Journal of Interdisciplinary Studies in Education

Volume 15, Issue 5 (2026), pp. 321-354
Journal of Interdisciplinary Studies in Education
ISSN: 2166-2681Print 2690-0408 Online | https://ojed.org/jise
https://doi.org/10.32674/jrw9d811

Leveraging Educational Chatbots to Train Future Worker: A Review of Applications, Outcomes, and Challenges

Alfaza Ranggana, Lala Septem Riza*, Ade Gafar Abdullah
Indonesia University of Education, Indonesia
**Corresponding author: lala.s.riza@upi.edu*

ABSTRACT

This study presents a systematic literature review and bibliometric analysis examining how chatbots and virtual agents are used in educational and professional training. Drawing from 52 articles (2021–2025) and a thematic synthesis of ten empirical studies, this review addresses eight research questions concerning chatbot applications, feedback mechanisms, challenges, design features, and effectiveness. The findings show that chatbots are widely deployed as virtual patients, simulated students, and interactive tutors across medicine, teacher education, programming, and risk assessment. Personalized, real-time feedback positively affects learning outcomes, while key challenges include limited empathy, inconsistent responses, and AI overreliance. The review also maps publication trends and emerging themes around LLMs and emotionally responsive interfaces, offering a framework for educators, developers, and policymakers in the design of ethical, sustainable AI-based training interventions.

Keywords: AI training, educational chatbot, feedback mechanism, systematic literature review, virtual agent, vocational education

INTRODUCTION

In the face of an era of technological disruption and digital transformation that is accelerating, the need for a digitally adaptive and skilled workforce is a global priority (Kohl et al., 2021; Dr. Obermayer et al., 2022). The World Economic Forum (2023) projects that more than 50% of workers worldwide will need to reskill or upskill in the next five years to keep up with technological developments, particularly in areas such as artificial intelligence, data analytics, and automation (Li, 2024; Vass & Kiss, 2023). In the midst of these challenges, educational chatbots are emerging as an innovative solution to support more flexible, personalized, and scalable training and skills development (Engeness et al., 2025; Labadze et al., 2023; Schmid-Meier & Schulz, 2025). This technology enables the delivery of learning materials in real time, interactive, and adaptable to individual needs, making them relevant in the context of vocational training and lifelong learning (Proske et al., 2023; Yaseen et al., 2025). In addition, the use of chatbots in education supports the global agenda of the Sustainable Development Goals (SDGs), especially the 4th goal (quality education) and the 8th goal (decent work and economic growth), by providing broader and more inclusive access to technology-based training. In this context, reviewing the literature on the application of chatbots for training in future work is crucial to understanding the extent to which this technology has contributed, the challenges it faces, and its future development opportunities.

While the potential of educational chatbots for future workforce training is promising, their implementation still faces a variety of conceptual, technical, and contextual challenges (Babu et al., 2025; Davar et al., 2025; Musundire, 2025). Prior studies have shown the complexity of designing chatbots that provide meaningful interactions, adapt to user needs, and align with set learning goals (Alabbas & Alomar, 2025; Villegas-Ch et al., 2024). These challenges include limited semantic understanding by the system, a lack of personalization, and a lack of optimal integration with formal and nonformal training curricula (Farah et al., 2022; Rajput, 2025). In addition, the diversity of study contexts in terms of industry sectors, user backgrounds, and digital infrastructure shows that the effectiveness of chatbots is strongly influenced by technological readiness, organizational culture, and policy support. Another constraint often identified in the literature is low end-user digital literacy and a lack of systematic evaluation of the impact of long-term learning (Woollaston et al., 2025). Nevertheless, rapid advances in natural language processing, machine learning, and conversational AI have resulted in many opportunities to develop more intelligent, adaptive, and contextual chatbots. In addition, increasing government and international attention to 21st century skills development and technology-based education has prompted the need for more in-depth research on the effectiveness and sustainability of chatbot implementation in vocational training and human resource development.

The literature on educational chatbots for workforce training has shown significant growth in recent years (Chen & Huang, 2023), with diverse focuses ranging from AI-based system development (Traymbak et al., 2024) and user interaction analysis to the evaluation of learning outcomes (Wollny et al., 2021). A number of studies have examined the effectiveness of chatbots in increasing motivation, engagement, and understanding of concepts, particularly in the context of higher education and technology-based training (Suciati et al., 2024; Huang et al., 2025; Luria, 2024). However, these studies are generally fragmentary and do not present a comprehensive synthesis of the relationships among chatbot design, implementation context, and learning outcomes. Most studies also tend to focus on limited case studies or technical approaches, while systematic exploration of pedagogical variables, sociocultural factors, or policy roles is minimal (Chen et al., 2025; Hwang & Chang, 2023; Stapić et al., 2020). In addition, few literature reviews have specifically compared chatbot applications across different sectors, geographic regions, or vocational learning models. This gap highlights the urgent need to integrate the findings in a systematic review that not only maps applications and outcomes but also identifies key challenges and strategic opportunities for the development of educational chatbots in the context of global workforce training.

In the context of the global need for technology-based education and job training transformation, this systematic study is becoming increasingly relevant and urgent (García-Peñalvo, 2022). This systematic literature review (SLR) is needed to consolidate scattered and heterogeneous research findings and to compile a complete understanding of how educational chatbots have been utilized, what results have been achieved, and what challenges are still being faced (Lefaivre & Slobogean, 2013). This study will not only make a theoretical contribution by mapping the key themes, methodologies, and variables that have been studied but also provide a strong empirical basis for the development of conceptual models and future research agendas. Researchers will benefit from identifying gaps that can be used as a basis for further research, while education practitioners and technology developers can understand best practices and the potential for more effective design. For policymakers, the findings of this study have the potential to serve as a strategic reference in formulating AI-based job training policies and digital technologies in an inclusive and sustainable manner. As such, this article makes important contributions to bridging the gap between technological innovation, vocational education needs, and global policy directions in the digital age.

To address the complexity and diversity of issues in the development and utilization of educational chatbots for workforce training, this study adopts a systematic literature review (SLR) approach supported by bibliometric analysis. The study is designed to answer eight key research questions that reflect the

dimensions of application, effectiveness, challenges, and the evolution of research in this field.

- **RQ1:** How are chatbots and virtual agents used to support training and skill development in both professional and educational contexts?
- **RQ2:** What are the types of integrated feedback mechanisms and their impact on learning outcomes?
- **RQ3:** What are the commonly reported challenges in implementing chatbots for training simulations?
- **RQ4:** How do chatbot design features such as conversational style, empathy, and AI sophistication affect user engagement and training effectiveness?
- **RQ5:** Is there any evidence regarding the effectiveness of virtual agents in simulating communication-based or interview-based training scenarios?
- **RQ6:** How have the number of publications on the application of chatbots in training or simulation evolved from year to year, and what are the main journal sources that publish the most on this topic?
- **RQ7:** What are some of the main trends and topics that are developing in the literature about chatbots for training or simulation?
- **RQ8:** What are the most influential (highly cited) documents in chatbot research for training or simulation, and how does it contribute to the intellectual structure of this field?

Through these questions, this study not only compiles a map of the latest knowledge but also provides an academic and practical foundation for the development of more strategic and impactful chatbot innovations in the context of future training.

LITERATURE REVIEW

Historical Context and Technical Evolution

Conversational agents have been a part of computing history since 1966 with the development of ELIZA (Wollny et al., 2021). Traditional educational chatbots are often rule-based systems that rely on predefined keyword detection and structured dialog flows (Davar et al., 2025). Recent advancements in natural language processing (NLP) and large language models (LLMs) have enabled the rise of generative AI (GenAI) chatbots, such as ChatGPT, which can engage in more fluid and contextually relevant academic interactions (Davar et al., 2025; Wollny et al., 2021).

Chatbots for Education

Current research on the pedagogical role of chatbots identifies three primary roles for chatbots in education:

- Learning Agents: These agents serve as virtual tutors that deliver instructions and answer content-specific questions (Wollny et al., 2021).
- Assisting Agents: Administrative support, such as guiding students through library resources or enrollment procedures, should be provided (Wollny et al., 2021; Kuhail et al., 2023).
- Mentoring Agents: Facilitating metacognition and self-regulated learning by prompting students to reflect on their own progress (Wollny et al., 2021; Kuhail et al., 2023).

The benefits and learning outcomes Chatbots in higher education enhance student performance by offering 24/7 accessibility and "just-in-time" feedback, which is particularly vital in remote or underprivileged contexts (Labadze et al., 2023; Jose & Jayaron Jose, 2024). Furthermore, they support personalization by adapting to individual learning pace and styles, often outperforming human feedback in specific domains such as emotional awareness evaluations (Huang et al., 2025; Elyoseph et al., 2023). The implementation of these tools is frequently driven by the desire to increase student motivation and engagement (Huang et al., 2025; Yaseen et al., 2025).

Challenges and Risks Critical concerns include the potential for "hallucinations" in the generation of factually incorrect or fabricated information and references (Davar et al., 2025; Macdonald et al., 2023). Moreover, educators express concern that overreliance on chatbots could negatively impact academic integrity through increased plagiarism or lead to a decline in students' critical thinking skills (Sánchez-Vera, 2025; Jose & Jayaron Jose, 2024).

Worker training and professional development

Simulation-based training

In professional sectors, chatbots serve as sophisticated simulation tools for high-stakes environments. In healthcare training, they are employed as virtual patients to help nursing and medical students practice history-taking and diagnostic reasoning without risking patient safety (Kaur et al., 2021; Chen et al., 2023). Similarly, forensics programs use chatbots to train students in juvenile offender risk assessment, allowing for safe, repetitive practices involving complex interview techniques (Raiche et al., 2023).

Teacher professional development

AI-driven scenario-based training is used to improve "teachers" ability to attend to and interpret students' mathematical thinking in real time (Galiç et al., 2025). By interacting with a chatbot acting as a "virtual student" with specific misconceptions, teachers can refine their rational questioning and instructional adaptation skills (Galiç et al., 2025).

Vocational and Administrative Support

In Technical and Vocational Education and Training (TVET), chatbots enhance worker productivity by automating repetitive helpdesk inquiries and technical support tasks (Alabbas & Alomar, 2025). This integration requires a high task-technology fit (TTF), ensuring that the AI tool aligns specifically with the professional tasks of the worker (Soodan et al., 2024).

Role of Digital Literacy

The success of worker training with AI is significantly moderated by the user's digital literacy (Yaseen et al., 2025). Employees with greater digital skills are better positioned to leverage the benefits of personalized feedback and adaptive learning technologies during their training (Yaseen et al., 2025).

RESEARCH METHOD

PRISMA plays a crucial role in increasing transparency, reducing selection bias, and supporting evidence-based decision-making, making it a highly relevant tool for fields influenced by education policy and training technology innovation (Boaye Belle & Zhao, 2023; Page et al., 2021). In this study, the PRISMA approach was used to systematically review the application of educational chatbots in the training of prospective professional workers. The literature search strategy was carefully designed using the TITLE-ABS-KEY QUERY ((“chatbot*” OR “virtual agent*” OR “conversational agent*”) AND (“training” OR “skill development” OR “upskilling”) AND (“education” OR “e-learning”) AND (“feedback” OR “engagement” OR “simulation” OR “learning outcome*”) to ensure the coverage of the relevant yet comprehensive information. The term “Chatbot” and its variations are intended to capture studies that address agents in terms of AI-based conversations or communications used in educational contexts. Moreover, the term “training” was chosen to identify research that highlights the implementation of chatbots as a means of learning or training, particularly in simulated software testing skills. This combination of keywords aims to balance specificity and reach, ensuring that the analyzed studies are truly relevant to the development and evaluation of chatbots in the context of professional training.

The selection of the right database is important for guaranteeing that the literature obtained is of high quality and comprehensively represents the scope of

relevant research. In this study, the search was conducted in the Scopus database, a widely recognized source of academic literature (Baas et al., 2020). The inclusion criteria are: English-language journal articles, open access, classified as scientific journal articles (not proceedings or reviews), and published from 2021 to 2025. This time frame was chosen to reflect the latest developments in the application of educational chatbots, especially in the training of prospective workers. From the initial search results, 303 relevant articles were obtained on the basis of matches against predesigned keywords. After screening on the basis of the inclusion criteria, a total of 52 articles were declared eligible and included in the further analysis process. These articles provide a comprehensive overview of chatbot applications in the context of training, including implementation trends, results achieved, and challenges faced in the context of software testing skills development.

Table 1: *Inclusion and Exclusion Criteria*

Inclusion Criteria
Articles published in peer-reviewed scientific journals
Written in English
Focus on Chatbots and Training
Published in the year range [2021 - 2025]
Accessible for free (*Open Access*)
Exclusion Criteria
Articles in the form of editorials, opinions, news, books or book chapters
Written not in English
Not focusing on Chatbots and Training
Published before 2021
Not accessible for free

Table 1 summarizes the selection criteria used to determine which articles to include and exclude in this systematic review. These criteria are designed to ensure that the selected studies are relevant to the research focus, have good methodological quality, and are reliable for further analysis. The inclusion criteria included articles published in peer-reviewed scientific journals, written in English, focused on the use of chatbots in the context of training or skills development, published from 2021 to 2025, and conducted in an organizational or corporate environment.

In contrast, studies excluded from the analysis include nonscientific articles such as editorials, opinions, news, books, or book chapters; studies that focus on the application of chatbots in nonvocational schools or communities that are not related to the work context; and those that are not in English. Through the application of these criteria, the selection process is carried out strictly and systematically, thus ensuring that only relevant and strong methodological articles

are included. Of the 303 articles identified at the initial stage, 52 met the inclusion criteria and were selected for further analysis. These articles form the basis of the synthesis in this review to describe the practices, impacts, and challenges in the utilization of chatbots for professional training in the work environment.

Table 2: *Quality criteria*

Quality 1: Is the purpose of the study clearly explained?
Quality 2: Is the role of chatbots or virtual agents in the training process explained in detail?
Quality 3: Are the training methods (e.g.: simulations, platforms, scenarios) described systematically?
Quality 4: Are the outcomes described and supported by verifiable data or findings?
Quality 5: Is the study conducted in an organizational/corporate context, and relevant to professional training?
Quality 6: Are the challenges or limitations of using chatbots in that context explicitly discussed?
Quality 7: Does the author provide any recommendations or suggestions for further research or implementation?

Stage 2 is the third step in the systematic review process, which includes a thorough evaluation of the quality of each journal article, as described in Table 2. These seven quality criteria were chosen to ensure that each study included in the review substantially contributes to an understanding of the application of chatbots in the context of future workforce training. The first criterion emphasizes the importance of the clarity of study objectives as a basis for assessing the relevance and focus of research. Furthermore, a detailed description of the role of chatbots or virtual agents is needed to understand how this technology is used in real life during the training process. Training methods that are described systematically through simulations, digital platforms, and work scenarios ensure that the intervention process can be evaluated and replicated. The assessment of the training results, especially if supported by empirical data, is an indicator of the effectiveness of chatbot implementation. In the context of on-the-job training, it is also important to identify whether the study is conducted in a professional or organizational setting rather than purely formal education. Explicit discussions of the challenges and limitations of technology help uncover potential obstacles to its application in the real world. Finally, the recommendations given by the authors show critical reflection on the findings and open the direction for further research development or implementation. As such, these seven criteria are designed to screen studies that are not only informative but also relevant, applicable, and of strategic value for the development of educational chatbots in professional training.

Among the 52 articles that passed the initial screening stage, a number of articles were removed because they did not meet the predetermined quality criteria. Table 3 presents a sample of how scoring in the screening stage works. The evaluation was conducted by considering the extent to which the articles were aligned with the objectives of the study, describing in detail the role and use of chatbots in training, the training methods used, and the data-supported results. Each abstract of the 10 articles collected was systematically evaluated on the basis of seven predetermined quality criteria. The assessment process was conducted by three researchers, who carefully read the abstracts to determine the extent to which each article met each criterion. Each criterion was assessed using a binary scale:

Score 1 (Yes): Criteria are met and can be explicitly identified in the abstract.
Score 0 (No): The criteria are not mentioned or not clear enough in the abstract.

Table 3: *Results of the quality criteria assessment*

Paper Code	Result							Total
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	
A-01	0	1	1	1	1	1	1	6
A-02	1	1	1	1	1	1	0	6
A-03	1	1	1	1	0	0	1	5
A-04	1	0	1	1	1	0	0	4
A-05	1	1	1	1	1	1	1	7
A-06	0	1	1	1	1	0	0	4
A-07	0	1	1	1	0	0	0	3
A-08	0	1	0	1	0	0	0	2
A-09	0	1	1	1	0	0	0	3
A-n	...	...	...	...	...	...	...	...

After the researcher's assessment is conducted, the results of the assessment are discussed collectively in team discussions to reach a consensus. Differences of opinion are resolved through discussion until a mutual agreement is reached. As a result of the discussion, it was agreed that the article would be retained for further analysis if it met 7 quality criteria. This approach ensures that only relevant, good-quality, and contextual articles on the main topic of the use of chatbots in workforce training are involved in an in-depth literature review. The results of the evaluation resulted in 34 articles that were selected for further reading and research.

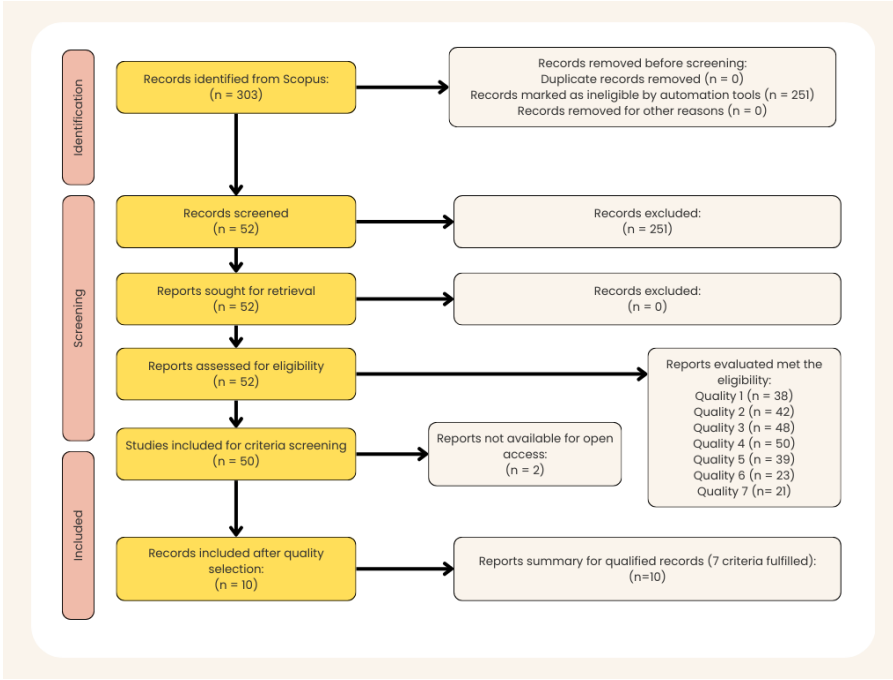
Table 4: *Summary of quality criteria*

Quality criteria	Paper Count	Percentage
Quality 1	38	73.08
Quality 2	42	80.77
Quality 3	48	92.31
Quality 4	52	100.0
Quality 5	39	75.0
Quality 6	23	44.23
Quality 7	21	40.38

Across 52 abstracts, Table 4 shows that the percentage of said quality criteria is fulfilled in this research, and the average criterion fulfillment was 72.25%. Reporting of outcomes was near-universal (quality 4: 100.00%, 52/52), and training methods were typically described in sufficient detail (quality 3: 92.31%, 48/52). Most abstracts articulated the role of chatbots/virtual agents in the training process (Quality 2: 80.77%, 42/52) and were situated in organizational/professional contexts (Quality 5: 75.00%, 39/52). A clear statement of purpose appeared in approximately three-quarters (Quality 1: 73.08%, 38/52). In contrast, fewer abstracts explicitly discussed challenges or limitations (quality 6: 44.23%, 23/52) or offered recommendations/future work (quality 7: 40.38%, 21/52). This pattern suggests that abstracts in this area prioritize methods and reporting results, whereas reflective elements (limitations and recommendations) are underreported at the abstract level.

This systematic review process involves three main stages, as illustrated in Figure 1: identification, screening, and inclusion. At the identification stage, literature sources are collected from various databases and academic registers. After that, duplication and exclusion of irrelevant documents are eliminated on the basis of the results of automatic detection and other considerations. Furthermore, at the screening stage, the titles and abstracts of all the remaining records are thoroughly examined to ensure their suitability with the focus of the study. Articles that pass this stage are then searched for the full-text version for a feasibility assessment. If there are documents that cannot be accessed, they are removed from the process. In the final stage, namely, the feasibility assessment, the available documents are evaluated in depth on the basis of the inclusion and exclusion criteria that have been set. Articles that did not meet methodological standards or were not relevant to the research topic were excluded. Only the articles deemed to meet the quality and relevance criteria were ultimately included in this systematic review analysis. This phased approach ensures that the results of the review are valid, transparent, and accountable.

Figure 1. *Flowchart of the screening process*



RESULTS

This section presents the findings and interpretation of the research results based on eight research questions (RQs) that have been formulated previously. The first five research questions were answered through a comprehensive content analysis of ten relevant empirical articles, which discussed the application of chatbots and virtual agents in the context of training and education. The analysis was carried out thematically to identify usage patterns, feedback mechanisms, system design, and their influence on learning processes and outcomes. Moreover, the last three research questions were answered through a bibliometric approach to map the scientific landscape, intellectual structure, and current research trends in related fields. The combination of these two approaches provides a comprehensive overview of the contribution of chatbot technology in supporting pedagogical innovation and skill development in various sectors of education and professional training.

How are chatbots and virtual agents used to support training and skill development in both professional and educational contexts?

Chatbots and virtual agents have been used extensively in a variety of educational and professional contexts as flexible, responsive, and sometimes superseding human-based simulations. In education, for example, Galiç et al. (2025) applied ChatGPT as a virtual student to train math teachers in the skills of "noticing rationality" through realistic teaching scenarios. This approach allows teachers to practice responding to different types of student rationality in simulations that resemble a real classroom. Moreover, in the realm of medical education, Kaur et al. (2021) and Chen et al. (2023) used chatbots as virtual patients for anamnesis training and clinical interviews, providing students with the opportunity to practice safely in clinical situations that are difficult to replicate in person.

In the field of higher education, Sánchez-Vera (2025) utilizes GPT-based chatbots as interactive tutors that support the independent learning of early childhood education students through the ability to answer questions and explain complex concepts. The use of ChatGPT in a more exploratory form is also reported by Christozov & Toleva-Stoimenova (2024) to help students and lecturers draft academic assignments, create outlines, and understand lecture material, although not formally, for training.

Moreover, in professional contexts such as nutrition and public health, Ashton et al. (2023) developed chatbots to support healthy diet behavior changes through interactive features such as recipe recommendations and healthy eating reminders. In the field of anesthesiology, Sardesai et al. (2024) trained doctors to use virtual patient avatars in preoperative consultation scenarios to improve procedural communication. With respect to risk assessment training in the field of criminology, Raiche et al. (2023) used 3D avatars to help students simulate assessments of juvenile lawbreakers.

Finally, in the field of educational technology, Essel et al. (2024) developed a WhatsApp-based VoiceBot to help students learn HTML and CSS independently, strengthening their technical skills outside of formal classrooms. Overall, chatbots are used for various forms of training, from scenario-based teaching and self-tutoring to clinical and professional simulations. Their roles are highly contextual, ranging from learning assistants and interactive dialog partners to temporary substitutes for human interaction in high-risk or expensive training if performed in person.

Table 5: *Chatbot Roles in the Certain Sector and Target Audience*

Studies	Context/Sector	Target Users	The Role of Chatbots
Galic et al. (2025)	Teacher Training	In-service math teacher	Virtual students in classroom scenarios
Kaur et al. (2021)	Medical Education	Medical students	Virtual patient for anamnesis
Chen et al. (2023)	Nursing Education	Nursing students	Virtual patient clinical practice
Sanchez-Vera (2025)	Early Childhood Education	Early Childhood Students	Self-tutor for exam preparation
Jose & Jose (2024)	General Education	Lecturers and educators	Content and language virtual assistant
Ashton et al. (2023)	Nutrition & Health	Young adults	Interactive diet advisor
Christozov & Toleva-Stoimenova (2024)	Higher Education	Students & lecturers	Informal academic companion
Sardesai et al. (2024)	Medical Education	Anesthesiologist	Patient avatar for consultation
Raiche et al. (2023)	Criminology	Criminology students	Risk assessment simulation
Essel et al. (2024)	Educational Technology	Student programming	Voice tutor for HTML/CSS

Table 5 provides a more convenient presentation in table format for mapping the role of chatbots in training by field and the target of the research conducted by each author.

What are the types of integrated feedback mechanisms and their impact on learning outcomes?

The feedback mechanisms integrated in the use of chatbots and virtual agents vary widely, depending on the purpose of the training and the type of skills developed. Most studies show that chatbots are able to provide real-time feedback, which encourages immediate reflection and instant improvement in the learning process. In the context of teacher training, Galiç et al. (2025) reported that two-way interaction with ChatGPT allows math teachers to evaluate and adjust their questions to rational or irrational responses provided by chatbots. This feedback has been shown to improve teachers' ability to "notify rationality", an important reflective skill in dialogical learning.

Moreover, Chen et al. (2023) reported that nursing students want real-time and end-of-session feedback that can help them correct errors in medical history retrieval. A similar finding was reported by Sánchez-Vera (2025), who reported that interaction with the GPT chatbot improved concept comprehension in 95.7% of students and improved their satisfaction with clarifying questions in 91.4%. Here, the effectiveness of the feedback is strongly influenced by the student's ability to compose the right prompt. In the context of professional training, Raiche et al. (2023) reported that chatbots with 3D avatars resulted in high levels of

satisfaction, especially with respect to ease of use and the perception that the system helps with risk assessment training. Nonetheless, inaccurate responses are still a challenge in some cases. Sardesai et al. (2024) also reported that trainee doctors rated the chatbot as quite accurate (median score 8/10) and easy to use (9/10), although it has not been quantitatively proven to improve clinical performance.

In the field of educational technology, Essel et al. (2024) reported that WhatsApp-based chatbots provide real-time verbal feedback that significantly improves posttest scores in HTML & CSS training. The bot also provides relevant contextual links and answers, reinforcing students' self-paced learning experiences.

In general, the most effective feedback mechanisms are as follows:

1. Real-time, thus encouraging direct reflection (e.g., Galiç et al., 2025; Sánchez-Vera, 2025; Chen et al., 2023).
2. Personal and contextual, so that it is relevant to the needs of the user (Essel et al., 2024; Ashton et al., 2023).
3. Accompanied by end-of-session support or postworkout reflection (Chen et al., 2023; Sardesai et al., 2024).

The impacts on learning outcomes include improved concept comprehension, user satisfaction, improved test scores, and increased confidence in real-life situations. Table 6 summarizes the types of feedback provided by chatbots and their impact on learning outcomes across studies.

Table 6: *Feedback type and its impact on the study results*

Studies	Feedback Type	Impact on Learning Outcomes
Galic et al. (2025)	Real-time interactive	Increased teacher noticing
Kaur et al. (2021)	Real-time (REST API)	Limitations of formal impact
Chen et al. (2023)	Real-time and end-of-session	Increased trust & readiness
Sanchez-Vera (2025)	Real-time question-based	Improved understanding of test concepts & values
Jose & Jose (2024)	Real-time (subjective perception)	Not quantitatively measured
Ashton et al. (2023)	Indirect, exploration-based	Increased diet awareness
Christozov & Toleva-Stoimenova (2024)	Non-specific	Useful in academic assignments
Sardesai et al. (2024)	Post-interaction surveys	High comfort, realistic perception
Raiche et al. (2023)	Surveys and user experience	Assisting with risk assessment exercises
Essel et al. (2024)	Real-time verbal	Posttest scores improved, positive experience

What are the commonly reported challenges in deploying chatbots for training simulations?

Various challenges arise in the implementation of chatbots for training simulations, both in educational and professional contexts. One of the main challenges that has been overcome in many studies is the inconsistent quality of responses, especially when chatbots fail to understand context or provide accurate information. For example, Galiç et al. (2025) reported that ChatGPT sometimes generates inconsistent mathematical responses, making it difficult for teachers to effectively guide virtual students. A similar thing was discovered by Essel et al. (2024), who noted that their WhatsApp chatbot is not always capable of providing complete or relevant answers, especially if the question is too complex or uses certain accents in voice input. In the field of medical education, Kaur et al. (2021) and Chen et al. (2023) highlight concerns over the lack of realism and empathy in chatbot interactions, as well as limitations in non-verbal interactions that are important in clinical contexts. This reduces the expected value of interpersonal training. Sardesai et al. (2024) reported that their chatbot exhibited "hallucination" (providing false or fictitious information), overly formal language, and copyright issues with medical content.

In addition to the quality of responses, users' dependence on chatbots is a concern in itself. Sánchez-Vera (2025) notes that excessive use can reduce learning outcomes, especially if students are unable to design effective prompts. Jose & Jose (2024) even reported that reliance on chatbots can reduce creativity and critical thinking and can pose ethical risks such as plagiarism.

Other challenges include the following:

1. Technical barriers include a system that is heavily run or unstable software (Raiche et al., 2023).
2. Trust issues the quality of answers and ethical confusion over the use of AI in academic assignments (Christozov & Toleva-Stoimenova, 2024).
3. Mismatches between task types and chatbot capabilities can lead to user frustration or confusion (Jose & Jose, 2024; Christozov & Toleva-Stoimenova, 2024).
4. User preferences for human interaction are real, especially in the context of sensitive topics such as eating disorders or emotional support (Ashton et al., 2023).

Overall, the challenges of chatbot implementation are rooted in the gap between user expectations of the realism of interaction and the limitations of current AI technologies, as well as the need for training in designing prompts and understanding the context of use wisely. Although the technology is promising, user readiness, system design, and pedagogical support remain the determining

factors of success. The number of obstacles that arise from the ten papers that have been analyzed is shown in Figure 1.

How do chatbot design features such as conversational style, empathy, and AI sophistication affect user engagement and training effectiveness?

Chatbot design features, including conversational style, empathy level, and responsiveness, play crucial roles in determining the extent to which users feel engaged and benefit from training. One of the design elements that is widely appreciated is the ability of chatbots to adjust language and conversational styles. Essel et al. (2024), for example, designed a VoiceBot with a motivational style that greets users like real lecturers (e.g., "Good job!"), uses 1,000+ personalized intents, and produces an interface that is considered to be very user-friendly. This has been proven to increase engagement and motivation to learn.

Ashton et al. (2023) also highlight the importance of empathetic design and positive conversational styles in chatbots for the promotion of healthy diets. They emphasize that users tend to be more engaged when chatbots include features such as nudging, personalization, and animated avatars, which create an authentic and communicative impression. However, not all designs have a positive effect. Kaur et al. (2021) and Chen et al. (2023) reported that the lack of empathic expression and nonverbal interaction in chatbots leads to less realistic experiences, especially in training that involves interpersonal communication. Chen suggested the use of more expressive and emotional humanoid or cyborg designs to improve this aspect.

In terms of AI sophistication, Sánchez-Vera (2025) shows that chatbot designs that utilize zero-shot, few-shot, and chain-of-thought prompting techniques yield high accuracy (98%) and consistency (96%). However, its performance remains highly dependent on the quality of the user's prompts. Such chatbots are less effective at building motivation or setting long-term learning strategies without additional pedagogical support. Moreover, Raiche et al. (2023) reported that user preferences for chatbot design are influenced by individual factors such as self-efficacy and learning style. Users with analytical learning styles and high anxiety are more receptive to chatbots, as long as the design is customizable.

In the context of teacher training, Galiç et al. (2025) reported that ChatGPT's design, which is capable of generating "rational" and "irrational" responses, allows trainers to hone reflective thinking skills. However, limitations in the dimension of "communicative rationality" actually cause some teachers to fail to guide virtual students optimally.

In general, the influence of chatbot design on training engagement and effectiveness includes the following:

1. Positive and contextual conversational styles improve user comfort and retention (Essel, Ashton).
2. Emotional responses and empathetic expressions reinforce the realism of interactions in interpersonal contexts (Chen, Kaur).
3. The sophistication of AI (prompting & natural language understanding) affects the accuracy, consistency, and depth of learning (Sánchez-Vera, Galiç).
4. Design flexibility that accounts for user characteristics drives adoption and trust (Raiche, Christozov).

A design that is not only technical but also pedagogical and affective is the key to the successful implementation of chatbots as effective training partners. Table 7 clarifies the features of chatbots that are expected to provide certain insights to users.

Table 7: *Design of Chatbot Features on the Impact Toward the User*

Key Design Features	Impact on Users
Language style & motivation	Increases comfort and motivation (Essel et al., 2024)
Empathy & emotional expression	Increase realism and empathy in communication (Chen et al., 2023)
Prompt sophistication (zero-shot few-shot)	Improve the accuracy and consistency of answers (Sanchez-Vera, 2025)
Rational & irrational response	Improving teachers' noticing skills (Galic et al., 2025)
Personalization & avatar	Increase user engagement and acceptance (Ashton et al., 2023
Interactivity & flexibility	Provide a real-life classroom-like experience (Sardesai et al., 2024)

Is there any evidence regarding the effectiveness of virtual agents in simulating communication or interview-based training scenarios?

There is strong evidence that virtual agents, whether in the form of text-based, voice-based, or avatar-based chatbots, are effectively used to simulate communication-based or interview-based training scenarios, especially in medical education, nursing, teacher training, and criminology. One of the standout studies is that of Sardesai et al. (2024), who developed a virtual avatar of a patient using the Convai platform for the training of anesthesiologists. In the pre-operative consultation simulation (interscalene block), participants gave a median score of 9/10 for ease of use and 8/10 for response accuracy and reported high levels of comfort (87%) during interaction. Although there are no quantitative data on performance improvement, the participants consider this simulation to be realistic and suitable for replacing human actors in training and exams (mock exams).

The study of Galiç et al. (2025) also provides strong evidence that ChatGPT, when used as a "virtual student" with misconceptions, is capable of creating an authentic practice environment for teachers to hone questioning skills and direct rational interactions. These simulations resemble real classroom

conditions, especially in practicing reflective responses to students' answers that are not entirely correct. In the context of nursing, Chen et al. (2023) reported that chatbots allow students to practice taking patients' health histories in scenarios that may not be available in real clinical practice. Students feel more prepared and confident before facing real patients, although the realism of interaction and empathic expression is still considered suboptimal.

Raiche et al. (2023) used a 3D avatar chatbot in risk assessment training for juvenile offenders. The participants stated that the system helps the learning process, although there are still technical shortcomings. The study also highlights that the effectiveness of simulations is highly dependent on the suitability of the chatbot design for the user's learning style. In the field of programming, Essel et al. (2024) reported that voice-based chatbots are also effective at providing answers to technical and practical questions, even though the interactions are more factual than interpersonal communication-based.

However, some limitations remain, especially regarding the lack of nonverbal and empathic expression, as reported by Kaur et al. (2021). Concerns have arisen that chatbot-based training is not capable of fully training interpersonal and empathy skills, which are crucial in medical education. In conclusion, evidence from various studies indicates that virtual agents are effectively used as a means of scenario-based communication training, especially to improve initial readiness, confidence, and mastery of question-and-answer formats. Nonetheless, for advanced interpersonal skills training, virtual agents still need further development, especially in terms of realism, emotional expression, and cultural context.

How have the number of publications on the application of chatbots in training or simulation evolved from year to year, and what are the main journal sources that publish the most on this topic?

Figure 2. *Evolution of Publication in Chatbots in Training or Simulation*

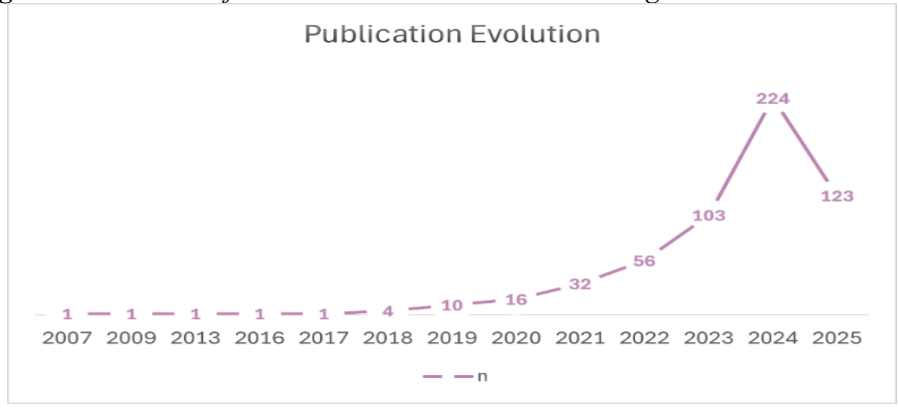
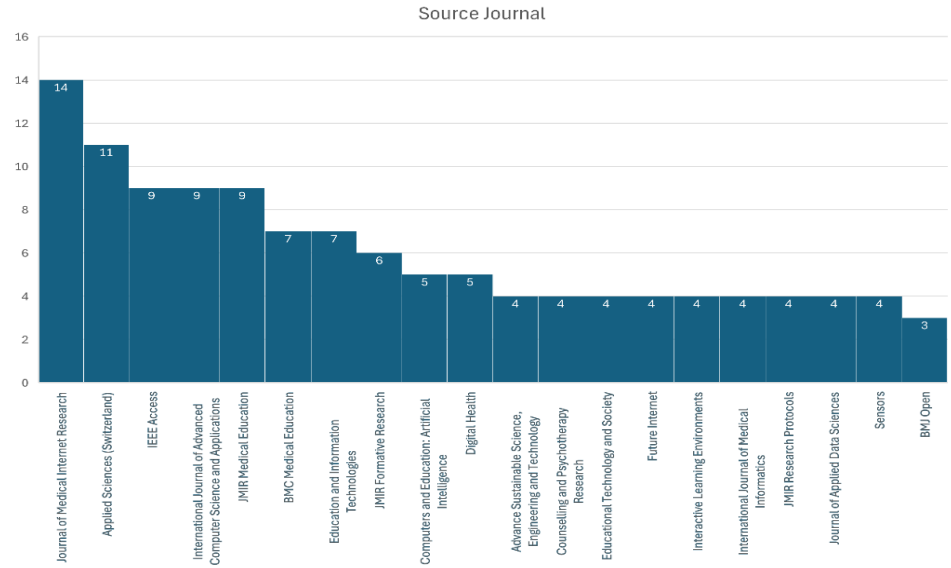


Figure 3. *Impactful Source Journal*



The evolution of the number of chatbot-related publications in the context of training and simulation from 2007 to 2025 is shown in Figure 3. For more than the first decade (2007–2017), the number of publications tended to stagnate, with only one publication per year recorded. The increase began in 2018 to 2020, indicating an initial phase of growth, with a gradual increase from 4 to 16 publications per year. Exponential growth has occurred since 2021, with the number of publications increasing significantly from 32 (2021) to 56 (2022) to 103 (2023). The year 2024 will be the peak of productivity, with 224 publications, indicating a very high surge in research interest in this topic. Moreover, the number of publications in 2025 is recorded at 123, but this figure is likely to still be temporary because data collection has not been completed for the current year. This pattern reflects the dynamics of the development of chatbot topics in training as one of the research areas that has grown rapidly in the past five years.

The distribution of scientific journals that publish the most chatbot-related studies in training and simulation is shown in Figure 4. The most prolific journal is the Journal of Medical internet Research (14 articles), followed by Applied Sciences (Switzerland) (11 articles), IEEE Access, the International Journal of Advanced Computer Science and Applications, and JMIR Medical Education, which each contribute 9 publications. In general, the dominant journals come from three main domains: health and medical (such as JMIR, BMC, and digital health), educational and computer technology (IEEE Access, Education and Information Technologies, Computers and Education: Artificial Intelligence), and applied sciences and engineering. This distribution reflects the cross-disciplinary nature of

chatbot research in training and simulation, which combines technological, educational, and applicative approaches across a wide range of professional sectors.

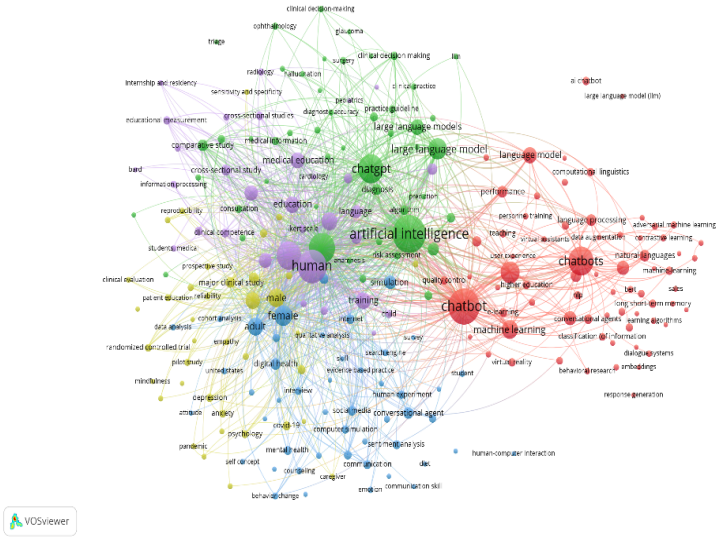
Figure 4. *Map Visualization for the 15 Countries with the Most Publication*

On the basis of the map in Figure 4 showing the geographical distribution of publications, 15 countries are most active in publishing scientific papers related to the application of chatbots in training or simulation for workers. These countries cover regions from North America, Europe, Asia, and Oceania, demonstrating the global spread of interest in AI-based training technologies. In the North American region, the United States and Canada occupy the top positions in terms of the contribution to the number of publications. In Europe, the UK, Germany, France, Spain, Italy, and the Netherlands are significant research centers on this topic, demonstrating the strong role of European educational and health institutions in adopting chatbot-based learning technologies.

Moreover, in Asia, China, India, South Korea, and Indonesia are countries with active contributions, indicating the rapid adoption of AI technology in the workforce training and education sectors in developing and developed countries. In the Oceania region, Australia recorded a number of important publications, particularly in the context of medical education and training. In addition, Malaysia and Singapore also showed regional involvement in Southeast Asia. This map emphasizes that research on chatbots for professional training is not just a local trend but also a global phenomenon involving cross-continental collaboration.

What are some of the main trends and topics that are developing in the literature on chatbots for training or simulation based on keyword co-occurrence analysis?

Figure 5. *Network Visualization*

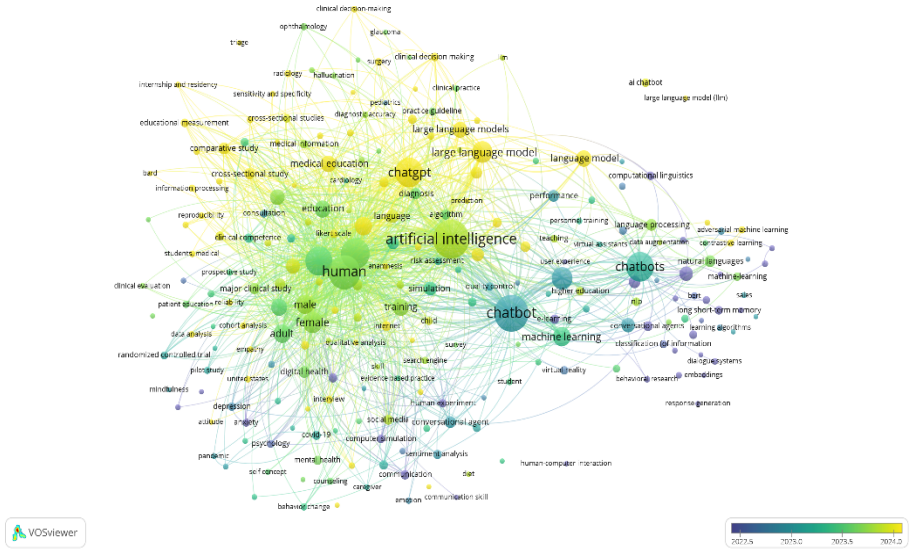


A map of keyword co-occurrence networks from chatbot-related publications in the context of training and simulation is shown in Figure 6 (network visualization). This visualization divides keywords into color clusters on the basis of their topic relatedness:

1. The red cluster focuses on the technological aspects of chatbots and artificial intelligence, with the dominant keywords such as chatbots, machine learning, language models, natural language processing, and conversational agents. This cluster shows the technical foundations and development of the system.
2. The green clusters are centered on keywords such as artificial intelligence, chatgpt, large language models, and diagnostics, which indicate the integration of LLM and AI in simulations, particularly in clinical or medical contexts.
3. The purple and yellow clusters emphasize evaluative and educational aspects, with terms such as medical education, clinical competence, the Likert scale, and comparative study.
4. The blue and light blue clusters focus on the social and psychological aspects of chatbot use, such as mental health, communication, sentiment analysis, and the user experience.

Network visualization reveals that chatbot research in training is multidisciplinary, with strong branching in terms of technology, clinical education, and human–machine interaction in emotional and social contexts.

Figure 6. *Overlay Visualization*



An overlay visualization based on the year the keyword appeared in the publication is shown in Figure 6 (Overlay Visualization), with a color scale from purple (old) to yellow (recent).

1. Bright yellow keywords such as chatgpt, large language models, language models, and hallucination are relatively new topics, signaling the latest trend (2023–2024) in the adoption of LLMs on chatbots for training.
2. Keywords such as mental health, pandemic, anxiety, and qualitative studies tend to be blue to green, indicating that the intensity of the research that emerged from approximately 2021–2022 is especially related to the impact of the pandemic and the use of chatbots in the context of emotional support.
3. Old terms such as human, training, education, and artificial intelligence are in the middle of the spectrum (green), indicating the consistency of these themes in recent years.

This overlay visualization shows the dynamics of the development of the topic. The research focus shifted from a post-pandemic psychology and public

health-based approach to further exploration of cutting-edge technologies such as ChatGPT and LLMs for AI-based training.

What are the most influential (highly cited) documents in chatbot research for training or simulation, and how does it contribute to the intellectual structure of this field based on cocitation analysis?

Table 8 presents the top articles by number of citations, which show the most influential contributions in the chatbot literature on various domains:

1. Dominant Year: The majority of influential articles were published in 2023, reflecting a surge in research and global attention to ChatGPT and large language model (LLM) applications in training and healthcare.
2. Most Cited Articles: An article by Lee et al. (2023) discussing the potential and risks of GPT-4 in the medical world occupies the top position, with 913 citations, published in the New England Journal of Medicine, a journal with a very high scientific impact.
3. Systematic Contribution: A systematic review by Kuhail et al. (2023), which examined the use of educational chatbots, received 431 citations, showing the importance of review studies in consolidating research directions.
4. Psychological and Emotional Aspects: An article by Ho et al. (2018) that explores the emotional effects of postinteraction with chatbots received 372 citations, showing the psychosocial dimension as a significant exploratory field.
5. Clinical and Medical Contributions: Articles such as Mihalache et al. (2023) and Rao et al. (2023) demonstrate a strong focus on chatbots in clinical assessments and workflows, signaling the integration of the technology in the AI-driven medical world.
6. Ethical and Social Issues: An article on creepiness (psychological discomfort) by Rajabolelina et al. (2021) underscores the importance of ethical aspects and user perception in chatbot adoption.

Table 8: *Top 10 Highest Cited Document*

Title	Author	Year	Cited By	Source Journal
Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine.	Lee P.; Bubeck S.; Petro J.	2023	913	New England Journal of Medicine
Interacting with educational chatbots: A systematic review	Kuhail M.A.; Alturki N.; Alramlawi S.; Alhejori K.	2023	431	Education and Information Technologies
Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot	Ho A.; Hancock J.; Miner A.S.	2018	372	Journal of Communication
Performance of an Artificial Intelligence Chatbot in Ophthalmic Knowledge Assessment	Mihalache A.; Popovic M.M.; Muni R.H.	2023	174	JAMA Ophthalmology
Bringing chatbots into education: toward natural language negotiation of open learner models	Kerly A.; Hall P.; Bull S.	2007	168	Knowledge-Based Systems
Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study	Rao A.; Pang M.; Kim J.; Kamineni M.; Lie W.; Prasad A.K.; Landman A.; Dreyer K.; Succi M.D.	2023	166	Journal of Medical Internet Research
ChatGPT outperforms humans in emotional awareness evaluations	Elyoseph Z.; Hadar-Shoval D.; Asraf K.; Lvovsky M.	2023	157	Frontiers in Psychology
Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot	Rajaobelina L.; Prom Tep S.; Arcand M.; Ricard L.	2021	140	Psychology and Marketing
A virtual counseling application using artificial intelligence for communication skills training in nursing education: Development study	Shorey S.; Ang E.; Yap J.; Ng E.D.; Lau S.T.; Chui C.K.	2019	128	Journal of Medical Internet Research
Large AI Models in Health Informatics: Applications, Challenges, and the Future	Qiu J.; Li L.; Sun J.; Peng J.; Shi P.; Zhang R.; Dong Y.; Lam K.; Lo F.P.-W.; Xiao B.; Yuan W.; Wang N.; Xu D.; Lo B.	2023	120	IEEE Journal of Biomedical and Health Informatics

Table 8 also shows the ten scientific articles with the highest citation rates that discuss the use of chatbots in various contexts, ranging from education and medicine to social psychology. The most influential article comes from Lee et al. (2023), who published in the New England Journal of Medicine, highlights the benefits and risks of using GPT-4 in the medical field and has garnered more than 900 citations. The dominant trend in 2023 suggests that the rise of large language model technologies such as ChatGPT is driving rapid exploration across various disciplines. Other articles explore important topics such as the effectiveness of

chatbots in clinical training (Rao et al., Mihalache et al.), the emotional experience of users (Ho et al.), and ethical factors such as the perception of "creepiness" (Rajabolelina et al.). Several studies have also highlighted the role of chatbots in communication skills education and training. Overall, these articles reveal that chatbot research has evolved from mere text interactions to complex applications in medical, educational, and psychological contexts and is becoming an important focus in the realm of applied AI.

DISCUSSION AND CONCLUSIONS

Discussion

The results of this review confirm that the landscape of educational chatbots has undergone a significant transformation, transitioning from simple rule-based systems to advanced generative AI (GenAI) architectures powered by large language models (LLMs) (Davar et al., 2025; Wollny et al., 2021). This technical evolution, which has been particularly evident since the surge of research in 2023, allows for fluid, contextually aware dialogs that are better suited for the complexities of professional training than traditional scripted interactions are (Huang et al., 2025; Kuhail et al., 2023).

Simulation-Based Training and Professional Development A dominant theme across the literature is the effectiveness of chatbots as simulation tools in high-stakes environments (Labadze et al., 2023; Raiche et al., 2023). In healthcare, chatbots serve as virtual patients, enabling nursing and medical students to practice history-taking and diagnostic reasoning without risking patient safety (Chen et al., 2023; Kaur et al., 2021). Similarly, scenario-based training using tools such as ChatGPT allows in-service teachers to act as "virtual tutors" for a simulated student, thereby improving their "professional noticing" skills and the ability to attend to and interpret student reasoning in real time (Galiç et al., 2025). These applications demonstrate that chatbots can bridge the gap between theoretical knowledge and practical skill acquisition (Sardesai et al., 2024).

Personalization and Adaptive Learning The integration of knowledge graphs with LLMs has optimized the precision of these tools, allowing for adaptive responses that align with specific academic content and individual student progress (Villegas-Ch et al., 2024). Research has indicated that while chatbots excel in structured subject areas, the ability of these agents to provide 24/7 accessibility and personalized feedback is a major driver of student motivation and engagement (Huang et al., 2025; Jose & Jayaron Jose, 2024). Furthermore, interacting with a chatbot can reduce learner anxiety, as the agent provides a nonjudgmental environment for making mistakes during the training process (Davar et al., 2025; Labadze et al., 2023).

The successful implementation of task–technology fit (TTF) and human factors in a workplace setting is heavily dependent on task–technology fit (TTF),

ensuring that the chatbot's capabilities are aligned with the professional tasks of the worker (Soodan et al., 2024). Furthermore, digital literacy acts as a critical moderator; workers with greater digital skills are more likely to engage with and benefit from AI-driven personalized feedback (Yaseen et al., 2025). While social networks and peer influence among professionals can accelerate the adoption of these tools, concerns regarding information accuracy (hallucinations) and academic integrity remain significant barriers (Davar et al., 2025; Macdonald et al., 2023; Soodan et al., 2024).

Conclusion

In conclusion, AI chatbots have evolved from simple administrative aid to pivotal pedagogical agents for workforce training (Kuhail et al., 2023; Wollny et al., 2021). They offer transformative opportunities for safe, simulation-based learning and personalized instruction (Galiç et al., 2025; Raiche et al., 2023). However, the rise of GenAI introduces significant risks related to misinformation (hallucinations), ethical integrity, and a potential decline in critical thinking (Davar et al., 2025; Jose & Jayaron Jose, 2024). The "clinical copilot" model illustrates that while AI can assist in complex workflows and decision support, human oversight remains indispensable for guaranteeing the accuracy and ethical validity of training outcomes (Macdonald et al., 2023; Rao et al., 2023). Ultimately, the success of chatbots in worker training depends on achieving a balance between technological innovation and human-centric pedagogical strategies (Alabbas & Alomar, 2025).

IMPLICATIONS

Implications for Corporate and Vocational Policy

This study implies that organizations must prioritize upskilling and AI literacy as a precursor to digital transformation (Li, 2024; Obermayer et al., 2022). Because digital literacy is a moderator of success, training programs should focus not only on job-specific skills but also on the ability to critically evaluate and "prompt" AI-generated content (Yaseen et al., 2025).

Implications for Instructional Design

Instructional designers should move toward "zero-coding" architectures, such as VoiceBots integrated into common platforms such as WhatsApp, to democratize access in resource-constrained environments (Essel et al., 2025). Furthermore, the integration of chatbots with learning management systems (LMS) such as Moodle using knowledge graphs can ensure that the AI remains a reliable and accurate "assisting agent" rather than a source of misinformation (Villegas-Ch et al., 2024).

Implications for pedagogical theory

This research extends TTF theory by demonstrating that professional adoption is influenced by social networks and psychological factors (Soodan et al., 2024). Chatbots that act as "Mentoring Agents" and support metacognition and self-regulated learning offer more comprehensive pedagogical value than those that merely deliver content do (Kuhail et al., 2023; Wollny et al., 2021).

REFERENCES

Alabbas, A., & Alomar, K. (2025). A Weighted Composite Metric for Evaluating User Experience in Educational Chatbots: Balancing Usability, Engagement, and Effectiveness. *Future Internet*, *17*(2), 64. https://doi.org/10.3390/fi17020064

Ashton, L. M., Adam, M. T., Whatnall, M., Rollo, M. E., Burrows, T. L., Hansen, V., & Collins, C. E. (2023). Exploring the design and utility of an integrated web-based chatbot for young adults to support healthy eating: A qualitative study. *International Journal of Behavioral Nutrition and Physical Activity*, *20*(1), 119. https://doi.org/10.1186/s12966-023-01511-4

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019

Babu, C. V. S., William, M. P., & Abdul Fazil, R. (2025). The Role of Chatbots in Educational Leadership: Enhancing Decision-Making and Efficiency Through AI Tools. In A. Musundire (Ed.), *Chatbots in Educational Leadership and Management* (pp. 291–320). IGI Global. https://doi.org/10.4018/979-8-3693-8734-4.ch012

Boaye Belle, A., & Zhao, Y. (2023). Evidence-based decision-making: On the use of systematicity cases to check the compliance of reviews with reporting guidelines such as PRISMA 2020. *Expert Systems with Applications*, *217*, 119569. https://doi.org/10.1016/j.eswa.2023.119569

Chang, C., Hwang, G., & Gau, M. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, *53*(1), 171–188. https://doi.org/10.1111/bjet.13158

Chen, S., & Huang, Y. (2023). Research and Application of Chatbots in the Field of Education: A Bibliometric Analysis Using VOSviewer. *2023 International Conference on Computer Engineering and Distance*

Learning (CEDL), 163–170. https://doi.org/10.1109/CEDL60560.2023.00039
Chen, X., Zou, D., Xie, H., & Wang, F. L. (2025). Educational chatbot research: Text mining and bibliometrics. *Interactive Learning Environments*, *33*(4), 2762–2780. https://doi.org/10.1080/10494820.2024.2430632
Chen, Y., Lin, Q., Chen, X., Liu, T., Ke, Q., Yang, Q., Guan, B., & Ming, W. (2023). Need assessment for history-taking instruction program using chatbot for nursing students: A qualitative study using focus group interviews. *DIGITAL HEALTH*, *9*, 20552076231185435. https://doi.org/10.1177/20552076231185435
Davar, N. F., Dewan, M. A. A., & Zhang, X. (2025). AI Chatbots in Education: Challenges and Opportunities. *Information*, *16*(3), 235. https://doi.org/10.3390/info16030235
Dr. Obermayer, N., Dr. Csizmadia, T., & Dr. Banász, Z. (2022). Companies on Thin Ice Due to Digital Transformation: The Role of Digital Skills and Human Characteristics. *International and Multidisciplinary Journal of Social Sciences*, *11*(3), 88–118. https://doi.org/10.17583/rimcis.10641
Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, *14*, 1199058. https://doi.org/10.3389/fpsyg.2023.1199058
Engeness, I., Nohr, M., & Fossland, T. (2025). Investigating AI Chatbots' Role in Online Learning and Digital Agency Development. *Education Sciences*, *15*(6), 674. https://doi.org/10.3390/educsci15060674
Essel, H. B., Vlachopoulos, D., Nunoo-Mensah, H., & Amankwa, J. O. (2025). Exploring the impact of VoiceBots on multimedia programming education among Ghanaian university students. *British Journal of Educational Technology*, *56*(1), 276–295. https://doi.org/10.1111/bjet.13504
Farah, J. C., Spaenlehauer, B., Ingram, S., & Gillet, D. (2022). A Blueprint for Integrating Task-Oriented Conversational Agents in Education. *Proceedings of the 4th Conference on Conversational User Interfaces*, 1–8. https://doi.org/10.1145/3543829.3544525
Galiç, S., Urhan, S., Dost, Ş., & Lavicza, Z. (2025). Examining Mathematics Teachers Noticing the Rationality: Scenario-Based Training with AI Chatbot. *Science & Education*. https://doi.org/10.1007/s11191-025-00618-3
García-Peñalvo, F. J. (2022). Desarrollo de estados de la cuestión robustos: Revisiones Sistemáticas de Literatura. *Education in the Knowledge Society (EKS)*, *23*, e28600. https://doi.org/10.14201/eks.28600
Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, Relational, and Emotional Effects of Self-Disclosure After Conversations With a Chatbot.

Journal of Communication, *68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Huang, W., Jiang, J., King, R. B., & Fryer, L. K. (2025). Chatbots and student motivation: A scoping review. *International Journal of Educational Technology in Higher Education*, *22*(1), 26. https://doi.org/10.1186/s41239-025-00524-2

Hwang, G.-J., & Chang, C.-Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, *31*(7), 4099–4112. https://doi.org/10.1080/10494820.2021.1952615

Jose, J., & Jayaron Jose, B. (2024). Educators' Academic Insights on Artificial Intelligence: Challenges and Opportunities. *Electronic Journal of E-Learning*, 00–00. https://doi.org/10.34190/ejel.21.5.3272

Kaur, A., Singh, S., Chandan, J. S., Robbins, T., & Patel, V. (2021). Qualitative exploration of digital chatbot use in medical education: A pilot study. *DIGITAL HEALTH*, *7*, 20552076211038151. https://doi.org/10.1177/20552076211038151

Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: Toward natural language negotiation of open learner models. *Knowledge-Based Systems*, *20*(2), 177–185. https://doi.org/10.1016/j.knosys.2006.11.014

Kohl, M., Knauer, S., & Fottner, J. (2021). Industry 4.0 in Logistics and Associated Employee Competencies – A Technology Providers' Perspective. In T. Ahram, R. Taiar, K. Langlois, & A. Choplin (Eds.), *Human Interaction, Emerging Technologies and Future Applications III* (Vol. 1253, pp. 377–383). Springer International Publishing. https://doi.org/10.1007/978-3-030-55307-4_57

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), 56. https://doi.org/10.1186/s41239-023-00426-1

Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *New England Journal of Medicine*, *388*(13), 1233–1239. https://doi.org/10.1056/NEJMsr2214184

Lefaivre, K. A., & Slobogean, G. P. (2013). Understanding Systematic Reviews and Meta-analyses in Orthopaedics: Journal of the American Academy of Orthopaedic Surgeons, 21(4), 245–255. https://doi.org/10.5435/JAAOS-21-04-245

Li, L. (2024). Reskilling and Upskilling the Future-ready Workforce for Industry 4.0 and Beyond. *Information Systems Frontiers*, *26*(5), 1697–1712. https://doi.org/10.1007/s10796-022-10308-y

Luria, E. (2024). Exploring the use of chatbots in English language learning among higher education students in Israel. In S. Papadakis & G. Lampropoulos (Eds.), *Intelligent Educational Robots* (pp. 213–232). De Gruyter. https://doi.org/10.1515/9783111352695-012

Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *Journal of Global Health*, *13*, 01003. https://doi.org/10.7189/jogh.13.01003

Mihalache, A., Popovic, M. M., & Muni, R. H. (2023). Performance of an Artificial Intelligence Chatbot in Ophthalmic Knowledge Assessment. *JAMA Ophthalmology*, *141*(6), 589. https://doi.org/10.1001/jamaophthalmol.2023.1144

Musundire, A. (2025). Chatbots as Virtual Leaders: Supporting Decision-Making and Strategic Planning in Australian Secondary Schools. In A. Musundire (Ed.), *Chatbots in Educational Leadership and Management* (pp. 1–28). IGI Global. https://doi.org/10.4018/979-8-3693-8734-4.ch001

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, n160. https://doi.org/10.1136/bmj.n160

Pattier, D. (2026). Perceptions, Uses, and Attitudes Towards ChatGPT Among University Students in the Field of Education. *Journal of Interdisciplinary Studies in Education*, *15*(2), 147–168. https://doi.org/10.32674/nh8hdt53

Proske, A., Körndle, H., Drexler, K., Kirsten, J., Schröder, K., Kucharski, S., Kubica, T., & Braun, I. (2023). Towards Designing and Evaluating an Adaptable Assistance System for Technology-Enhanced Vocational Education. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou, & T. Papathoma (Eds.), *Responsive and Sustainable Educational Futures* (Vol. 14200, pp. 618–623). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42682-7_51

Raiche, A.-P., Dauphinais, L., Duval, M., De Luca, G., Rivest-Hénault, D., Vaughan, T., Proulx, C., & Guay, J.-P. (2023). Factors influencing acceptance and trust of chatbots in juvenile offenders' risk assessment

training. *Frontiers in Psychology*, *14*, 1184016. https://doi.org/10.3389/fpsyg.2023.1184016

Rajaobelina, L., Prom Tep, S., Arcand, M., & Ricard, L. (2021). Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot. *Psychology & Marketing*, *38*(12), 2339–2356. https://doi.org/10.1002/mar.21548

Rajput, R. (2025). Use of Artificial Intelligence to Solve Problems in the Classroom: In G. Reyes Ruiz (Ed.), *New Technological Applications in the Flipped Learning Model* (pp. 137–168). IGI Global. https://doi.org/10.4018/979-8-3373-0437-3.ch005

Rao, A., Pang, M., Kim, J., Kamineni, M., Lie, W., Prasad, A. K., Landman, A., Dreyer, K., & Succi, M. D. (2023). Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study. *Journal of Medical Internet Research*, *25*, e48659. https://doi.org/10.2196/48659

Sampah, S. N. A., Nabang, M., Krampa, E. K., Issah, M., Koduah, F., Essel, H. B., & Adu-Agyem, J. (2026). Generative AI Competence and Student Engagement in Higher Education: Mediating Roles of Utilization, Autonomy, and Formal Learning. *Journal of Interdisciplinary Studies in Education*, *15*(2), 109–146. https://doi.org/10.32674/3s3cx241

Sánchez-Vera, F. (2024). Subject-Specialized Chatbot in Higher Education as a Tutor for Autonomous Exam Preparation: Analysis of the Impact on Academic Performance and Students' Perception of Its Usefulness. *Education Sciences*, *15*(1), 26. https://doi.org/10.3390/educsci15010026

Sardesai, N., Russo, P., Martin, J., & Sardesai, A. (2024). Utilizing generative conversational artificial intelligence to create simulated patient encounters: A pilot study for anaesthesia training. *Postgraduate Medical Journal*, *100*(1182), 237–241. https://doi.org/10.1093/postmj/qgad137

Schmid-Meier, C., & Schulz, L. (2025). Exploring the Inclusive Potential of AI Chatbots: Engagement and Barriers in Foreign Language Training. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction* (Vol. 15780, pp. 377–395). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-93848-1_25

Soodan, V., Rana, A., Jain, A., & Sharma, D. (2024). AI Chatbot Adoption in Academia: Task Fit, Usefulness and Collegial Ties. *Journal of Information Technology Education: Innovations in Practice*, *23*, 001. https://doi.org/10.28945/5260

Stapić, Z., Horvat, A., & Plantak Vukovac, D. (2020). Designing a Faculty Chatbot Through User-Centered Design Approach. In C. Stephanidis, D. Harris, W.-C. Li, D. D. Schmorrow, C. M. Fidopiastis, P. Zaphiris, A. Ioannou, X. Fang, R. A. Sottilare, & J. Schwarz (Eds.), *HCI International 2020 – Late Breaking Papers: Cognition, Learning and Games* (Vol.

12425, pp. 472–484). Springer International Publishing. https://doi.org/10.1007/978-3-030-60128-7_36
Suciati, S., Silitonga, L. M., Wiyaka, Huang, C.-Y., & Anggara, A. A. (2024). Enhancing Engagement and Motivation in English Writing Through AI: The Impact of ChatGPT-Supported Collaborative Learning. In Y.-P. Cheng, M. Pedaste, E. Bardone, & Y.-M. Huang (Eds.), *Innovative Technologies and Learning* (Vol. 14786, pp. 205–214). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65884-6_21
Traymbak, S., Sharma, M., Anand, A., & Shukla, A. (2024). Chatbot technology in the education sector: A bibliometrics analysis using VOS viewer. *International Journal of System Assurance Engineering and Management*. https://doi.org/10.1007/s13198-023-02230-6
Vass, V., & Kiss, F. (2023). Implementation of a Talent Development Program in Higher Education. In D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Innovative Approaches to Technology-Enhanced Learning for the Workplace and Higher Education* (Vol. 581, pp. 446–450). Springer International Publishing. https://doi.org/10.1007/978-3-031-21569-8_42
Villegas-Ch, W., Govea, J., & Gutierrez, R. (2024). Optimizing Language Model-Based Educational Assistants Using Knowledge Graphs: Integration with Moodle LMS. *IEEE Access*, 1–1. https://doi.org/10.1109/ACCESS.2024.3518952
Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are We There Yet? - A Systematic Literature Review on Chatbots in Education. *Frontiers in Artificial Intelligence*, *4*, 654924. https://doi.org/10.3389/frai.2021.654924
Woollaston, S., Flanagan, B., Ocheja, P., Toyokawa, Y., & Ogata, H. (2025). ARCHIE: Exploring Language Learner Behaviors in LLM Chatbot-Supported Active Reading Log Data with Epistemic Network Analysis. *Proceedings of the 15th International Learning Analytics and Knowledge Conference*, 642–654. https://doi.org/10.1145/3706468.3706556
Yaseen, H., Mohammad, A. S., Ashal, N., Abusaimeh, H., Ali, A., & Sharabati, A.-A. A. (2025). The Impact of Adaptive Learning Technologies, Personalized Feedback, and Interactive AI Tools on Student Engagement: The Moderating Role of Digital Literacy. *Sustainability*, *17*(3), 1133. https://doi.org/10.3390/su17031133

Alfaza Ranggana, is a researcher and educator affiliated with Indonesia University of Education (UPI), Bandung. He completed his undergraduate degree of Bachelor of Computer (B.Comp.) at Parahyangan Catholic University (UNPAR) and earn his Master of Education (M.Ed.) degree from Universitas Pendidikan Indonesia. Right now, Alfaza is currently pursuing His PhD studies in UPI, with hist focus subject in Technology and Vocational Education. His research

interests include Computer Science Education, Computational Thinking, and Extracurricular Activities.

Prof. **Lala Septem Riza**, PhD. is a full professor at the Faculty of Mathematics and Natural Sciences Education (FPMIPA), Universitas Pendidikan Indonesia (UPI), Bandung. He holds a professorship in the field of data science and big data analysis. He completed his undergraduate studies at the ITB (Institut Teknologi Bandung) and earned his doctoral degree in informatics from the Universidad de Granada. His research interests span artificial intelligence, machine learning, soft computing, and big data analysis. He has been widely published in international journals and has contributed to research spanning educational technology, computer vision, NLP-based question generation, and astronomical data analysis.

Prof. **Ade Gafar Abdullah**, PhD, was born in Bandung on November 13, 1972. He completed his undergraduate studies in electrical engineering education at IKIP Bandung (now Universitas Pendidikan Indonesia/UPI), followed by a master's degree in instrumentation physics and a doctoral degree in nuclear physics, both from Institut Teknologi Bandung (ITB). He is a Professor of Energy Computation at Universitas Pendidikan Indonesia and is affiliated with the Department of Electrical Engineering Education, Faculty of Technology and Vocational Education (FPTK). He has been serving as a lecturer at UPI since 1999, and his research spans a wide range of topics, including artificial intelligence, renewable energy, machine learning, and vocational education.