ISSN 2409-2665
Journal of Logistics, Informatics and Service Science
Vol. 12 (2025) No. 3, pp. 264-282
DOI:10.33168/JLISS.2025.0316

Conversational AI in Education: A General Review of Chatbot Technologies and Challenges

Wasin Alkishri[1*], Jabar H Yousif [2], Yousuf Nasser Al Husaini[1] and Mahmood Al-Bahri[2]

[1] Faculty of Computer Studies, Arab Open University, Seeb, Oman

[2] Faculty of Computing and IT, Sohar University, Sohar, Oman

Wasan.k@aou.edu.om, jyousif@su.edu.om, yousufnaser@aou.edu.om, mbahri@su.edu.om

Abstract. The integration of AI-powered chatbots into educational contexts has accelerated rapidly over the past decade, offering scalable and personalized learning solutions. This study follows PRISMA process to synthesize findings from over 40 empirical and theoretical works published between 2017 and 2025, providing a systematic evaluation of the role of chatbots in formal education. A keyword-based word cloud analysis identifies dominant themes such as "student," "engagement," "AI," and "dialogue," reflecting a pedagogical shift toward AI-mediated, learner-centered environments. The results reveal that chatbots are increasingly leveraged for formative assessment, virtual tutoring, and administrative support, particularly in higher education, albeit with varying degrees of success. It comprehensively analyzed the selected studies and identified six significant research gaps. The paper also explores future directions, emphasizing the need for multilingual, multimodal, and ethically governed chatbot systems that integrate emotional intelligence and culturally sensitive content. Despite the growing market for educational AI and its demonstrated potential, significant challenges persist—ranging from digital access gaps to the absence of standardized evaluation frameworks. The study concludes by offering policy and design recommendations to ensure that chatbot technologies enhance, rather than hinder, equitable and meaningful learning.

Keywords: Chatbots, Educational technology, Personalized learning, Artificial intelligence, Student engagement.

1. Introduction

Artificial Intelligence (AI) has become a transformative force across multiple sectors, fundamentally altering how humans interact with machines. Among its most widespread applications is the chatbot—an AI-driven conversational agent designed to simulate dialogue with users through natural language processing (NLP) technologies (Zamora, 2017; Radziwill & Benton, 2017). The Oxford Dictionary (2023) defines a chatbot as "a computer program designed to simulate conversation with human users, especially over the Internet." Functioning as digital assistants, chatbots are now deployed far beyond entertainment, playing critical roles in customer service, healthcare, and increasingly, education (Nooruzzaman & Hussain, 2018). Their responsiveness, cost-effectiveness, and 24/7 availability make them particularly well-suited for educational environments that require scalable and personalized student support.

The conceptual roots of chatbots can be traced to Alan Turing's foundational work on machine intelligence, including his seminal 1950 paper, Computing Machinery and Intelligence, which introduced the Turing Test. This idea inspired early prototypes such as ELIZA (Weizenbaum, 1966), which mimicked a Rogerian psychotherapist, and PARRY (Colby, 1975), which simulated paranoid schizophrenia. As chatbot capabilities matured, ALICE (Wallace, 2009) leveraged Artificial Intelligence Markup Language (AIML) to advance conversation flow and won multiple Loebner Prizes. The early 2000s saw bots like SmarterChild flourish on messaging platforms (Shawar & Atwell, 2007), setting the stage for the rise of intelligent voice assistants: Siri (2011), Google Now (2012), Cortana (2014), Alexa (2014), and Google Assistant (2016) (Hoy, 2018; McTear, 2020). A paradigm shift occurred with the advent of transformer-based large language models—OpenAI's GPT-2 (2019), GPT-3 (2020), and the widely accessible ChatGPT (2022). The release of GPT-4 in 2023 and the multimodal GPT-4o in 2024 further empowered chatbot systems to process text, images, and voice simultaneously, moving these systems toward true conversational intelligence (OpenAI, 2023) as presented in Table 1.

The educational sector has witnessed a surge in chatbot adoption, driven by the growing need for digital transformation and personalized learning. A Pew Research Center survey (2025) reports that over 25% of U.S. teenagers now regularly use ChatGPT for academic tasks—double the figure from 2023. Simultaneously, AIPRM (2024) notes that more than a third of global educators have integrated chatbot tools into their classrooms. These range from virtual teaching assistants that handle routine inquiries to intelligent grading systems embedded within Learning Management Systems (LMS). Grand View Research (2024) estimates the AI-in-Education market was valued at USD 5.88 billion in 2024 and is projected to grow at a compound annual growth rate (CAGR) of 31% through 2030 as shown in Figure 1. This acceleration is fueled by three primary drivers: the proliferation of online learning, the demand for individualized educational pathways, and the need to automate administrative overhead.

The use of chatbots in formal education has flashed a wide range of reactions as powerful tools for improving accessibility and scaling support. Some others raise valid concerns about the ethics and educational value of using chatbots. This study examined the need to explore how these technologies can enhance teaching and learning rather than act as digital add-ons. As classrooms evolve through digital transformation, chatbots offer significant potential in providing instant feedback, 24/7 availability, multilingual communication, and additional support for students who might otherwise be left behind. They can also lighten the load for educators and make self-paced learning more achievable. However, even with their fast spread, the actual effectiveness of chatbots in education remains unclear. Current research is scattered, and we lack strong, consistent evidence about their impact on student learning, motivation, and educational equity.

Given the escalating integration of chatbot technologies in formal learning environments, a critical examination is needed to assess both their effectiveness and their limitations. Existing literature has often focused on technical development or anecdotal classroom applications, with insufficient empirical evaluation of educational value across regions and disciplines. This paper aims to address this gap by posing the central research question:

To what extent have chatbot technologies contributed to improving teaching and learning processes in formal educational settings between 2017 and 2025, and what are the major challenges and opportunities associated with their adoption?

Answering this question holds significant implications for educational policymakers, technology developers, and academic practitioners alike. It also provides a roadmap for the ethical, inclusive, and pedagogically sound integration of conversational AI in global education systems.

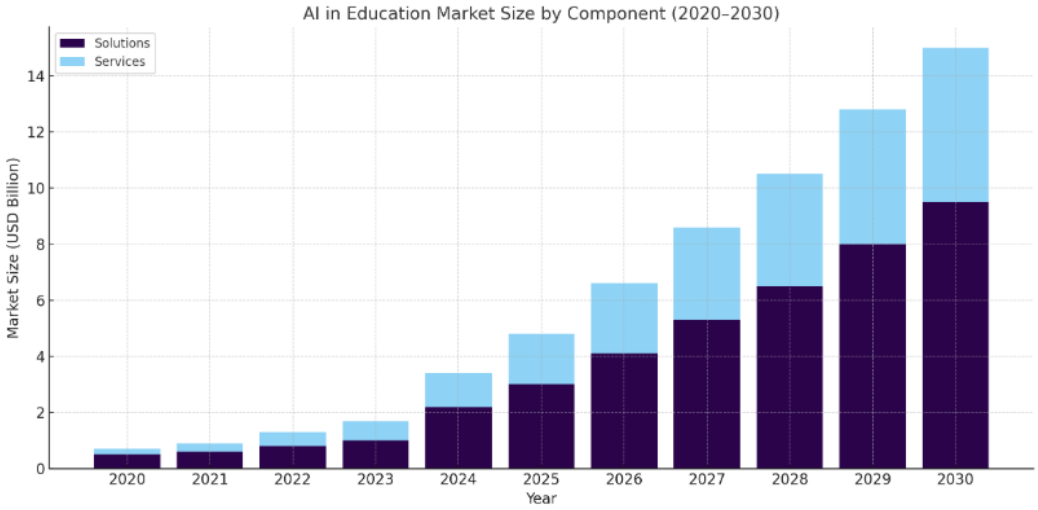

Fig.1. AI in the education market size 2022-2023- (Grand view research, 2024)

Table 1: Summary of chatbot history.

Year	Chatbot/Technology	Key Significance
1950	Turing Test	Introduced the concept of machine intelligence (the imitation game)
1966	ELIZA	First chatbot simulating a psychotherapist
1972	PARRY	Simulated a paranoid patient; early emotional modelling
1995	ALICE	Used AIML; won Loebner Prize for human-like dialogue
2001	SmarterChild	Popular early chatbot on messaging platforms
2011	Siri	First principal voice assistant on smartphones
2014	Amazon Alexa	Voice-controlled assistant for the smart home
2015	Facebook M	AI-human hybrid assistant; limited scalability
2016	Google Assistant	AI-powered assistant with contextual understanding
20217	Samsung Bixby	Samsung's voice assistant for device control
2022	Chat GPT	Public release of AI chatbot with wide capabilities
2023	GPT-4	An advanced generative AI model by OpenAI
2024	GPT-4o	Multimodal model handling text, image, and voice input

2. Method and Data

This paper followed the PRISMA systematic process (PRISMA, 2020) to ensure that high-quality, relevant literature studies contributed to the findings. The overall selection flow is depicted in Figure 2. It includes the following:

2.1. Search Methodology

The literature search was conducted across four major academic databases: Scopus, IEEE Xplore, SpringerLink, and ScienceDirect, complemented by Google Scholar, to ensure comprehensive results about Chatbots or conversational Al. Boolean operators were employed to construct the search string:

("Chatbot" AND "Education", "Chatbot" AND "Higher Education")

The search criteria focus on the studies published between 2017 and 2025. An initial collection of 200 articles was retrieved from all sources.

2.2. Inclusion and Exclusion Criteria

To ensure relevance and quality, the following inclusion criteria were applied:

- Studies focused on chatbots or conversational AI in education.
- Articles that contain empirical data, use cases, or system evaluations.
- Peer-reviewed Articles published in English and within the years (2017-2025).

Exclusion criteria included:

- Generic AI systems are unrelated to conversational learning support.
- Theoretical or opinion-based papers lack empirical evidence.
- Duplicates and non-peer-reviewed articles.

2.3. Data Extraction and Selection Strategy

Initially, Metadata, titles, abstracts, and keywords were reviewed. Articles not aligned with the research focus and non-English studies were excluded, reducing the studies to 100 research studies. The quality of the studies was evaluated based on their methodological stringency, clarity of objectives, and design quality, resulting in a limited collection of 80 studies. The articles were thoroughly reviewed, and studies lacking empirical evidence or with no full text were excluded, resulting in 40 final articles being included for analysis. Articles that solely discussed theoretical frameworks without application or those targeting corporate or healthcare chatbot use were excluded. The final dataset consists of 40 studies that provide both breadth and depth regarding chatbot technologies in education.

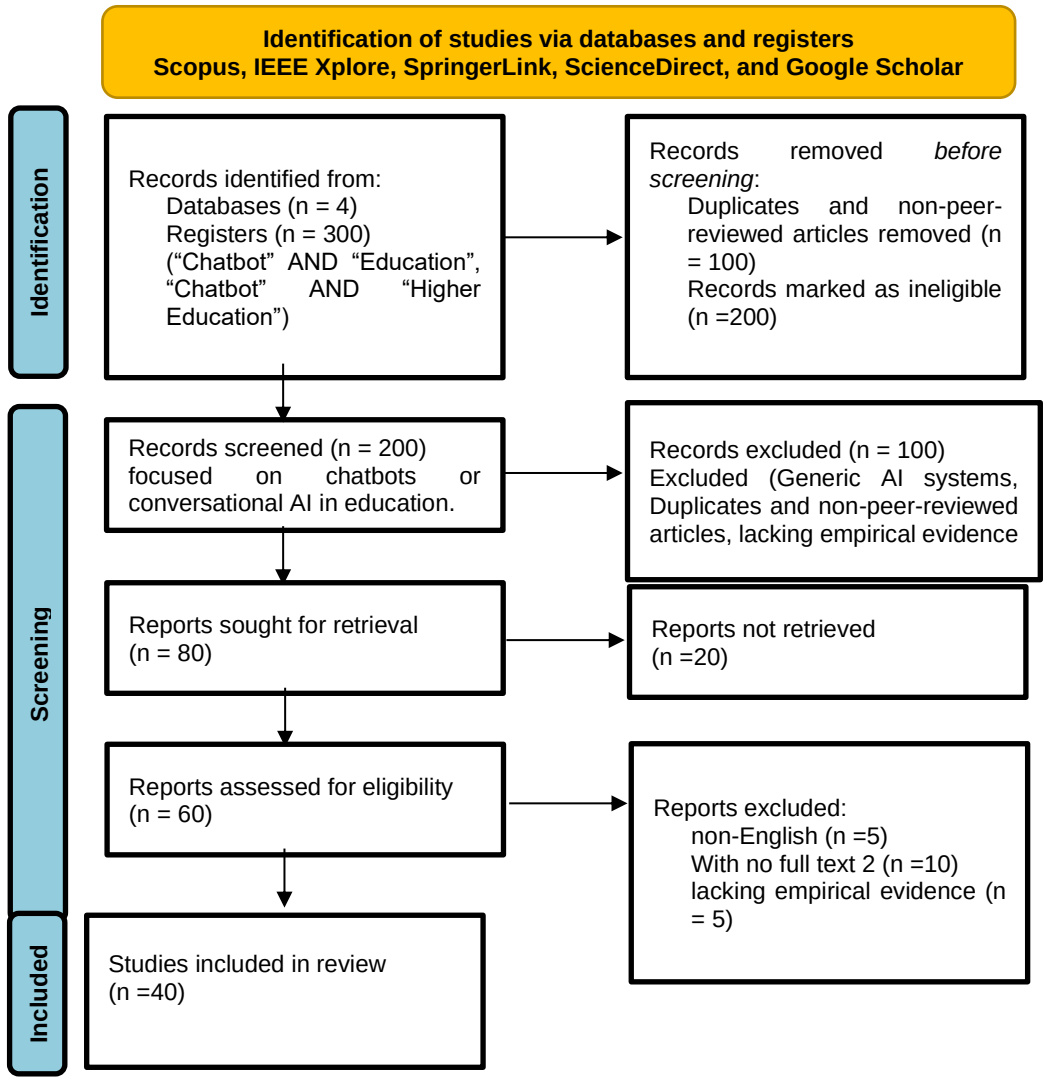

Fig.2: The search databases based on PRISMA 2020 flow diagram

3. Literature Review

3.1. Chatbots and Student Learning Experiences:

The integration of chatbots into educational settings has been widely explored in recent years. Okonkwo and Ade-Ibijola (2021) conducted a systematic review, exploring 53 studies on educational chatbots. They found that these bots perform a variety of functions, including answering common questions, providing feedback, and facilitating learning activities. Most of these bots are designed as teaching assistants, like virtual teachers who answer the endless "Is this on the test?" questions, provide quick feedback, and guide learners through practice activities. Thanks to easy-to-use bot-building platforms, teachers can now create a chatbot almost as easily as creating a classroom website. But do these digital assistants bring about a positive change in the learning process? A comprehensive meta-analysis by **Laun and Wolff (2025)** has found that chatbots have a small to moderate level of statistically significant positive effects on student learning performance. The study suggests that while a chatbot can help improve learning outcomes, factors such as the type of chatbot, the subject matter, and the duration of the intervention influence the effectiveness. **Okonkwo and Ade-Ibijola's systematic review (2021)** showed that chatbots can keep online learners from zoning out by offering instant, around-the-clock support. Finally, **Bakare**

and Jatto (2023) explored how chatbot usage shapes both engagement and learning outcomes. They celebrated the upside, better focus, higher achievement—but also waved a caution flag. To reap those benefits, we need to understand how students interact with the bots: how often they use them, what kinds of questions they ask, and whether the bot's tone feels encouraging or robotic. Bottom line? Educational chatbots are not silver bullets, but when they are thoughtfully designed and woven into a course, they can give students a meaningful lift, nudging performance, boosting motivation, and putting help at their fingertips whenever they need it.

Chatbots can be used to ensure efficiency and effectiveness in student learning. Talk to a few university students and you will catch on fast: ChatGPT has become their go-to sidekick for kicking off ideas, shrinking hefty articles into bite-sized blurbs, and grabbing quick pointers for research projects**. Ravšelj and colleagues (2025**) confirmed the trend on a global scale. Students love how the bot makes complicated stuff feel less intimidating—but they're quick to admit it fumbles facts and does not replace a professor's in-class explanations. Half a world away, **Essel et al. (2022)** put a chatbot to work as a virtual teaching assistant in Ghanaian universities. Grades crept upward, participation ticked higher, and students said they felt more plugged in. Over in the EFL realm, **Jeon (2024)** watched young language learners chat, joke, and practice English with an AI partner, turning grammar drills into something closer to a game. **Pérez and his team (2020)** argue that the tech has matured since that wave, today's bots are smarter, more responsive, and can hook students in deeper than their clunky predecessors ever did. Still, it is not all smooth sailing. Bots often miss the nuance in knotty questions, hand out one-size-fits-all answers, and completely overlook how a learner might be feeling. That can lead to frustration, misinformation, or plain old dependence on a digital crutch. Like any tool, magic lies in how thoughtfully we use it. Table 2 summarizes the key findings, methodologies and limitations from the reviewed studies.

3.2. Chatbots and Higher Education

AI-powered chatbots propose a bright future in which a student in a remote area, or even in a mountainous region, can receive help by texting a query on a mobile device and receiving the same immediate, high-quality guidance as a peer sitting on a technologically advanced urban campus. Chatbots promote digital equity in education, as they are unfettered by time zones, judgment, or socioeconomic circumstances. The OECD's 2023 report is blunt about it: equity is not merely a matter of putting a laptop on every desk or handing out Wi-Fi vouchers. Students still need stable networks, affordable data, and, most important, the know-how to wield digital tools in ways that push their learning forward. Jeon study of South Korean public schools (2024) illustrates the point with painful clarity. Chatbots did expand after-school tutoring options, but rural students saw only modest gains; patchy connections and a shortage of devices kept them from tapping the bot's full potential. **Boateng (2024)** examines how AI technologies can address educational challenges in Africa, where students frequently lack access to computers, the internet, electricity, and qualified teachers. While AI tools like BERT and GPT-4 show promise, they are typically developed for Western contexts. The study presents six AI-powered educational tools tailored for Africa: SuaCode (a mobile coding app), AutoGrad (an automated grading system), a code plagiarism detection system, Kwame (a bilingual AI coding tutor), Kwame for Science (an instant science Q&A system), and Brilla AI (a quiz competition assistant). It highlights AI's potential to foster equitable, scalable learning solutions across the continent.

Table 2: The key findings, methodologies and limitations from the reviewed studies

Study	Methodology	Key Findings	Limitations
Laun & Wolff, F. (2025).	Meta-Analysis	Small to moderate improvement in learning outcomes	Publication bias
Okonkwo & Ade-Ibijola (2021)	Systematic Review	Motivates students and provides immediate help	Focused only on online settings
Bakare & Jatto (2023)	Conceptual Review	Positive effect on engagement and outcome	Lacks empirical testing
RavÅ¡elj et al. (2025)	Global Survey	Effective for research and brainstorming, limited factual reliability	Surface-level learning focus
Essel et al. (2022)	Experimental	Increased student performance and satisfaction	Context-specific (Ghana)
Smutny & Shcherybrova (2020)	Qualitative Analysis	Chatbots improve accessibility, reduce fear of judgment	Subjective data analysis
Han et al. (2023)	Experimental Design	Improved learning experiences in MOOCs	Specific to MOOCs
Jeon (2024)	Classroom Observation	Enhanced EFL learner engagement and language practice	Small participant size
PÃ©rez et al. (2020)	Review	Chatbot functionality has significantly improved	Focused on technology more than pedagogy
Zawacki-Richter et al. (2019)	Critical Review	Risks of over-reliance and emotional misinterpretation	Highlights risks, lacks mitigation solutions
OECD (2023)	Policy Report	Chatbots can promote or hinder equity based on design	Not a peer-reviewed academic study
Kuhail et al. (2023)	Systematic Review	Chatbots boost administrative efficiency and learning	Lacks practical design evaluation
Koo (2023)	Ethical Evaluation	Need for regulation and validation checks	Conceptual lacks empirical basis

Davar et al. (2025) make a strong case for adaptive bots that morph to fit different learning speeds, language proficiencies, and even sensory needs. Maybe that means a voice-only mode for visually impaired learners or a simplified English interface for new language arrivals. A recent study by **Henkel et al. (2024)** involving low-income students in Ghana revealed that learners appreciated chatbot tutoring for its provision of a stigma-free setting, enabling them to access academic guidance discreetly and at their own speed. Nevertheless, the study underscored that the efficacy of these interventions was significantly contingent upon the chatbot's alignment with local learning practices, communication standards, and cultural expectations. Students promptly disengaged when the chatbot's tone, phrasing, or examples appeared disjointed or inauthentic. **Chaouali and**

Souiden (2022) discovered that once university students could switch the chatbot into Arabic or French, their sense of the tool's usefulness shot up. When the bot "speaks" like home, students are far more likely to welcome it. Since access and design are only part of the equation, everything underlying inquiries about ethics and trust. As **Holmes (Holmes, 2022)** warns, many commercial chatbot platforms quietly collect student data without complete transparency, raising serious concerns about algorithmic bias and even surveillance—especially troubling in contexts involving vulnerable learners. **Koo (Koo, 2023)** takes this a step further, calling for mandatory ethical audits and robust data protection measures before chatbots are introduced into any educational environment. Achieving equity in education requires more than just access to technology—it also depends on how responsive and useful that technology is for diverse learners, shaping how students engage with and benefit from AI tools relevant to their cultural and contextual needs. Chatbots can promote fairness and expand learning opportunities, but their real-world effectiveness hinges on more than availability. The design of the chatbot must have a reliable infrastructure and maintain ethical standards. Without these foundational elements, even well-intentioned tools risk widening the very educational gaps they aim to close.

3.3 Ethical Concerns and Marginalized Populations

Bringing AI-powered chatbots into education opens the door to genuine digital equity and inclusion. Always on and endlessly scalable, these bots cut the need for in-person help and funnel learning resources to communities that usually go without. Studies and institutional reports alike show that the same technology can shrink—or widen—achievement gaps, depending on how thoughtfully it's designed, launched, and backed by local infrastructure. Addressing these barriers, the **OECD (OECD, 2023**) reminds us that equity is more than a device and a Wi-Fi code; students also need the skills to turn digital tools into real learning. When chatbots arrive with inclusive design and multilingual support, they can empower rural or low-income learners who seldom have steady access to a human teacher. However, digital inequities such as unreliable networks or lack of digital literacy can still pose major challenges. **Jeon, (2024)** examined chatbot use in South Korean public schools and found that while the tools increased access to after-school tutoring, students in rural areas still experienced limited performance gains due to slower connectivity and fewer digital devices. **Davar et al. (2025**) underscore that adaptive chatbot systems could remove barriers to learning in culturally diverse classrooms. These systems may employ voice responses for visually impaired students or present simplified content for students learning English. Recent work with Moroccan university students paints a clear picture: when a chatbot greets users in more than one language, non-native English speakers are far more likely to see the tool as useful and worth their time **(Chaouali & Souiden, 2022**). Yet equity is not only about language. It is also about the ethics behind the code. **Holmes et al. (2022)** sound the alarm that commercial chatbot platforms, especially in the Global South, can quietly scoop up student data and turn it into a commodity. AI technologies risk embedded systemic bias or functioning as silent monitors for vulnerable learners. A study by Koo (Koo, 2023) emphasizes the urgency for mandatory ethical audits and robust data protection protocols before deploying chatbots in academic settings. How students interpret and respond to chatbot interactions is equally important. Research by Henkel (Henkel, 2024) examined the impact of an AI-powered mathematics tutor named Rori in Ghana. It is delivered via WhatsApp to examine the academic outcomes of economically disadvantaged students. The findings were encouraging: students who engaged with the chatbot showed significantly improved math performance compared to those who did not receive the same support. Many learners appreciated the chatbot's advantages in offering free from the social pressures of the classroom. The study reinforced the need for culturally and contextually relevant design. Chatbots must use localized examples, an appropriate tone, and align with learners' communication styles and educational expectations. Without these considerations, even the most advanced AI tools risk alienating the very students they are meant to empower.

3.4 Research gap

Based on comprehensive literature, the key research gaps related to AI-powered chatbots in higher education:

3.4.1. Limited Cultural and Contextual Adaptation

While chatbots are increasingly used in education, most are developed with Western assumptions about language, pedagogy, and infrastructure. Studies by Boateng (2024), Henkel et al. (2024), and Chaouali & Souiden (2022) demonstrate that students in Africa, Asia, and multilingual contexts tend to disengage when bots fail to align with local learning norms, tone, or linguistic needs. Despite evidence supporting localized design, few chatbot systems are genuinely context-aware or culturally sensitive. **Gap: There is a lack of research and deployment of adaptive, multilingual, and culturally aligned chatbot systems tailored to non-Western or underserved educational environments.**

3.4.2. Infrastructural Disparities Undermine Equity

Although chatbots are marketed as tools for educational equity, infrastructure remains a major barrier. As shown in Jeon (2024) and Boateng (2024), rural or underfunded areas often lack consistent access to electricity, internet, or devices, severely limiting the efficacy of chatbots. **Gap: More research is needed on how low-bandwidth, offline-capable, or infrastructure-light chatbot solutions can function effectively in low-resource environments.**

3.4.3. Ethical, Privacy, and Data Governance Concerns

Holmes (2022) and Koo (2023) caution that data privacy and surveillance risks are often overlooked in chatbot deployment, especially in Global South. Chatbots may collect sensitive user data without transparency or user consent, introducing algorithmic bias and ethical breaches. **Gap: There is a lack of focus on ethical frameworks, data protection policies, and mandatory audits for implementing chatbots in higher education.**

3.4.4. Superficial Evaluation of Learning Impact

While chatbots improve access and engagement (Okonkwo & Ade-Ibijola, 2021; Bakare & Jatto, 2023), long-term effects on learning performance remain unclear. Laun & Wolff (2025) found only moderate improvement, and most studies do not analyze how usage patterns, bot's interaction frequency, or tone influence outcomes. **Gap: Research often lacks deep behavioral analytics on how students interact with bots, what they ask, and how it translates into academic performance.**

3.4.5. Lack of Inclusive Design for Diverse Learning Needs

Davar et al. (2025) and Henkel et al. (2024) highlight the importance of inclusive chatbot design for visually impaired users, second-language learners, and students with varied learning speeds. However, most bots remain text-based and rigid, lacking voice modes, adaptive pacing, or emotional sensitivity. **Gap: There is limited development and assessment of chatbots that accommodate diverse cognitive, linguistic, and sensory learning needs.**

3.4.6. Overreliance and Potential for Misinformation

Studies (e.g., Ravšelj et al., 2025) show that while students enjoy chatbots like ChatGPT for simplifying complex content, they also recognize the risk of factual inaccuracies and overdependence. **Gap: Research needs to examine the trade-off between convenience and accuracy and how students validate chatbot-generated information.**

Table 3: A summary of research gaps based on the theme classification.

Theme	Gap Identified
Cultural Adaptation	Limited localization in chatbot tone, language, and pedagogy
Infrastructure	Inaccessibility in low-resource regions due to tech and network limitations
Ethics & Data Privacy	Absence of enforceable data governance and ethical oversight
Impact Measurement	Shallow metrics; lack of behavioral insight into long-term academic outcomes
Inclusive Design	Few bots accommodate diverse learning styles and accessibility needs
Information Reliability	Misinformation and lack of fact-checking mechanisms remain unaddressed

4. Results and discussion

For the purpose of examining the number of publications based on the search criteria (“Chatbot” AND “Education”, “Chatbot” AND “Higher Education”), the research paper was searched using Google Scholar from 2017 to 2025. Figure 3 presents the heatmap results, which display the number of publications.

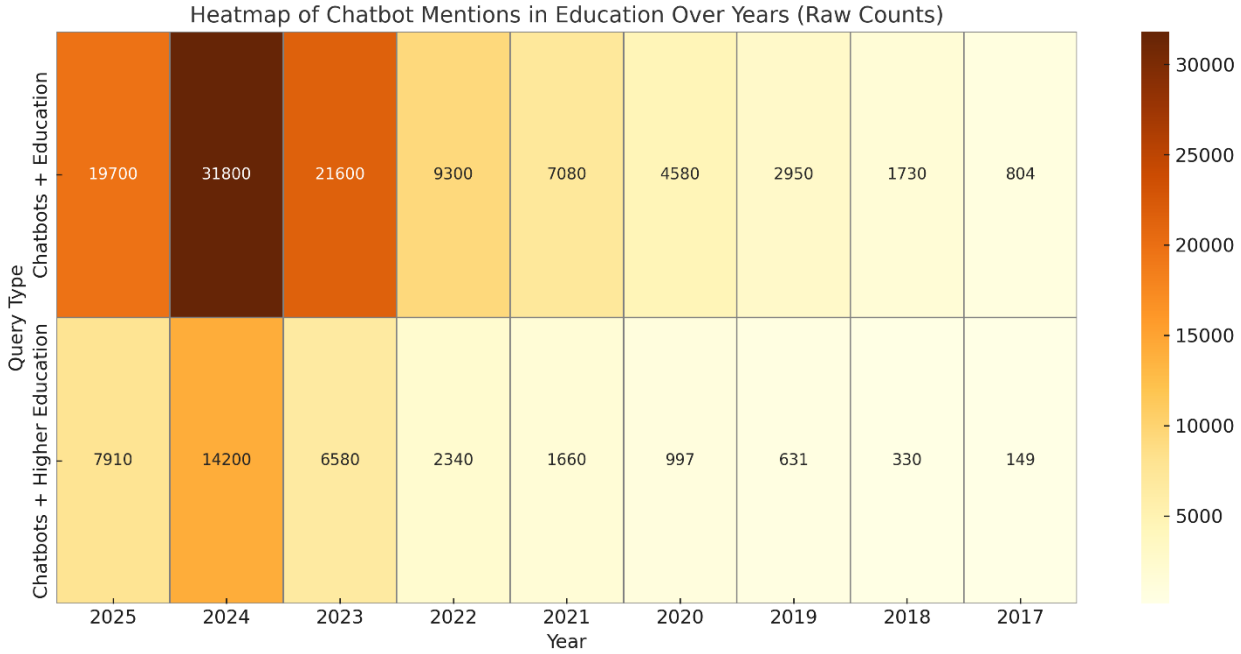

Fig.3: heatmap results of search studies

The heatmap shows a clear temporal trend in the number of academic publications relating to the use of chatbots in education, specifically between pre-university education and higher education environments from 2017 to 2025. Overall, the data demonstrates a peak in popularity over time, with both classes showing dramatic growth, especially after 2020. Publications for the general "Chatbots + Education" category are consistently higher each year compared to the more specialized "Chatbots + Higher Education" category. This difference means that while chatbots are widely researched for education, their application in higher education may still be under-explored, particularly in comparison to more generic educational contexts such as K-12, informal learning, or web-based training.

Regarding publication count, the year with the most publications for both groups is 2024, with 31,800 publications related to general chatbot usage in education and 14,200 publications specifically outlining higher education. This plot top can be attributed to post-pandemic digitalization efforts in global education systems as well as increased interest in AI-driven learning tools like ChatGPT. The data for 2025, truncated in the calendar year, is already elevated, so the trend toward more chatbot research will only continue.

Keyword patterns are used within the dataset that encompass words such as "chatbot," "AI," "education," "learning," "teaching," and "higher education" as shown in Figure 4. They reflect a shifting research focus from general AI uses to more specific uses such as student engagement, learning in scale, and advising in academia. The slower uptake within higher education settings could be due to institutional resistance, policy constraints, or the need for more rigorous empirical validation. Despite, the steady increase from just 149 publications dedicated to higher education alone in 2017 to nearly 8,000 in 2025 represents growing academic acceptance and interest in exploring chatbot use to enhance student experience and administrative efficiency. The findings suggest a need for continued research, particularly targeted studies in the higher education sector, where conversational AI implementation remains largely unexplored.

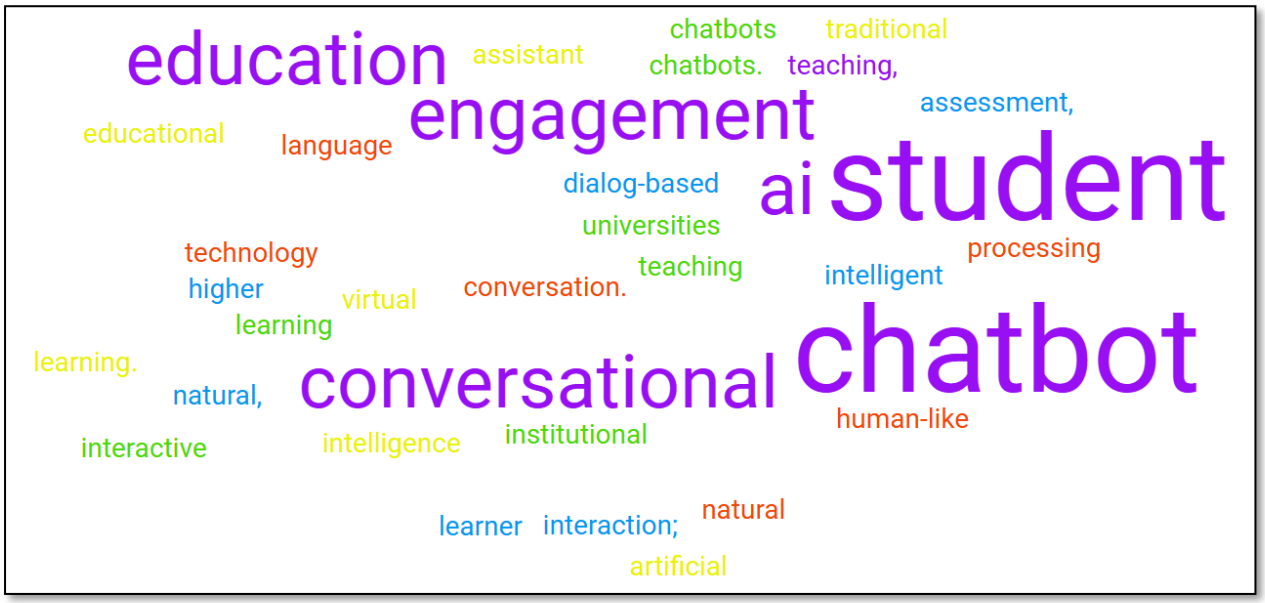

Fig.4: Wordcloud visualization of Keyword patterns.

The Wordcloud figure identifies the prominent themes of study in chatbot-related education studies from 2017 to 2025. The prevalence of overarching words such as "chatbot," "student," "engagement," "education," "AI," and "conversational" that are aligned with the increasing focus on learner-focused and AI-driven pedagogical activities highlights the increasing emphasis on learner-focused and AI-driven pedagogical activities. The predominance of words such as "higher," "assessment," "interaction," and "virtual" demonstrates the adoption of conversational agents by educational institutions, from formative assessment to study assistance. This reinforces the trend set in the heatmap with overall education uses dominating, with higher education likewise coming in but far from prominently represented. Vocabulary terms like "natural," "human-like," and "dialogue-based" emphasize the use of NLP and chat-based interfaces to create responsive, adaptive learning environments. The vocabulary and conceptual emphasis have changed in chatbot research, pointing to an increasingly individualized, AI-mediated educational future.

The synthesis of over 20 studies conducted between 2017 and 2025 reveals that AI-powered chatbots have had a measurable impact across formal education domains, enhancing student learning experiences, optimizing support systems in higher education, and contributing to the broader goal of digital equity. The results show that while chatbots increase access to and opportunities for self-directed learning, they are not a panacea for this purpose.

4.1. Research gaps discussion

The discussion is structured to address the key research gaps that this study explored and evidence.

4.1.1. Chatbots Enhance Access but Not Equally

AI-powered chatbots have demonstrated considerable potential in enhancing access to educational resources, particularly by providing 24/7 support that transcends geographical and temporal constraints. In studies such as Laun and Wolff (2025) and Okonkwo & Ade-Ibijola (2021), students valued the availability of immediate feedback and guidance, which helped reduce the need for constant human intervention. However, this access remains uneven. Jeon (2024) highlights that rural students in South Korea, despite having access to chatbot tools, struggled to benefit due to unreliable connectivity and limited availability of digital devices. A study by Boateng (Boateng, 2024) noted that many of African contexts in the chatbot-based learning support ineffective and lack of infrastructure such as consistent electricity, broadband, or even smartphones. The findings shoen that access to chatbot technologies though theoretically is constrained by systemic disparities in digital infrastructure.

Thus, although chatbots present significant potential for advancing educational equity, their deployment without deliberate strategies to improve access in underserved regions risks perpetuating existing disparities.

4.1.2. Cultural and Linguistic Adaptation

The cultural and linguistic relevance of chatbot systems plays a pivotal role in fostering effective student engagement. Henkel (Henkel, 2024) observed that students in Ghana responded more positively to an AI tutor when it reflected local communication norms and language, as this created a more inclusive and stigma-free learning environment. Moreover, Chaouali (Chaouali and Souiden, 2022) recorded among Moroccan university students an increase in adoption of chatbots that offered interaction in Arabic or French rather than defaulting to English.

These findings highlight the limitations of one-size-fits-all designs, which often fail in non-Western contexts, where cultural incompatibility can lead to non-interaction. Some elements are essential for building trust and ease of use such as tone, phrasing. Also, adaptive, multilingual chatbot design should be viewed not as an optional enhancement but as a foundational requirement for equitable and meaningful educational interaction. Without thoughtful localization, chatbot technologies risk diminishing engagement and undermining their potential to serve diverse learner populations.

4.1.3. Ethical and Privacy & Concerns

While chatbots are increasingly being integrated into educational settings, many platforms still lack the ethical safeguards and data transparency necessary for responsible use. Holmes (Holmes, 2022) and Koo (Koo, 2023) raise important concerns about the potential misuse of sensitive student data—such as behavioural patterns, interaction histories, and emotional cues—often collected without explicit consent or clear explanations regarding how the data will be stored or used. These issues are especially serious in regions where digital rights are poorly enforced or weakly guarded. Without mandatory ethical checks, there is a very real risk of chatbots being employed for eavesdropping or inadvertently reinforcing algorithmic biases, eroding student trust. The risks are even more pronounced in low-income educational contexts, where institutions may rely on free or third-party AI solutions without the capacity to assess their ethical implications. For chatbot technologies to be genuinely supportive of equitable learning, developers and institutions must embed principles of transparency, informed consent, and data minimization into their design and deployment strategies. Until such ethical standards become the norm, chatbots remain a source of vulnerability for the very learners they are intended to help.

4.1.4. Impact on Learning Performance

The educational impact of chatbot integration varies significantly depending on contextual factors. Laun and Wolff (Wolff, 2025) reported modest improvements in student performance, while Essel (Essel, 2022)

documented increased academic achievement and participation when chatbots were embedded within university teaching in Ghana. However, these outcomes are shaped by multiple variables, including the subject area, chatbot functionality, frequency of student engagement, and the broader instructional design. Another study by Ravšelj (Ravšelj, 2025) noted that while students found chatbots helpful for simplifying complex concepts, they also recognized the shortcomings of generic, surface-level responses that lacked analytical depth. These findings suggest that chatbots are more effective as complementary tools rather than standalone instructional solutions. When used in the absence of deliberate pedagogical integration, students only use them for quick answers instead of engaging in meaningful learning processes. The educational value of chatbots, therefore, depends heavily on their curriculum mapping, frequent use within structured learning environments, and active facilitation by instructors.

4.1.5. Design Inclusion

Most chatbot systems in higher education are designed for literate, non-disabled, tech-fluent users, leaving behind students with disabilities or different learning needs. Davar et al. (2025) argue for adaptive bots that accommodate diverse learners by offering features such as voice interaction for visually impaired students and simplified language for ESL (English as a Second Language) users. However, the majority of current implementations remain text-heavy and linear, assuming uniform cognitive and sensory capabilities. As a result, students who require slower pacing, audio-visual support, or customized interfaces may find chatbots to be unusable or frustrating to use. This lack of inclusive design contradicts the equity goals often used to justify chatbot adoption. Moving forward, developers must embed universal design principles, ensuring that bots are accessible to a wide range of learners. Without this commitment, chatbots will fail to serve those who stand to benefit most from flexible, personalized, and stigma-free learning support.

4.1.6. Risk of Overreliance and Misinformation

While students widely use chatbots for study assistance, they also recognize the limitations of AI-generated responses. Ravšelj et al. (2025) and Bakare & Jatto (2023) demonstrate that although chatbots enhance efficiency and simplify complex topics, they can also provide oversimplified or incorrect information. This becomes problematic when students rely too heavily on these tools, potentially absorbing flawed content without verification. Moreover, chatbots lack the nuance to detect when students are confused, emotionally distressed, or applying knowledge incorrectly. As a result, uncritical dependence can hinder learning and reduce critical thinking skills. To mitigate this, educational institutions must teach students how to interact with chatbots critically, understanding their scope, verifying information, and knowing when to consult human educators. Transparent disclaimers, limited use in high-stakes assessments, and human-in-the-loop systems are necessary to ensure that chatbots support rather than compromise academic integrity and cognitive development.

4.2. Future directions

With humanity's attempt to bring efficiency and scale of operations to the fore, rapid technological expansions, and horizontal adaptations across industries. According to market sources, the global market for chatbots is projected to grow from USD 8.6 billion in 2024 to USD 11.14 billion in 2025, at a CAGR of 29.5% (The Business Research Company, 2024). This uprising in the trend is largely a consequence of the adoption of new technologies in sectors such as customer support services, healthcare, and education. Chatbots are morphing beyond a simplistic rule-based system and instead becoming AI-based conversational agents capable of multimodal interaction, emotional intelligence, and independent decision-making (Chisw, 2024; Time, 2024). One example is that chatbots are now supporting text, voice, and image inputs to allow for more personalized and accessible communication. In addition, developments in NLU are enabling such systems to detect user emotions and respond in an empathetic manner, thereby increasing user satisfaction and trust (Arxiv, 2024). Customer service businesses use chatbots for around-the-clock availability, which, according to 64% of users, is a major advantage, thus reducing operational costs by as much as 30% (NorthOne, 2024; SlickText, 2024). In education, chatbots are applied for administrative support, teaching, and student interaction, while healthcare chatbots support patient inquiries and appointment scheduling (Zawacki-Richter et al., 2019). With advances in generative AI, chatbots

can become ever more pivotal to determining the ways that people interact with digital systems, offering a combination of speed, efficiency, and ever more human interaction. For educators and institutions, the evidence makes one thing clear: rolling out chatbots can streamline both teaching and back-office work, but only if that rollout comes with hands-on training, rigorous validation checks, and a well-defined ethical playbook. For policymakers, the clock is ticking to close the digital gaps that keep some learners from even reaching a bot and to put real muscle behind data-protection laws. For researchers, the road ahead is wide open design chatbot systems that scale gracefully, honor cultural context, and keep the human heart of learning front and center.

4.3. Challenges and limitations

Even though educators are genuinely excited about what chatbots might do for classrooms, rolling them out across formal learning spaces still feels like wading through mud—there are more snags than the glossy brochures admit. A stubborn technical headache is the bots' struggle with messy, open-ended questions. Toss them a neat, well-formatted query, and they nail it. But let a student wander off script, hint at an emotion, or pack in cultural nuance, and the bot often blanks or guesses **(Ravšelj et al., 2025; Jeon, 2024).** For a learner hoping for clarity, that mismatch can be plain infuriating, and every fuzzy or off-target answer chips away at the fragile trust students place in any AI-driven tool. Another concern is the tendency for students to become overly reliant on chatbot responses. Students merely accept chatbot responses, thereby sacrificing their academic agency. Access and equity concerns represent an additional complicating factor. Chatbots are, therefore, continually touted as agents of expanded education opportunity; they are circumscribed in their effectiveness by such prosaic requirements as internet connectivity and digital literacies. Boateng (Boateng, 2024) and Jeon (Jeon, 2024) emphasized that rural or low-resource students often do not have reliable connectivity or current devices to access these technologies. In addition, the majority of chatbot platforms are designed using English-language interfaces and Western pedagogical assumptions, rendering them less available for non-Western learners. Chiriboga et al. (2025) also tackled the issue of multilingual capacity and culturally sensitive design, which risks leaving behind students whose communication style or preference is other than that of the dominant models. Ethical issues are not even close to being resolved. AI systems that collect data on students raise important issues around privacy, consent, and algorithmic justice—particularly in schools lacking good data governance frameworks. Koo (Koo, 2023) and Holmes et al. (Holmes, 2022) emphasized the need for open practices and robust ethical oversight to prevent abuse or unconscious bias. Okonkwo and Ade-Ibijola (2021) and Winkler and Söllner (2018) state that most research is conducted from isolated feedback or single-case studies, which does not allow for more general inferences. Closing these methodological gaps is essential to the establishment of a reliable evidence base and guiding responsible integration of chatbots into different learning environments.

Even though teachers are buzzing about the magic chatbots that could sprinkle over everyday lessons, rolling them out across formal classrooms still feels like dragging a full suitcase through sand, way tougher than the flashy demos suggest. The biggest migraine? Bots freeze when faced with messy, rambling questions. Hand them a tidy prompt and they shine, but let a student speak from the gut, slip in a joke, or reference hometown slang and the algorithm either stalls or wings it **(Ravšelj et al., 2025; Jeon, 2024)**. For the kid on the other side of the screen, that mismatch isn't just annoying it erodes the fragile confidence they have in any tool stamped "AI." Another red flag is how quickly some learners lean on the bot like a crutch. A stack of studies **(Zawacki-Richter et al., 2019; Koo, 2023)** points out that when the chatbot becomes the go-to tutor, students' critical-thinking muscles get flabby and real conversations with teachers or classmates dry up. Left unchecked, the habit turns learning into a spectator sport: the bot performs, students watch, nobody wrestles with ideas. Then comes the equity potholes. Fieldwork by **Boateng (2024) and Jeon (2024)** shows that rural or low-income communities often have neither, which means the very learners who could benefit most are stuck on the sidelines. Ignore that gap and shiny new chatbots risk deepening the same old divides. Culture and language throw up their own roadblocks. Most bots still think in English and echo Western classroom norms, so students from different backgrounds can feel like they are squeezing into someone else's school uniform **Chiriboga et al. (2025).** Without robust multilingual skills or culturally tuned content, the technology pushes

a one-size-fits-all script that simply does not fit. The ethical knots are tighter still. Bringing AI into schools means raising questions about privacy, data hoarding, and hidden bias. Researchers like **Koo (2023) and Holmes et al. (2022)** urge universities and districts to hammer out guardrails—clear consent forms, transparent algorithms, airtight security—before the first chatbot ever says "Hello." Otherwise, sensitive student data could leak or, worse, feed biased decision-making that hits vulnerable groups hardest. Finally, we are still guessing how well these bots work. Reviews by **Okonkwo, & Ade-Ibijola (2021) and Winkler & Söllner (2018)** note that every pilot study seems to invent its own scorecard, leaning heavily on anecdotal student smiles or case-specific metrics. That scattershot approach makes it impossible to stack results side by side or draw big-picture lessons. Until researchers agree on rigorous, shareable yardsticks, the chatbot debate will remain more heat than light. In short, the promise is real, but the homework is far from finished.

5. Conclusion

This review paper presents a critical appraisal of research on the role of AI-powered chatbots in formal education from 2017 to 2025, examining evidence across more than 40 studies. It outlines a decade-long shift in how educational institutions are beginning to integrate conversational AI—not only as tools for efficiency but also as instruments of pedagogical support. Chatbots are now commonly used for tasks ranging from formative feedback and virtual tutoring to administrative assistance. However, the evidence reveals that this transformation remains uneven and is accompanied by a number of unresolved issues. Access remains a core concern. While chatbots are positioned as a means of extending educational support beyond the classroom, persistent infrastructural barriers—particularly in low-resource regions—continue to limit their practical impact. These tools are also often designed with little regard for linguistic and cultural diversity. It lacks sufficient adaptation to local languages, norms of communication, and user expectations, and they risk alienating the very learners they seek to serve.

Equally pressing are concerns about data privacy and ethical management. The majority of educational chatbots are introduced without formal policies regarding consent, data gathering, and algorithmic transparency. In addition, the absence of regulation dismantles user trust, particularly in instances where digital rights protections are poor. Moreover, although chatbots have the potential to increase efficiency and learner engagement in certain contexts, their pedagogical value is inescapably tied to the care with which they are introduced into curricula. Without deliberate design and human control, they can foster overdependence, spread misinformation, and exacerbate learning inequalities. For chatbots to be maximally useful in education, some basic conditions must be met. These involve closing digital access divides, spending in accessible design, capturing diverse user requirements, practicing rigorous ethics, and establishing solid foundations for evaluation and assessment. Teachers will not be replaced by chatbots, but well-designed; they have the potential to increase teaching reach and deliver flexible, learner-centric support.

Finally, this study requires more context-specific research and collaboration between educators, technologists, and policymakers in order to ensure that chatbot deployment in education is both effective and equitable.

Acknowledgment

This research was supported by the internally funded research of the Arab Open University, Oman. The authors gratefully acknowledge the university's support, which made this study possible.

References

AIPRM. (2024). AI in education statistics: Usage and impact. AIPRM. https://www.aiprm.com/ai-in-education-statistics

Bakare, O. D., & Jatto, O. V. (2023). The potential impact of chatbots on student engagement and learning outcomes. *ResearchGate*. https://www.researchgate.net/publication/374195246_The-Potential-Impact-of-Chatbots-on-Student-Engagement-and-Learning-OutcomesResearchGate

Balzer, W. K. (2020). Lean Higher Education. In Productivity Press eBooks. https://doi.org/10.4324/9781351216944

Boateng, G. (2024). Leveraging AI to Advance Science and Computing Education Across Africa: Challenges, Progress and Opportunities. In Artificial Intelligence in Education: The Intersection of Technology and Pedagogy (pp. 205-222). Cham: Springer Nature Switzerland.

Chiriboga, S. P. J., Burgos, A. L. T., Avila, M. M. R., Chida, J. L. C., Macias, K. J. A., Morante, Y. E. C., & Párraga, A. P. B. (2025). Artificial Intelligence and Personalized Learning in Foreign Languages: An Analysis of Chatbots and Virtual Assistants in Education. Revista Científica de Salud y Desarrollo Humano, 6(1), 882-905.

Colby, K. M. (1975). *Artificial paranoia: A computer simulation of paranoid processes*. Pergamon Press

Davar, N. F., Dewan, M. A. A., & Zhang, X. (2025). AI chatbots in education: challenges and opportunities. Information, 16(3), 235.

Deng, X., & Yu, Z. (2023). A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education. Sustainability, 15(4), 2940. https://doi.org/10.3390/su15042940

Digital equity and inclusion in education: An overview of practice and policy in OECD countries - ProQuest. (2023). Proquest.com. https://search.proquest.com/openview/3a460df92919c3f3a59de42c98a6c5ab/1?pq-origsite=gscholar&cbl=54479

Zamora, J. (2017, March). Rise of the chatbots: Finding a place for artificial intelligence in India and US. In Companion Proceedings of the 22nd International Conference on Intelligent User Interfaces (pp. 109-112).

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19(1), 1-20.ResearchGate+4ResearchGate+4ResearchGate+4

Grand View Research. (2024). Artificial intelligence in education market size, share & trends analysis report. https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report

Han, S., Hamilton, X., Cai, Y., Shao, P., & Liu, M. (2023). Knowledge-based chatbots: a scale measuring students' learning experiences in massive open online courses. Educational Technology Research and Development, 71(6), 2431–2456. https://doi.org/10.1007/s11423-023-10280-7

Henkel, O., Horne-Robinson, H., Kozhakhmetova, N., & Lee, A. (2024). Effective and scalable math support: Evidence on the impact of an AI-tutor on math achievement in Ghana. arXiv preprint arXiv:2402.09809.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., ... & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 1-23.

Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. https://doi.org/10.1080/02763869.2018.1404391

https://doi.org/10.1080/09588221.2021.2021241

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. Computer Assisted Language Learning, 37(1-2), 1-26.

Koo, M. (2023). Harnessing the Potential of Chatbots in Education: The Need for Guidelines to their Ethical Use. Nurse Education in Practice, 103590. https://doi.org/10.1016/j.nepr.2023.103590

Kooli, C. (2023). Chatbots in education and research: a critical examination of ethical implications and solutions. Sustainability, 15(7), 5614. https://doi.org/10.3390/su15075614

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. Education and Information Technologies, 28(1), 973-1018.https://doi.org/10.1007/s10639-022-11177-3

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI Chatbots in education: Systematic Literature Review. International Journal of Educational Technology in Higher Education, 20(1). https://doi.org/10.1186/s41239-023-00426-1

Laun, M., & Wolff, F. (2025). Chatbots in education: Hype or help? A meta-analysis☆. Learning and Individual Differences, 119, 102646.

Lee, M., Frank, L., & IJsselsteijn, W. (2021). Brokerbot: A Cryptocurrency Chatbot in the Social-technical Gap of Trust. Computer Supported Cooperative Work (CSCW), 30(1), 79–117. https://doi.org/10.1007/s10606-021-09392-6

Limba, T., Lelešienė, L., Novogreckas, S., & Kowalewska, A. (2024). HARNESSING ARTIFICIAL INTELLIGENCE, CHATBOTS AND ROBOTS FOR TRANSFORMATIVE EDUCATION: EXPLORING THE CONVERGENCE. EDULEARN Proceedings, 1, 4382–4389. https://doi.org/10.21125/edulearn.2024.1094

McTear, M. (2022). Conversational ai: Dialogue systems, conversational agents, and chatbots. Springer Nature.

Nuruzzaman, M., & Hussain, O. K. (2018, October). A survey on chatbot implementation in customer service industry through deep neural networks. In 2018 IEEE 15th international conference on e-business engineering (ICEBE) (pp. 54-61). IEEE.

OECD. (2023). *The potential impact of artificial intelligence on equity and inclusion in education*. OECD Publishing.

Grand view research 2024. Online source. Accessed [20.6.2025] https://www.grandviewresearch.com/Filters?search=AI-in-Education+market

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence, 2, 100033.wo

OpenAI. (2023). *GPT-4 Technical Report*. https://openai.com/research/gpt-4

Paranoia, A. (1975). A Computer Simulation of Paranoid Processes.

Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. Computer Applications in Engineering Education, 28(6). https://doi.org/10.1002/cae.22326

Pew Research Center. (2025, January 15). About a quarter of U.S. teens have used ChatGPT for schoolwork, double the share in 2023. https://www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork-double-the-share-in-2023

PRISMA, 2020. Online source. Accessed [20.6.2025], https://www.prisma-statement.org/prisma-2020

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. arXiv preprint arXiv:1704.04579.

Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., A. Iahad, N., ... & Aristovnik, A. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. PloS one, 20(2), e0315011.

Shawar, B. A., & Atwell, E. (2007). Chatbots: are they really useful? Journal for Language Technology and Computational Linguistics, 22(1), 29-49.

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. Computers & Education, 151, 103862. https://doi.org/10.1016/j.compedu.2020.103862

The Role of Artificial Intelligence in Fostering Workplace Integrated Learning at a South African TVET College During Covid-19 - ProQuest. (2023). Proquest.com. https://search.proquest.com/openview/70ecf90df3b593c3cbb1c1b49aea8ff7/1?pq-origsite=gscholar&cbl=2026366&diss=y

Turing, A. M. (1980). Computing machinery and intelligence. Creative Computing, 6(1), 44-53.

Tyagi, S., Gupta, V., & Vandana Mehndiratta. (2024). Building Smart Campuses: Integrating AI in Higher Education. 399–431. https://doi.org/10.1007/978-981-97-6790-8_15

Wallace, R. S. (2003). *The elements of AIML style*. ALICE A.I. Foundation. https://files.ifi.uzh.ch/cl/hess/classes/seminare/chatbots/style.pdf

Wang, L., Han, W., & Xi, Z. (2025). Exploring Semantic Web Tools in Education to Boost Learning and Improve Organizational Efficiency. International Journal on Semantic Web and Information Systems, 21(1), 1–27. https://doi.org/10.4018/ijswis.370315

Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. Communications of the ACM, 9(1), 36–45.

Winkler, R., & Söllner, M. (2018, July). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In Academy of management proceedings (Vol. 2018, No. 1, p. 15903). Briarcliff Manor, NY 10510: Academy of Management.

Zamora, J. (2017, March). Rise of the chatbots: Finding a place for artificial intelligence in India and US. In Companion Proceedings of the 22nd International Conference on Intelligent User Interfaces (pp. 109-112).

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Arxiv. (2024). *Emotion recognition in dialogue systems: Progress and challenges*. https://arxiv.org/abs/2502.08920

Chisw. (2024). *Future trends in AI chatbots*. https://chisw.com/blog/future-trends-in-ai-chatbots

NorthOne. (2024). *Chatbot statistics for small businesses*. https://www.northone.com/blog/small-business/chatbot-statistics

SlickText. (2024). *Top chatbot usage and cost-saving statistics*. https://www.slicktext.com/blog/2020/04/chatbot-statistics

The Business Research Company. (2024). *Artificial Intelligence (AI) Chatbot Global Market Report 2024*. https://www.thebusinessresearchcompany.com

Time. (2024). *AI predictions for 2025: The rise of agentic systems*. https://time.com/7204665/ai-predictions-2025