Faculty Perspectives," *Computers in Human Behavior: Artificial Humans* 2, no. 2 (2024): 100090.

44. N. Randall, D. Wäckerle, N. Stein, D. Goßler, and S. Bente, "What an AI-Embracing Software Engineering Curriculum Should Look Like: An Empirical Study," *IEEE Software* 41, no. 2 (2023): 36–43.

45. J. P. Remoto, "ChatGPT and Other AIs: Personal Relief and Limitations Among Mathematics-Oriented Learners," *Environment and Social Psychology* 9, no. 1 (2024): 1–13.

46. C. K. Looi and F. Jia, "Personalization Capabilities of Current Technology Chatbots in a Learning Environment: An Analysis of Student-Tutor Bot Interactions," *Education and Information Technologies* 30, no. 10 (2025): 1–31.

47. S. Guizani, T. Mazhar, T. Shahzad, W. Ahmad, A. Bibi, and H. Hamam, "A Systematic Literature Review to Implement Large Language Model in Higher Education: Issues and Solutions," *Discover Education* 4, no. 1 (2025): 14165–14195.

48. J. Lademann, J. Henze, and S. Becker-Genschow, "Augmenting Learning Environments Using AI Custom Chatbots: Effects on Learning Performance, Cognitive Load, and Affective Variables," *Physical Review Physics Education Research* 21, no. 1 (2025): 010147.

49. K. Peyton, S. Unnikrishnan, and B. Mulligan, "*A Review of University Chatbots for Student Support*: *FAQs and Beyond*," *Discover Education* 4, no. 1 (2025): 21.

50. C. Xie, X. Ding, and R. Jiang, "Using Computer Graphics to Make Science Visible in Engineering Education," *IEEE Computer Graphics and Applications* 43, no. 5 (2023): 99–106.

51. J. Park, T. W. Teo, A. Teo, J. Chang, J. S. Huang, and S. Koo, "Integrating Artificial Intelligence into Science Lessons: Teachers' Experiences and Views," *International Journal of STEM Education* 10, no. 1 (2023): 61.

52. A. Hazzan-Bishara, O. Kol, and S. Levy, "The Factors Affecting Teachers' Adoption of AI Technologies: A Unified Model of External and Internal Determinants," *Education and Information Technologies* 30, no. 11 (2025): 15043–15069.

53. J. Kim, "Leading Teachers' Perspective on Teacher—AI Collaboration in Education," *Education and Information Technologies* 29, no. 7 (2024): 8693–8724.

54. Y. Song, L. Huang, L. Zheng, M. Fan, and Z. Liu, "Interactions With Generative Ai Chatbots: Unveiling Dialogic Dynamics, Students' Perceptions, and Practical Competencies in Creative Problem-Solving," *International Journal of Educational Technology in Higher Education* 22, no. 1 (2025): 12.

55. J. A. Delello, W. Sung, K. Mokhtari, J. Hebert, A. Bronson, and T. De Giuseppe, "AI in the Classroom: Insights From Educators on Usage, Challenges, and Mental Health," *Education Sciences* 15, no. 2 (2025): 113.

56. O. Filiz, M. H. Kaya, and T. Adiguzel, "Teachers and AI: Understanding the Factors Influencing AI Integration in K–12 Education," *Education and Information Technologies* 30, no. 13 (2025): 17931–17967.