

AI-Powered Conversational Agents in the Context of Learner Engagement in Engineering Education: A Systematic Review

Roshni R¹, Dr. Anitha D², Srimathi S³, Vishnupriya R⁴

¹Research Scholar, Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering, Madurai, Tamilnadu, India (roshniramachandran08@gmail.com)

²Associate Professor, Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering, Madurai, Tamilnadu, India (anitha.dm07@gmail.com)

³Assistant Professor, Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering, Madurai, Tamilnadu, India (srima2209@gmail.com)

⁴Assistant Professor, Department of Computer Science, Lady Doak College, Madurai, Tamilnadu, India (vishnupriyamadurai26791@gmail.com)

HOW TO CITE:

Roshni R, Anitha D, Srimathi S, Vishnupriya R (2026). AI-Powered Conversational Agents in the Context of Learner Engagement in Engineering Education: A Systematic Review. International Journal of Special Education 41(25s), 338-367

ABSTRACT:

Online education is a critical part of today's education, allowing learners to learn on-the-go through digital platforms. But online learners often lack motivation, interaction, and ongoing support primarily because of the absence of human-like interaction in conventional online learning systems. To overcome these challenges, there is a growing trend in educational technologies to incorporate Artificial Intelligence (AI), such as conversational agents and virtual tutors, to offer interactive support and assistance. Conversational agents mimic human interaction and may serve as tutors, peers or assistants in online learning systems. Recent developments in Generative Artificial Intelligence (GenAI) and Large Language Model (LLM), such as GPT-based systems, have greatly improved their ability to interpret and respond to complex questions, and engage in multi-turn interactions. This paper is a systematic literature review of 54 papers from 2022 to 2025 on the elements and features of educational conversational agents. The review recognizes important inputs and outputs, recent AI approaches like Machine Learning (ML), Deep Learning (DL) and Large Language Model (LLM) approaches. It also discusses interaction approaches, Generative AI (GenAI) integration and metrics to assess learning performance, engagement, adaptability and technical aspects. The review suggests that the development of conversational educational agents should focus on long time personalization strategies with changing learner attributes and learning outcomes, integration of emotional intelligence, recommendation and ethical considerations.

Keywords: Conversational Agents, Generative AI, Large Language Models, Educational Chatbots, Human-AI Interaction, Virtual Tutoring, Artificial Intelligence

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The "Online learning" is teaching and learning that occurs online, using digital technologies, where teachers and students can be separated by time and/or space (Akpan et al., 2025; Al-Abri, 2025). It includes e-learning platforms, Learning Management Systems (LMS), online classrooms, massive open online courses (MOOCs) and blended learning. The Online learning is increasing rapidly as it allows students to study at their own pace, time and place, so that everyone can access education. There is a lot of research showing online students often struggle with issues of motivation, interaction, feedback and isolation, which all affect learning (Albornoz-De Luise et al., 2023; Baabdullah et al., 2022; Întorsureanu et al., 2025; Oanh & Nguyen, 2025).

The Learning is not only about receiving information, it is about interacting, getting support and feedback. When students learn in person, teachers communicate with students by asking questions, explaining and adapting what they are teaching based on students' reactions. This human-like interaction means talking and listening to each other, in a natural, social, and beneficial way, just like humans do (Javaid et al., 2024). It includes talking, helping, listening, giving and receiving feedback and adapting to individual needs. Online learning systems still use mostly passive content such as videos, presentations and quizzes, which do not provide this human element and do not encourage active learning (Khosrawi-Rad et al., 2022; Krstić et al., 2022).

Conversational Agents are employed to help make online learning more interactive. Conversational Agents are computer programs that are able to communicate with humans in natural language (text or speech) (Khosrawi-Rad et al., 2022; Mariani et al., 2023). Conversational agents in education are used as tutors, teaching assistants, learning partners or coaches. They are able to respond to student questions, explain concepts, give feedback, remind students of their responsibilities, and motivate students (Bozkurt, 2023; Dai et al., 2024; Yusuf et al., 2025). Conversational agents have a few key components. One of those is inputs, which are what the system gets, such as learner inputs (text questions, voice commands, multiple-choice questions, clicks, time on task, and inactivity) and system data (exam scores, learning progress, and activity logs) (Akhuseyinoglu & Brusilovsky, 2022; Nazempour

& Darabi, 2023). Another part is the outputs, which are things that the system provides, like responses (text feedback and voice explanations), feedback (feedback, hint, and praise), guidance (learning plans, next-step recommendations), and motivation (reminders, badges, and positive messages) (Krstić et al., 2022; Rizwan et al., 2025; Younas et al., 2025). The third part is cues (signals for understanding the learner) which can be emotional (like frustration, confusion, or happiness from text or voice messages), cognitive (like incorrect answers, making repeated errors or answering slowly), or behavioral (like skipping lessons, inactivity or asking for lots of help) (Arguedas et al., 2024; Ortega-Ochoa et al., 2024; Zhai et al., 2024).

There are various types of conversational agents in education. Tutor agents teach and provide feedback on answers (Bozkurt, 2023). Peer agents are like study buddies who can help you discuss and practice what you learn (Mariani et al., 2023). Coach agents help set goals, track progress and motivate students (Yusuf et al., 2025). Some agents are rule-based (they follow a script and can choose from a decision tree with a set number of paths) and some are AI-based. Systems use GenAI and LLMs to generate explanations, questions, examples and feedback in a human-like manner (Al-Abri, 2025; Bozkurt, 2023). Conversational Agents are particularly useful in virtual tutoring, in which learners need help at any time and place. They provide instant support (fast response and step-by-step guidance), adaptive support (adjusting difficulty according to the learner's performance) and social support (empathetic phrases and encouragement) (Arguedas et al., 2024; Dai et al., 2024; Ortega-Ochoa et al., 2024). This makes them ideal for independent learning, online learning and large online classes where human tutors are not available to support all learners (Întorsureanu et al., 2025; Younas et al., 2025).

Conversational agents have several limitations and challenges. Conversational agents do not always have long-term memory, which hinders personalization (Al-Abri, 2025; Kusal et al., 2022). Understanding of emotions tends to be basic and keyword-driven (Javaid et al., 2024). Most systems are primarily text-based and lack multimodal interaction and expression such as tone of voice, facial expressions, and gestures (Ortega-Ochoa et al., 2024). Privacy concerns of personal data, bias in AI, transparency and over-reliance on AI are

also ethical concerns (Akpan et al., 2025; Al-Abri, 2025; Zhai et al., 2024). For these reasons, it's necessary to review the research in this area. This research reviews an emphasis on inputs and outputs and AI approach, GenAI in conversational agents. It clarifies four research questions related to the inputs and the outputs for low-engagement learners, how conversational agents interact with the users to provide real-time support and engagement, GenAI integration and evaluation.

The paper is structured as follows: Section 1.1 discusses the study background. Section 1.2 describes the research questions. Section 2 outlines the experiment. Section 3 listed the Result and Discussion. Section 4 notes the limitations and provides challenge gap analysis. Lastly, the paper is concluded in Section 5 with the findings and future work.

1.1 Basic Literature Study

The evolution of technology in education has come a long way from computer-based training to online learning. The earliest e-learning platforms were mostly concerned with the informative needs of learners with limited interaction between the system and the learner (Khosrawi-Rad et al., 2022). This tended to focus more on informative rather than interactive, supportive and emotional aspects. This resulted in low interaction and high drop-out rates for many of the students, especially with purely online and MOOCs (Întorsureanu et al., 2025; Rizwan et al., 2025).

To increase interaction, researchers created Intelligent Tutoring Systems (ITS). They tried to emulate some aspects of human tutors, such as explanations, questions and feedback (Dai et al., 2024). However, these first ITS systems used the expert-constructed, rule-based knowledge. They were inflexible and difficult to adapt. They were only able to answer limited questions and had difficulties with natural language input (Khosrawi-Rad et al., 2022). This led to the development of conversational interfaces, allowing students to interact using natural language (Mariani et al., 2023). With the advancements in ML and Natural Language Processing (NLP), conversational agents improved. These models, such as DL and transformer models, enabled them to understand sentences and generate human language (Al-Abri, 2025). This created more human-like chatbots that could explain, ask questions and have longer

conversations (Bozkurt, 2023; García-Méndez et al., 2025).

Educational conversational agents began as simple chatbots that used pre-scripted responses and keywords matching (Khosrawi-Rad et al., 2022). They could answer common queries or guide the user through a learning curriculum. Such chatbots provided basic support, but were not flexible, intelligent or personalized. The Learners were dissatisfied if the system didn't comprehend their questions or give them useful responses (Krstić et al., 2022). At the same time, learning analytics and educational data mining started to log and analyze the learner data - clicks, time-on-task, quizzes, interactions and so on (Akhuseyinoglu & Brusilovsky, 2022; Nazempour & Darabi, 2023). This allowed the system to learn about the learner and to provide personalized content. The Adaptive learning systems began to adapt the difficulty, ordering and feedback of the content based on learner's performance and interaction (Sayed et al., 2023; Yousef et al., 2025). And conversational agent systems recommend learning resources based on the learners' low-performing areas and difficulty levels to ease learning (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Gao, 2025; Younas et al., 2025). Based on learner interactions, these systems reinforce past learning content or offer new content according to the learner's progress (Arguedas et al., 2024; Rizwan et al., 2025; Yildirim-Erbasli et al., 2025). Agents can also identify boredom via language use and interactions and provide motivation to help the learner (Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025; Sajja et al., 2025).

These programs monitor learners' academic performance (e.g., quizzes, assignments, course progress) and behavioral patterns (e.g., motivation, frequency of interactions, time of participation, communication, and engagement) (Hu et al., 2024; Reddig et al., 2025; Yousef et al., 2025). Furthermore, emotional states (learner activities, questions, frustration, and engagement) are also tracked to adapt learning (Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025). Furthermore, question-answer similarity is assessed to measure concept comprehension and storytelling explanations are created to assist learners with low levels of engagement (Javaid et al., 2024; Rizwan et al., 2025; Yusuf et al., 2025). Finally, teacher dashboards are used to assist teachers monitor learners' progress and motivate them (Reddig et al.,

2025; Yildirim-Erbasli et al., 2025), while platform content generators assist teachers in providing personalized learning materials (Casheekar et al., 2024; Choi & Kim, 2025; Love et al., 2025).

In recent years, GenAI has had a major impact on conversational agents. Now, models like LLM can generate explanations, examples, feedback and even assessments in natural language (Al-Abri, 2025; Bozkurt, 2023; Prathigadapa & Daud, 2025). These agents help with virtual tutoring by providing immediate assistance (immediate feedback, step-wise guidance) and conversational feedback for a wide range of topics (Dai et al., 2024; García-Méndez et al., 2025; Naveed et al., 2025). Other research explores multi-agent or agentic AI for scaling feedback, assessment, motivation and content generation (Alfehaid & Hammami, 2023; Cai et al., 2025).

The next step is that many organizations are adopting affective computing. Research shows that emotions such as confusion, boredom and frustration are critical in learning (Javaid et al., 2024). The Emotion-aware conversational agents detect emotional cues (like vocal or text cues of emotions), cognitive cues (like incorrect answers and response time), and behavioral cues (like lack of activity and skipping) to understand what students are learning. They then respond with kindness (kind words), encouragement (motivating words), or changes to their teaching methods (Arguedas et al., 2024; Ortega-Ochoa et al., 2024; Zhai et al., 2024). This helps students feel more connected and engaged with online learning.

The previous literature reflects a clear shift from static e-learning to intelligent, conversational and generative systems. It has evolved from content delivery to engagement, personalization and emotional support. This work builds on this literature by looking at the design of conversational

agents, inputs (data given to the system), outputs (data provided by the system), and AI techniques; how conversational agents engage with the users to provide real-time support and engagement, use of GenAI and the evaluation of their effectiveness in education.

1.2 Research Questions

The review of online learning, conversational agents, virtual tutoring, and GenAI leads to the following research questions:

- RQ1:** What are the key inputs, outputs, and AI-based methodologies in implementing conversational agents to support learners with lower engagement?
- RQ2:** In what are the ways conversational agents interact with the users to provide real-time support and engagement?
- RQ3:** How is the Generative AI incorporated into the interaction mechanisms of conversational agents?
- RQ4:** What are the key metrics for evaluating the performance of conversational agents?

The review explores these questions to better understand the trends, challenges and opportunities in designing effective virtual tutoring systems powered by AI for online learning.

2. Methodology

This systematic review seeks to identify, categorize and assess the main inputs, outputs, AI Methodology, interaction techniques, GenAI technologies, and evaluation performance of the latest conversational agents for education. The studies reviewed are mostly from the years 2022 to 2025, due to the rapidly increasing research in conversational agents and GenAI for education.

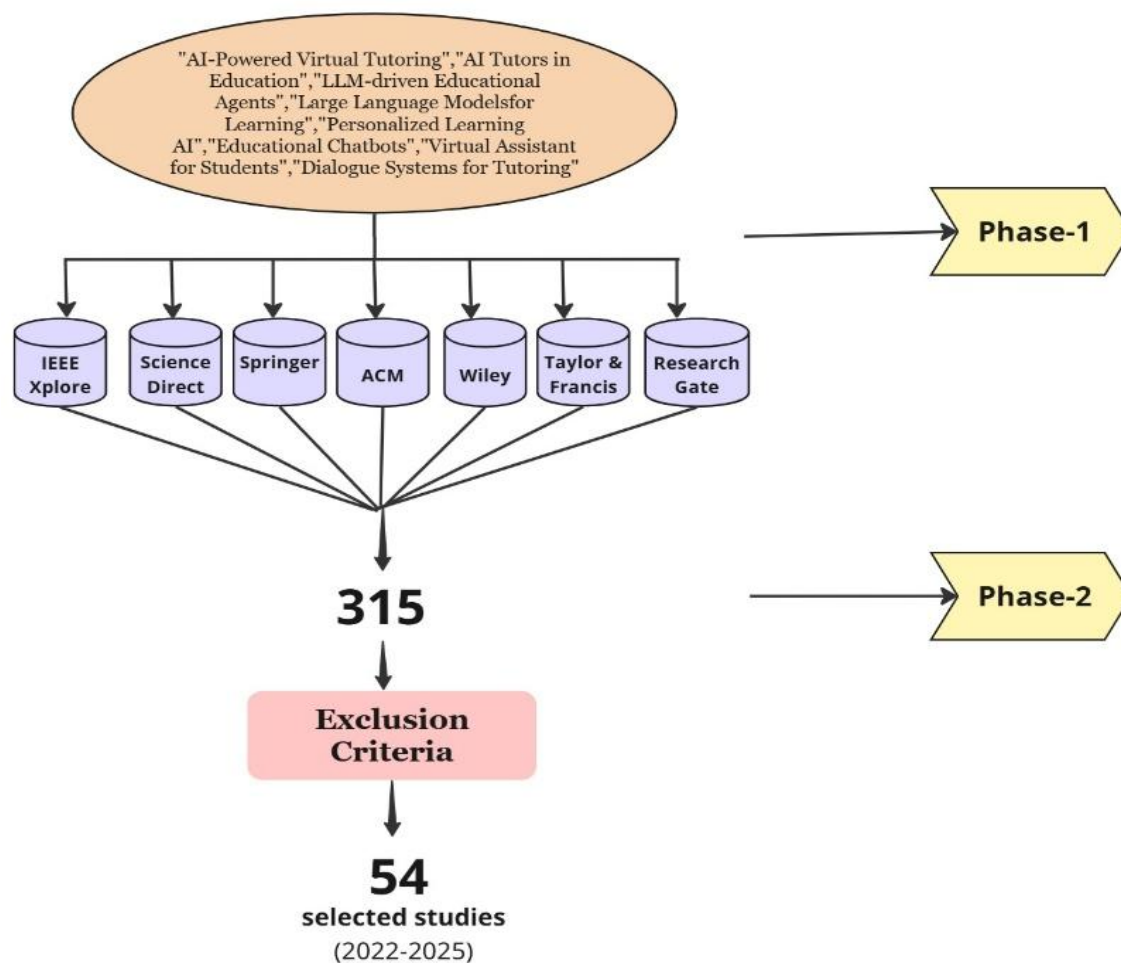


Figure 1. Process of this Study

2.1 Systematic Filtering Process

To conduct a thorough and rigorous review of the use of AI conversational agents in education, a systematic filtering process was used, as shown in Figure 1. involving 2 Phases are:

- Search Strategy and Data Sources
- Screening and Filtering

2.1.1 Phase-1: Search Strategy and Data Sources

During the identification phase, various databases, such as IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, Wiley, Taylor & Francis and ResearchGate, were searched. The following search related terms and keywords were used to search across these databases and identified 315 records.

“AI-powered Chatbots” / “Virtual Agents”, “Pedagogical Conversational Agents”, “Dialogue

Systems” / “Conversational Partners”, “Large Language Models (LLMs) in Education”, “Generative AI-powered Agents”, “GPT-3 / ChatGPT as Virtual Tutor”, “Adaptive and Personalized Tutoring”, “Learner Modelling” / “Individual Differences”, “Self-Directed Learning Support”, “Automated Feedback Generation”, “Formative Feedback Systems”, “AI-based Performance Prediction”, “AI Teaching Assistants”, “Real-time Classroom Behaviour Analysis”, “Deliberate Teaching Practice Frameworks”, “Virtual Learning Environments (VLE)”, “Intelligent Tutoring Systems (ITS)”, “Open Distance e-Learning (ODEL)”, “Over-reliance on AI Dialogue Systems”, “Student

Engagement & Motivation”,
“Human-AI Interaction (HAI)”

2.1.2 Phase-2: Screening and Filtering

In Phase 2 of the filtering funnel, Figure 1, papers were selected based on the exclusion criteria. Table 1 Provides the exclusion criteria which included non-English language articles, duplicate studies and those with no full text. Papers discussing general AI concepts and not directly related to education or student engagement were also excluded.

Table 2 shows the extracted key words from the selected studies from 2022 to 2025, based on the categorization of the concepts introduced in the previous studies (Akpan et al., 2025; Al-Abri, 2025;

Alfehaid & Hammami, 2023; Casheekar et al., 2024; Întorsureanu et al., 2025; Khosrawi-Rad et al., 2022; Kusal et al., 2022; Mariani et al., 2023; Rodrigues et al., 2022; Younas et al., 2025). These papers classified essential AI technologies in education, such as conversational agents, GenAI, adaptive learning, learner engagement and ITS. The clustered key terms in Table 2 correspond with typical research areas and conceptual themes described in several systematic reviews and technology systems (Akpan et al., 2025; Al-Abri, 2025; Bozkurt, 2023; Kusal et al., 2022; Mariani et al., 2023). The grouping reflects the main research trends and sub-trends, and enables a structured view of the evolution of conversational agents and GenAI technologies in digital learning.

Table 1. Exclusion Criteria

Exclusion Criteria
Non-English papers
Non-peer-reviewed sources (blogs, reports, white papers, theses without review)
Papers not related to AI in education or conversational agents
Papers with insufficient data, unclear methodology, or incomplete results
Research not addressing learner interaction, engagement, or tutoring aspects
Studies limited to general AI applications outside education
Articles without full-text access
Studies lacking evaluation or validation of results
Research focusing only on theoretical concepts without practical relevance

Table 2. Synthesis of AI in Education Key Terms (2022–2025)

Category	Key Term	Primary References
Conversational Agents (CA)	“AI-powered Chatbots” / “Virtual Agents”	(Akpan et al., 2025; Alfehaid & Hammami, 2023; Baabdullah et al., 2022; Belda-Medina & Calvo-Ferrer, 2022; Casheekar et al., 2024; Yildirim-Erbasli et al., 2025)
	“Pedagogical Conversational Agents”	(Arguedas et al., 2024; Ortega-Ochoa et al., 2024; Yusuf et al., 2025)
	“Dialogue Systems” / “Conversational Partners”	(Albornoz-De Luise et al., 2023; Love et al., 2025; Zhai et al., 2024)
GenAI & LLMs	“Large Language Models (LLMs) in Education”	(Al-Abri, 2025; Cai et al., 2025; Parker et al., 2025; Reddig et al., 2025)
	“Generative AI-powered Agents”	(Aperstein et al., 2025; Bozkurt, 2023; Întorsureanu et al., 2025; Milana et al., 2024)
	“GPT-3 / ChatGPT as Virtual Tutor”	(Bozkurt, 2023; Mariani et al., 2023; Prathigadapa & Daud, 2025)

Category	Key Term	Primary References
Personalization & Adaptation	“Adaptive and Personalized Tutoring”	(Champaneria, 2025; Sayed et al., 2023; Singh et al., 2022; Wah, 2025)
	“Learner Modelling” / “Individual Differences”	(Akhuseyinoglu & Brusilovsky, 2022; Javaid et al., 2024; Nazempour & Darabi, 2023; Sajja et al., 2025)
	“Self-Directed Learning Support”	(Younas et al., 2025)
Assessment & Feedback	“Automated Feedback Generation”	(Estévez-Ayres et al., 2025; Yousef et al., 2025)
	“Formative Feedback Systems”	(Alrajhi, 2024)
	“AI-based Performance Prediction”	(Gao, 2025; Rizwan et al., 2025; Wang & Luo, 2024)
Teaching & Classroom Support	“AI Teaching Assistants”	(García-Méndez et al., 2025; Rodrigues et al., 2022; Teng et al., 2025)
	“Real-time Classroom Behaviour Analysis”	(Hu et al., 2024; Krstić et al., 2022)
	“Deliberate Teaching Practice Frameworks”	(Kusal et al., 2022)
Learning Environments	“Virtual Learning Environments (VLE)”	(Bandi et al., 2025)
	“Intelligent Tutoring Systems (ITS)”	(Khosrawi-Rad et al., 2022)
	“Open Distance e-Learning (ODEL)”	(Maphalala et al., 2025)
Ethical & Cognitive Impact	“Over-reliance on AI Dialogue Systems”	(Yusuf et al., 2025)
	“Student Engagement & Motivation”	(Choi & Kim, 2025; Dai et al., 2024; Javaid et al., 2024; Oanh & Nguyen, 2025; Yin et al., 2024)
	“Human-AI Interaction (HAI)”	(Kim et al., 2025)

2.2 Distribution of Selected Studies

Following the full-text assessment, 54 primary studies were chosen for the qualitative synthesis and data extraction 338-367. The studies are from high-quality journals with the majority published in

Q1, followed by Q2 and Q3, and only a few in Q4 using SCImago indexing. A few studies are from non-indexed publications, such as conferences and arXiv preprints. Figure 2 shows the number of selected papers published per year.

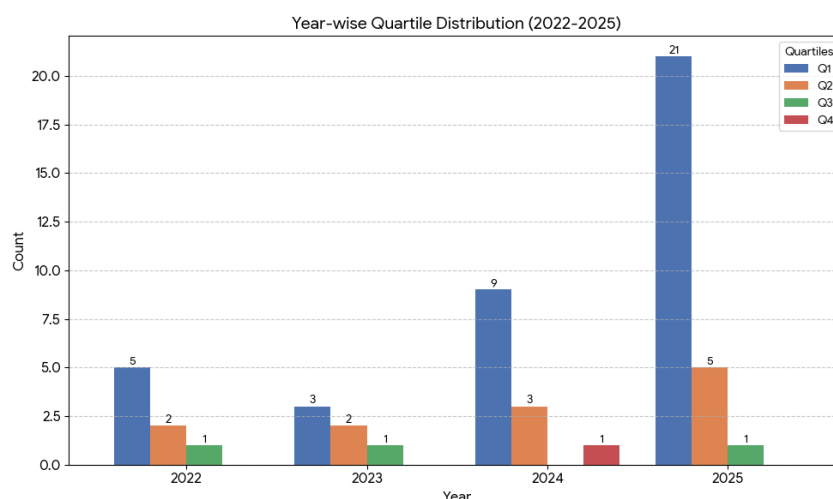


Figure 2. Distribution of Selected Articles by Year

The graph shows a significant increase in publications since 2020, and a steep rise from 2023 to 2025. This recent growth indicates a growing trend in the interest in intelligent, adaptive and generative conversational systems in education.

The analysis of the reviewed studies based on their method is presented in Table 3. The findings reveal that the majority of studies adopt quantitative methods (24 studies) that mostly use experimental

designs and surveys to assess the effectiveness and efficiency of AI educational systems. By comparison, 17 studies adopt a qualitative approach, using interview, case studies and focus groups techniques to investigate the educational implications, interaction quality and user experience. There are also 13 mixed-method studies, which use both quantitative and qualitative methods to explore the benefits and challenges of using conversational agents in education.

Table 3. Distribution of Selected Studies by Research Design and Methodology

Research Design	Number of Articles	Research Method	No. of Research Studies
Qualitative	17	Interviews	6
		Case Studies	5
		Focus Groups	2
		Other	4
Quantitative	24	Experiments	13
		Surveys	8
		Meta-analyses	1
		Text Analysis	1
		Other	1
Mixed Methods	13	Combined Approaches	10
		Meta-analyses	1
		Text Analysis	1
		Other	1
Total	54		54

To better understand the development and impact of conversational agents in education, recent research has increasingly focused on determining the key components, interaction patterns, GenAI technologies and evaluation methods that constitute these systems that are explored in RQ1, RQ2, RQ3 and RQ4, respectively.

3. Results

Based on the analysis of selected studies, research questions were refined. Literature shows clear trends that answer the research questions, and describe how conversational agents support learning and engagement in real-time.

3.1 RQ1: What are the key inputs, outputs, and AI-based methodologies in implementing conversational agents to support learners with lower engagement?

Conversational agents rely on a variety of inputs, outputs, and smart AI techniques to enhance

learners' engagement in online learning. These three components interact to understand, support and engage learners. Some studies highlight the need to access different inputs including learner interactions, logs, emotions, and performance data to support learners with low engagement, and outputs like adaptive feedback, motivational cues, and learning pathways (Akhuseyinoglu & Brusilovsky, 2022; Nazempour & Darabi, 2023; Sajja et al., 2025; Yousef et al., 2025). These papers describe how AI technologies such as ML, deep learning (DL), and LLM-based systems interact to process and transform learner data into pedagogical interventions (Al-Abri, 2025; Reddig et al., 2025; Yousef et al., 2025).

3.1.1 Inputs Used in Conversational Agents

The figure Figure 3 shown that input data for conversational agent.

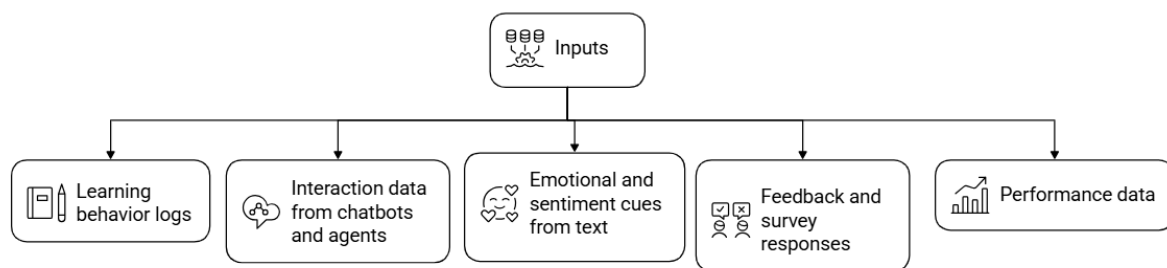


Figure 3. Types of Input Data used in Conversational Agent Systems

3.1.1.1 Learning Behaviour Logs

Initially, when disengagement of learners is detected through patterns, these patterns are related to clicks, time-on-task, and inactivity. Disengagement is detected if a learner is not clicking, not spending time on tasks, or is inactive. Several studies (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Sajja et al., 2025; Yousef et al., 2025) examined learner interaction logs, looking at activity levels to identify disengaged learners. This is a key input, such as clicks, time-on-task and patterns of inactivity, as it reflects learner engagement. Learning behavior logs are commonly used as an input to detect learner engagement levels in conversational agent systems. For example, Akhuseyinoglu and Brusilovsky (2022) examined learners' interaction patterns by capturing individual differences in learning using interaction logs and showed that learners' activity

patterns such as time-on-task and frequency of navigation can be used to predict their engagement levels. Likewise, Al-Abri (2025) described how the use of large language models can benefit from interaction data collected from previous sessions to personalize responses and track user engagement throughout the interaction.

3.1.1.2 Interaction Data from Chatbots and Agents

If chatbots are used to collect learners' interaction data, such as if learners send fewer messages, ask fewer questions, or stop responding, these chat messages, questions, and replies reflect learner disengagement. Researchers (Albornoz-De Luise et al., 2023; Bozkurt, 2023; García-Méndez et al., 2025; Parker et al., 2025; Yusuf et al., 2025) examined conversation logs to understand difficulties and interaction data. This input relates to engagement in communication, rather than just

actions. Learning conversation data from chatbots is also a significant input to detect learner engagement. For example, Albornoz-De Luise et al. (2023) created conversational models for intelligent tutoring systems that captured learner-agent interactions to enable natural language dialogues and understand where learners needed more guidance. Their research showed that by examining the conversation flow, learner confusion and disengagement can be detected. Likewise, Bozkurt (2023) discussed GenAI based conversational educational agents and the importance of dialogue logs to inform learner needs and ensure ongoing interaction between the learner and the agent. The research highlighted that high learner responses and questions signify high engagement, whereas low learner responses signify low engagement.

3.1.1.3 Text Emotion and Sentiment Analysis

Learner disengagement can be caused by boredom or frustration. Emotions are identified from learner messages using sentiment analysis. These emotions are used by conversational agents to engage and prevent disengagement (Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025). Emotions indicate hidden disengagement. Another source of input for disengagement detection is emotional and sentiment cues from learner utterances. Several studies have looked at the effects of emotions such as frustration, boredom, confusion and motivation on learner engagement. For instance, Krstić et al. (2022) reviewed the literature on AI in education and identified the use of affective computing techniques to recognize learner emotions while interacting with AI systems (Naveed et al., 2025). They noted that emotion recognition is valuable to detect subtle disengagement that is not evident from the learner's behaviour. Similarly, Kim et al. (2025) examined human-AI interactions in education and noted the importance of capturing learner emotions to improve learning engagement and learner satisfaction. In addition, Parker et al. (2025) used methods based on large language models to analyze learner responses and feedback, and demonstrated that sentiment detection approaches can detect learners' dissatisfaction or confusion. They found their approach enabled conversational agents to detect patterns of emotions to motivate learners and prevent disengagement.

3.1.1.4 Survey and Feedback

Learner-reported experiences, learning challenges or satisfaction levels are used by conversational agents to adjust teaching strategies and engagement strategies (Javaid et al., 2024; Nazempour & Darabi, 2023; Prathigadapa & Daud, 2025). Analysis of feedback forms and surveys is often used to increase system responsiveness. Survey and feedback responses provide direct and learner-reported input to conversational agents, helping them to identify user experiences, learning difficulties and satisfaction levels. For example, Nazempour and Darabi (2023) examined personalized learning environments through analyses of student feedback and responses to determine user preferences and problems. Research by Javaid et al. (2024) examined computational models associated with emotional and behavioural development and noted the value of user feedback to learn about student reactions and learning outcomes. They found that feedback-based knowledge helps to guide adaptive responses and detect learners who need extra support. Furthermore, Prathigadapa and Daud (2025) surveyed virtual tutoring systems that used learner feedback and survey responses to assess system effectiveness and enhance system interactions. This research highlighted the importance of collecting ongoing user feedback to allow conversational agents to improve their responses and increase user satisfaction.

3.1.1.5 Performance Analytics

Low learner engagement can be due to low learner performance. Performance analytics are used to detect poor learners. Learner performance data (such as quizzes, assignments and assessments) is a significant input used to identify learners who may have low engagement because of low performance. A number of studies have used performance analytics to detect at-risk learners and offer learning support. For example, Hu et al. (2024) created an AI-enhanced classroom behavior analytics system that tracked learner performance and activity to identify learners needing extended classroom support. Their research showed that using performance outcomes and behavioral features allows the identification of learners with deteriorating academic performance and participation. Similarly, Reddig et al. (2025) developed large language model-based intelligent tutoring systems, in which learner performance data was used to provide tailored feedback and

advice. This study demonstrated that conversational agents can deliver context-aware explanations and personalized feedback by examining performance data. Similarly, Yousef et al. (2025) proposed a framework for AI-based grading and feedback systems in programming, where student performance data was used to

provide automated feedback and determine areas where students were deficient.

3.1.2 Outputs Generated by Conversational Agents

The figure Figure 4 shown that key components of output from AI-based personalized learning.

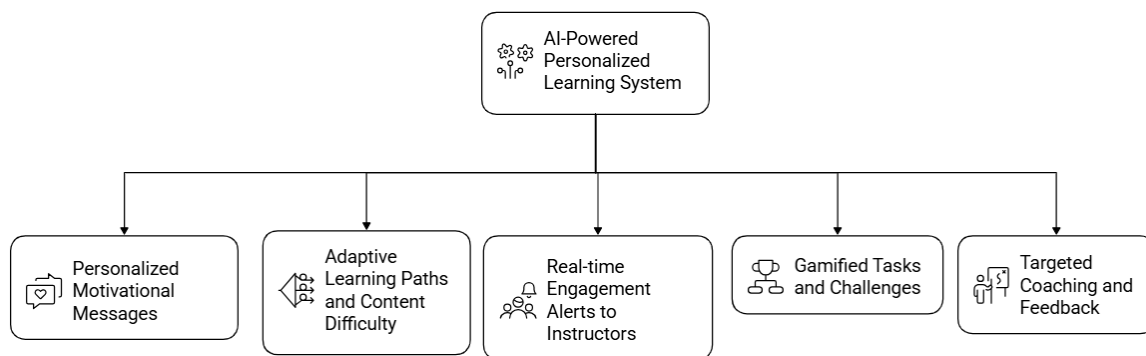


Figure 4. Types of Output Generated by AI-Powered Personalized Learning Systems

3.1.2.1 Personalized Motivational Messages

Conversational agents often use motivation messages to encourage the learner when he or she displays signs of low engagement and inactivity. There have been studies that have demonstrated how AI systems automatically generate motivational and encouragement messages when learners are showing signs of disengagement, frustration or a decrease in performance (Naveed et al., 2025). For instance, Krstić et al. (2022) reviewed research on AI in education, and found intelligent systems use motivational messaging strategies to provide emotional support for learners and to motivate them during learning tasks. They observed that encouragement and motivational messages help maintain learner engagement, particularly when learners encounter difficulties. Similarly, Kim et al. (2025) investigated AI enhanced learning environments and how personalized responses and motivational feedback affect learners' satisfaction and engagement. They discovered that AI systems that offer personalized encouragement and support responses to learners can increase motivation and keep learners engaged. Additionally, Parker et al. (2025) developed LLM based systems to respond to learner interactions and give feedback. They found that the motivational responses they automatically generate based on learner sentiment and interactions help maintain learner engagement and confidence.

3.1.2.2 Adaptive Learning Paths and Content Selection

Learning paths and difficulty level adjustments are generated outputs that dynamically adjust the learning path according to learner performance, helping learners who are falling behind stay engaged. The system analyses performance, behavior and engagement. The system dynamically alters the learning sequence and the altered learning path is the output. Numerous studies have used adaptive tutoring systems that automatically modify the difficulty of instruction and learning sequences (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Arguedas et al., 2024; Gao, 2025; Younas et al., 2025). Personalized learning sequences and content difficulty are key outputs of conversational agents to adapt learning paths according to learner's performance. Other researchers have implemented systems capable of changing learning content for learners who display disengagement or other difficulties. For example, Akhuseyinoglu and Brusilovsky (2022) studied learner behavior patterns to model learner differences and showed that adaptive systems can modify learning content based on learner interactions. They found that adapting the difficulty of the content to learner behavior enhances engagement and continuity of learning. Likewise, Arguedas et al. (2024) studied pedagogical agents that offered adaptive guidance and cognitive assistance in response to learners' behavior. They found that adaptive guidance and

personalized learning path support learners' learning performance and learning motivation.

3.1.2.3 Real-Time Engagement Alerts to Instructors

Instructor engagement alerts are automatic notifications to educators that learners are disengaged, and to provide timely support. A number of studies have adopted early warning systems to track learner engagement and trigger alerts when engagement is low. For instance, Suresh Babu and Dhakshina Moorthy (2024) investigated the use of AI-driven gamified learning platforms and suggested the importance of monitoring learner engagement to provide alerts when learners' engagement or motivation levels are low. They emphasized the value of alerts to educators to prevent disengagement. Similarly, Reddig et al. (2025) developed intelligent tutoring systems with large language models, which gather interaction and performance data to give real-time notifications and personalized teaching recommendations. They demonstrated that automatic notifications and suggestions based on performance trends enable instructors to reach out to struggling students. Similarly, Yildirim-Erbasli et al. (2025) investigated the use of conversational agents for higher education assessment systems, and showed that systems that monitor learner progress are able to notify instructors if learners are not engaged enough. Similarly, Yin et al. (2024) presented a longitudinal study on chatbot-based formative feedback systems and demonstrated that systems that track learner interactions can provide notifications and alerts about feedback to learners who need more instruction. The systems notify educators of disengagement or at-risk learners, allowing educators to intervene, provide personalized instruction and reduce retention.

3.1.2.4 Gamified Tasks and Challenges

Gamified tasks and challenges are key outputs from conversational agents to increase learner engagement through gamification strategies such as badging, levels, points and achievement tasks. These are aimed at maintaining learner engagement and avoiding disengagement through interactive and fun learning experiences. A number of studies have shown the success of gamification approaches in enhancing learner engagement. For example, Yildirim-Erbasli et al. (2025) examined the application of conversational agents in higher education assessment environments and found

that gamified feedback features (such as achievement levels and rewards) boosted learner motivation and engagement. They found that reward-based mechanisms embedded in AI based learning systems motivates learners to persist and engage with learning systems. Similarly, Younas et al. (2025) reviewed recent literature on AI based self-directed learning systems and found gamification is a popular approach to engagement. These systems, through the use of reward-driven activities and goal setting, support in maintaining interest, increasing engagement and helping learners at risk of disengagement.

3.1.2.5 Targeted Coaching and Feedback

Personalized coaching and feedback are critical outcomes of conversational agents that deliver personalized explanations, corrections and improvement strategies to support learners in overcoming learning obstacles and keeping them engaged. A number of researchers have developed automated feedback systems to interpret learner responses and provide personalized instruction. The paper by Javaid et al. (2024) investigated computational models that analyze learner responses and emotional cues and provide corrective and supportive feedback, showing that individualized support enhances learner comprehension and emotion. Their research highlighted the role of adaptive feedback in sustaining learner motivation. Likewise, García-Méndez et al. (2025) conducted a review of the use of large language models as tutors and found these systems provide real-time explanations, hints and step-by-step guidance to questions posed by learners. They found that automated feedback generated by natural language processing enhances learning, improves understanding, and promotes self-learning. Similarly, Rizwan et al. (2025) reviewed the prediction of student engagement and found that personalized feedback is an effective intervention to support students that were identified as disengaged. Also, Yusuf et al. (2025) proposed a framework for pedagogical conversational agents and reported that adaptive coaching models provide formative feedback (for improvement) and corrective feedback (for identifying learning errors and providing alternatives). In addition, Sayed et al. (2023) designed adaptive content presentation systems that adapt to learner success and failure and generate bespoke exercise tasks and navigation support, and reported that automated coaching

systems enhance learner engagement and understanding.

3.1.3 AI-Based Methodologies Used in Conversational Agents

This section outlines the key AI technologies, algorithms and models used in conversational agents to support learners with low engagement. The studies examined demonstrate an evolution, as shown in Figure 5, from rule-based dialogue systems to ML, DL, large language models, GenAI and multi-agent systems. These algorithms and models play a role in learner behavior prediction, response generation, personalization, emotional analysis and generation of feedback. Dialogue systems initially used rule-based methods in which scripts and pattern matching on keywords were used to generate responses. In studies such as (Bandi et al., 2025; Yin et al., 2024), it was found that early dialogue systems relied on pattern matching and pre-determined responses, which were not adaptive to the context and complex queries. These systems were effective for tasks with well-defined structure but not for complex learner questions and personalized.

In recent years, conversational systems have been powered by large language models (LLMs), providing human-like language understanding and response generation. Several studies (Alfehaid & Hammami, 2023; Belda-Medina & Calvo-Ferrer, 2022; Casheekar et al., 2024; Maphalala et al., 2025; Mariani et al., 2023) documented the development of conversational tutors powered by LLMs that interpreted learner queries, provided explanations and engaged in conversational exchanges. The systems provided better responses and supported contextual interactions in teaching and learning settings. The advent of GenAI also added to the potential of conversational tutors by allowing them to generate content, provide feedback and offer tutoring support. Research by Gao (2025), Al-Abri (2025), and Younas et al. (2025) demonstrated that GenAI approaches enable the generation of explanations, examples and adaptive learning content. These approaches enable real-time tutoring support and enhance engagement through personalized learning. In more sophisticated applications, multi-agent AI systems have been developed to orchestrate learning environments using multiple AI agents. Research by Nazempour and Darabi (2023) developed multi-agent systems with monitoring agents, tutoring agents and feedback agents that collaboratively assist learners.

3.1.3.1 Rule-Based Algorithms and Dialogue Models

Early conversational agents were implemented using rule-based algorithms that relied on predefined decision logic and structured dialogue control mechanisms. The study by Yin et al. (2024) implemented a pattern-matching algorithm using keyword-based Natural Language Processing (NLP) techniques. The study by Bandi et al. (2025) developed an early conversational agent using finite-state machine (FSM) dialogue models and if-then rule-based algorithms. The study by Teng et al. (2025) proposed an early-stage conversational agent architecture using script-based dialogue models and menu-driven interaction. These studies demonstrate that early conversational agents primarily relied on pattern-matching algorithms, if-then rule-based decision trees, finite-state machine (FSM) models, and script-based dialogue models to manage structured learner interactions.

3.1.3.2 Machine Learning for Engagement Prediction

Early detection of learners with low engagement is critical for offering early support. ML is commonly used to predict learners' engagement and disengagement, dropout risk, and inactivity patterns through interactions and performance data. The study by Sajja et al. (2025) developed an AI-based learning analytics model that used supervised ML algorithms such as Decision Tree, Random Forest, and Support Vector Machine (SVM) model to process student interaction data. The study by Yousef et al. (2025) designed a ML-driven assessment and feedback system with learning analytics to track student learning performance. The study used classification models such as Random Forest models and gradient-based predictive models. These studies show that prediction of engagement using ML mainly relies on supervised classification algorithms to process behavioral and learning performance data. These models allow the identification of at-risk learners and proactive responses to support engagement, such as automatic feedback, warnings, and recommendations for adaptive learning, to improve engagement and prevent learner dropouts in conversational learning systems.

3.1.3.3 Deep Learning for Behaviour and Pattern Analysis

DL is extensively used to study complex learner interaction patterns and to inform personalized

interactions in conversational agents. These approaches are well suited to processing data containing temporal and sequential information about the learner, allowing systems to learn hidden patterns of learner engagement and learning behaviors. The studies by Akhuseyinoglu and Brusilovsky (2022) and Stein and Pollak (2024) used data-driven learner behavior modeling approaches using Artificial Neural Network (ANN) models to process learner interaction data and detect individual learning differences. The study by Suresh Babu and Dhakshina Moorthy (2024) explored the application of deep learning neural network models in gamified learning systems to detect learning engagement patterns. The study adopted Deep Neural Network (DNN) models to analyze user interaction data including task completion, reward engagement and participation rate. The system used the identified patterns to adapt gamification strategies such as reward and difficulty levels of activities. The study by Akpan et al. (2025) surveyed conversational and GenAI systems and noted the use of Recurrent Neural Network (RNNs) and Long Short-Term Memory (LSTM) models to process conversational interactions. The study noted increased adaptive conversational behavior using RNN and LSTM models because of their learning from ongoing user interaction sequences. The study by Sajja et al. (2025) deployed DL-based learning analytics models (using LSTM networks), to analyze temporal sequence of user behavioral data such as login time, time on task, inactivity and path. These studies show that DL models for behavior and pattern recognition by the models to process sequential learner data. These models allow conversational agents to identify implicit patterns, disengagement risk, and to provide tailored feedback, which supports adaptive learning strategies for disengaged learners.

3.1.3.4 Conversational Agents with LLMs

LLMs are widely used in contemporary conversational agents because they are the technology that enables such conversation. LLMs can comprehend questions (Mariani et al., 2023), produce responses and mimic tutoring conversations (Alfehaid & Hammami, 2023). They have been used in AI tutoring (Belda-Medina & Calvo-Ferrer, 2022), conversational feedback systems (Casheekar et al., 2024), and automated explanation systems (Maphalala et al., 2025). LLM powered conversational agents have emerged as a

key approach to implementing intelligent tutors and learner support tools. The study (Casheekar et al., 2024), which surveyed state-of-the-art chatbot technologies, including GenAI systems, found that automated conversational feedback systems are broadly implemented with large language models. The studies by Belda-Medina and Calvo-Ferrer (2022), Alfehaid and Hammami (2023), Mariani et al. (2023), Casheekar et al. (2024), and Maphalala et al. (2025) reported the use of transformer-based language models such as GPT-based systems, to generate conversational responses and explanations. These studies show LLM based conversational agents are a key AI method used in modern educational systems. Through the generation of intelligent conversation and automated explanations, continuous interaction with the learner, LLM help personalize tutoring and increase engagement for low-participation learners.

3.1.3.5 Multi-Agent AI Frameworks

Conversational agents can utilize multi-agent AI frameworks to support multiple intelligent units to work together to support learners in real-time. Rather than processing information on a single unit, the systems distribute tasks among various types of agents such as monitoring agents, tutoring agents and feedback agents, to provide scalable and distributed learning support. The research by Nazempour and Darabi (2023) implemented personalized learning support using a multi-agent system architecture, which involved multiple intelligent agents collaborating to interpret learner actions and provide personalized learning support. A feedback agent generated personalized feedback and suggestions based on learner behavior. The system employed models of collaborative agent communication to share information between agents and coordinate decision-making. These models enable conversational agents to multitask, to assess when a user is disengaging, and to offer personalized instruction, improving the efficiency and flexibility of AI-powered learning systems.

3.1.3.6 GenAI Models for Generating and Personalising Content

Conversational agents extensively use GenAI models for content generation and personalization. They can produce explanations, questions, feedback and adaptive content in response to learners' performance and interaction data. The studies by Gao (2025) and Stein and Pollak (2024)

developed DL-based generative models that were integrated with predictive learning analytics algorithms to assess teaching quality and provide adaptive learning materials. The model employed DNNs for learner performance analysis. The study by Al-Abri (2025) proposed a full implementation of Generative LLMs using Transformer models for conversational agents that can generate context-specific teaching responses. The study by Younas et al. (2025) created adaptive recommendation systems incorporating GenAI and Retrieval-Augmented Generation (RAG) models for self-directed learning. These studies show the use of

GenAI models for content generation and personalisation) models and adaptive recommendation algorithms.

In Figure 5, shows the evolution of AI methods from rule-based and ML-based methods to DL, large language models, and GenAI-based methods. The findings indicate a decrease in rule-based methodologies with a rapid increase in LLM and GenAI methodologies, especially since 2023. Table 4 presents the key inputs, outputs and AI methodologies for supporting low-engagement learners.

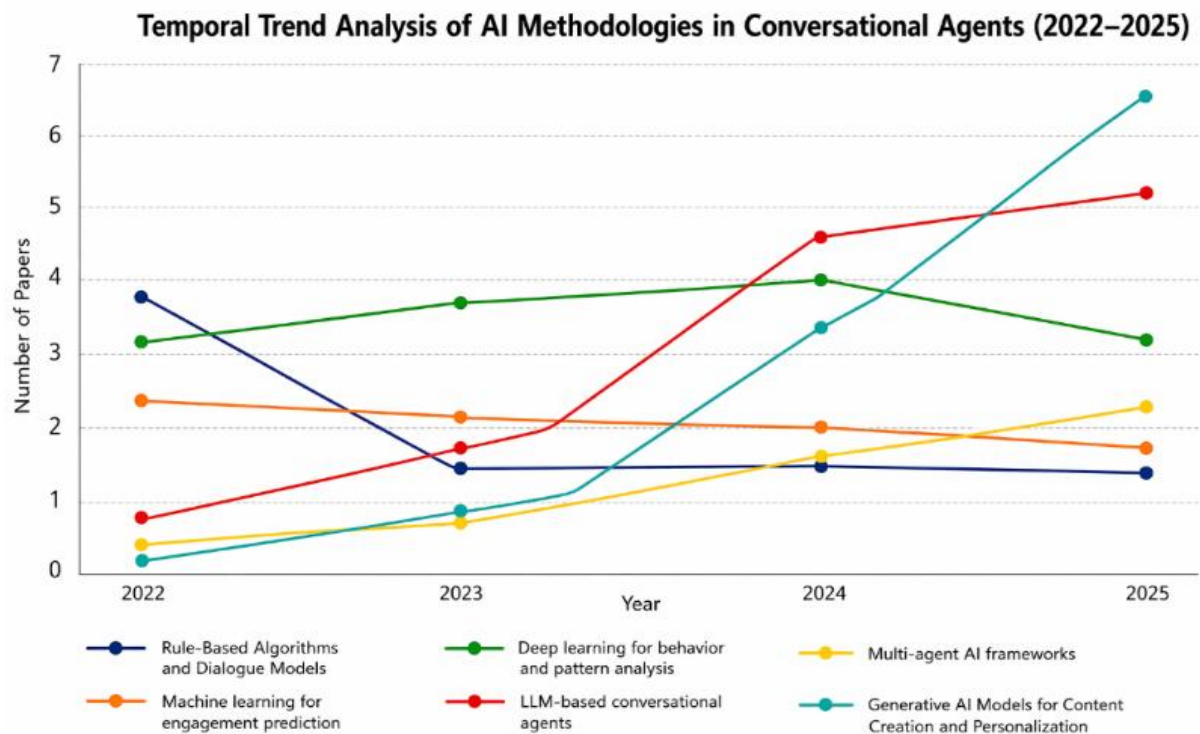


Figure 5. Temporal Trend Analysis of AI Methodologies in Conversational Agents (2022–2025)

Table 4. Key Inputs, Outputs, and AI Methodologies for Supporting Low-Engagement Learners

Category	Strategy	References
Inputs	Learning behaviour logs (clicks, time-on-task, inactivity patterns)	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Sajja et al., 2025; Yousef et al., 2025)
	Interaction data from chatbots and agents	(Albornoz-De Luise et al., 2023; Bozkurt, 2023; García-Méndez et al., 2025; Parker et al., 2025; Yusuf et al., 2025)
	Emotional and sentiment cues from text	(Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025)
	Feedback and survey responses	(Javaid et al., 2024; Nazempour & Darabi, 2023; Prathigadapa & Daud, 2025)
	Performance data (quizzes, assignments, assessments)	(Hu et al., 2024; Reddig et al., 2025; Yousef et al., 2025)

Category	Strategy	References
Outputs	Personalized motivational messages	(Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025)
	Adaptive learning paths and content difficulty	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Arguedas et al., 2024; Gao, 2025; Younas et al., 2025)
	Real-time engagement alerts to instructors	(Reddig et al., 2025; Suresh Babu & Dhakshina Moorthy, 2024; Yildirim-Erbasli et al., 2025; Yin et al., 2024)
	Gamified tasks and challenges	(Yildirim-Erbasli et al., 2025; Younas et al., 2025)
	Targeted coaching and feedback	(García-Méndez et al., 2025; Javaid et al., 2024; Rizwan et al., 2025; Sayed et al., 2023; Yusuf et al., 2025)
AI Methodologies	Rule-based algorithms and dialogue models	(Bandi et al., 2025; Teng et al., 2025; Yin et al., 2024)
	Machine learning for engagement prediction	(Sajja et al., 2025; Yousef et al., 2025)
	Deep learning for behaviour and pattern analysis	(Akhuseyinoglu & Brusilovsky, 2022; Akpan et al., 2025; Sajja et al., 2025; Suresh Babu & Dhakshina Moorthy, 2024)
	LLM-based conversational agents	(Alfehaid & Hammami, 2023; Belda-Medina & Calvo-Ferrer, 2022; Casheekar et al., 2024; Maphalala et al., 2025; Mariani et al., 2023)
	Multi-agent AI frameworks	(Nazempour & Darabi, 2023)
	Generative AI models for content creation and personalization	(Alfehaid & Hammami, 2023; Belda-Medina & Calvo-Ferrer, 2022; Maphalala et al., 2025; Mariani et al., 2023; Reddig et al., 2025; Yildirim-Erbasli et al., 2025)

In general, the studies reviewed in this article show that conversational agents depend on a mixture of behavioral, emotional, and performance inputs to produce adaptive outputs (feedback, guidance, and motivation). These outputs are driven by sophisticated AI-based approaches such as ML, DL, learning analytics and LLM. Thus, these results answer RQ1, which examines the key inputs, outputs, and AI-based methodologies for implementing conversational agents for supporting low-engagement learners in digital learning.

3.2 RQ2: In what are the ways conversational agents interact with the users to provide real-time support and engagement?

Conversational agents engage with users in several interactive and intelligent ways to provide real-time assistance and engagement. Firstly, they employ natural language interaction, enabling learners to interact with the system in a human-like way using text or voice, making it easier and more intuitive to learn (Albornoz-De Luise et al., 2023; Bozkurt,

2023; Yusuf et al., 2025). They also offer real-time feedback, allowing learners to receive immediate explanations, feedback and assistance during tasks, removing any lag and enhancing understanding (García-Méndez et al., 2025; Javaid et al., 2024; Stein & Pollak, 2024). Moreover, according to the literature, conversational agents create real-time interaction by employing multi-turn dialogue, real-time feedback, personalized interaction, emotion- or sentiment-based interaction, interactive prompts and reminders and multimodal interaction. These agents provide human-like interactions by sustaining the conversation flow, delivering real-time feedback, and responding to learners' input to improve engagement and learning outcomes (Albornoz-De Luise et al., 2023; Hu et al., 2024; Kusal et al., 2022; Love et al., 2025; Yusuf et al., 2025). Furthermore, conversational agents personalize interactions through the analysis of user actions, performance and preferences to adapt responses, the difficulty of tasks, and the learning path (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025). They also use emotion and

sentiment analysis to detect the user's mood (e.g., confusion, frustration) and provide feedback with encouraging messages to increase engagement (Krstić et al., 2022; Parker et al., 2025). Conversational agents also use interactive elements such as prompts, reminders and follow-up questions to engage students and avoid disengagement (Akpan et al., 2025; Albornoz-De Luise et al., 2023). Certain systems use multimodal communication through the integration of text, speech and visuals to enhance user experience. These systems also integrate with learning management systems and dashboards to provide real-time notifications, progress reports and instructor intervention, if required (Reddig et al., 2025; Yildirim-Erbaşlı et al., 2025). Real-time interactions with conversational agents enhance the development of cognitive abilities, directly responding to RQ2 by showing how conversational agents interact with users through dialogue, feedback, and support. RQ2 explores how conversational agents interact with learners, delivering real-time support and engagement. The interaction mechanisms provide not only instructional support but also strengthen cognitive abilities for successful learning.

Conversational agent interactions enhance three key cognitive skills: decision-making, critical thinking, and analytical thinking. By engaging in real-time conversations, generating feedback and adapting to learners' responses, conversational agents support learners to assess information, make decisions and draw conclusions. As such, the conversational agent interaction designs are linked with cognitive skill development and enhance learner engagement and learning outcomes. Conversational agents support decision-making skills when they provide real-time advice, hints and feedback to support learners in choosing suitable solutions or learning approaches. Critical-thinking skills are supported when agents provide follow-up questions, explanations and clarification questions that prompt learners to assess information and think critically. Critical-thinking abilities are enhanced through problem-solving strategies and performance feedback that help learners solve problems in a step-by-step manner, and understand cause and effect. Conversational agents also employ various conversational roles, known as multiple communication models, such as tutor (instructs), peer (learns with others), and coach (motivates) (Bandi et al., 2025; Love et al., 2025; Mariani et al., 2023; Milana et al., 2024). The agent

can teach like a teacher, practice like a peer, or motivate like a coach, as required by the learner. Depending on performance, agents can coach learners by evaluating performance data and developing individual learning plans and recommendations (Al-Abri, 2025; Hu et al., 2024; Krstić et al., 2022; Rizwan et al., 2025; Yousef et al., 2025). To enhance motivation, many conversational agents also use gamification elements such as points (scores for completed tasks), badges (rewards for achievements) and levels (progress stages) (Yildirim-Erbaşlı et al., 2025). These features enhance the interaction and engagement, encouraging learner participation in activities. Feedback is another important element of conversational interaction. Agents offer corrective feedback (feedback on what is wrong and why) and formative feedback (feedback on how to improve) to help learners reflect on their mistakes and improve their learning over time (Javaid et al., 2024; Krstić et al., 2022; Prathigadapa & Daud, 2025; Rizwan et al., 2025; Stein & Pollak, 2024).

Table 5 outlines how conversational agents employ various interaction mechanisms to support users and enhance interaction in online learning settings. These methods serve a unique purpose in enhancing the effectiveness, interactivity and personalization of interactions. These interaction mechanisms include natural language interaction, multi-turn conversation, real-time feedback, personalization, awareness of learner emotions and prompting strategies. A number of researchers have created conversational systems that support interaction with human-like conversation and learning support.

Multi-turn dialogue allows conversational agents to engage in multiple rounds of conversation. For example, Kusal et al. (2022), Milana et al. (2024), and Bandi et al. (2025) have developed conversational systems that have the capacity to remember learner input via contextual memory, enabling the agent to respond to earlier conversations. Likewise, Kim et al. (2025) and Love et al. (2025) have shown conversational systems that enabled long dialogue sessions, enhancing learning and dialogue quality. These systems demonstrated that multi-turn dialogue supports reasoning and provides continuous support.

Instant feedback systems offer immediate feedback during exercises. Research by Javaid et al.

(2024), García-Méndez et al. (2025), and Parker et al. (2025) developed feedback mechanisms that provided real-time explanations and explanations when learners made mistakes. Furthermore, Reddig et al. (2025), Yildirim-Erbasli et al. (2025), and Yousef et al. (2025) created AI-based models for feedback that detect incorrect answers and provide corrective strategies in real time. These systems showed increased task performance accuracy and learning rates due to timely instruction.

Tailored interaction provides learning responses based on learner progress. In Akhuseyinoglu and Brusilovsky (2022) and Al-Abri (2025), studies implemented adaptive learning systems that monitored learner progress to make decisions on course content and teaching strategies. Likewise, Cai et al. (2025), Gao (2025), and Rizwan et al. (2025) used learner modeling to provide personalized explanations and recommendations. These found that tailored responses enhance motivation and facilitate personal learning experiences.

Emotion and sentiment-based interaction is related to detecting learners' emotions and giving support. Research by Krstić et al. (2022) and Parker et al. (2025) applied sentiment analysis to identify emotions like frustration or confusion in learning interactions. Further, Ortega-Ochoa et al. (2024),

Întorsureanu et al. (2025), and Wah (2025) investigated emotion-aware chatbots that provided encouragement when negative emotions were identified. These methods improved learner satisfaction and engagement by catering to emotional needs.

Prompts and reminders engage learners with interactive messages. Research by Albornoz-De Luise et al. (2023), Hu et al. (2024), and Akpan et al. (2025) developed an automated reminder system, providing prompt messages and notifications of tasks to learners. Likewise, Champaneria (2025), Oanh and Nguyen (2025), and Sajja et al. (2025) created activity monitoring tools that sent notifications if the learners were inactive. These techniques successfully diminished the risk of attrition and kept learners engaged.

Multimodal interaction allows communication in multiple modes (text, speech, visual). Several studies (Bozkurt, 2023; Choi & Kim, 2025; Milana et al., 2024) developed multimodal conversational agents, featuring speech recognition, text entry and visual content. Moreover, Kim et al. (2025), Maphalala et al. (2025), and Teng et al. (2025) implemented systems that combined voice tutoring with visual dashboards to enhance interaction options. These systems showed that multimodal interaction increases accessibility, and engagement.

Table 5. Interaction Mechanisms used by Conversational Agents for Real-Time Support and Learner Engagement

Interaction Mechanism (IM)	Cognitive Ability	Performance	Behaviour Tutoring	Explicit User Survey
Multi-Turn Conversation	Enhances analytical-thinking and critical-thinking through reasoning and dialogue (Bandi et al., 2025; Kusal et al., 2022; Mariani et al., 2023; Milana et al., 2024)	Improves answer accuracy and task completion through clarification (Aperstein et al., 2025; Love et al., 2025; Naveed et al., 2025; Prathigadapa & Daud, 2025)	Provides guided tutoring through explanations (Akpan et al., 2025; Khosrawi-Rad et al., 2022; Kim et al., 2025)	Dialogue satisfaction and clarity evaluated using user surveys (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Love et al., 2025)

Interaction Mechanism (IM)	Cognitive Ability	Performance	Behaviour Tutoring	Explicit User Survey
Real-Time Feedback Mechanisms	Improves decision-making ability by providing timely feedback and guidance (Arguedas et al., 2024; Javaid et al., 2024; Parker et al., 2025)	Provides improved learning scores and task completion time (Estévez-Ayres et al., 2025; García-Méndez et al., 2025; Reddig et al., 2025; Yousef et al., 2025; Yusuf et al., 2025)	Provides instant tutoring support for errors (Yildirim-Erbasli et al., 2025; Yin et al., 2024)	Feedback satisfaction measured using learner feedback surveys (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Parker et al., 2025)
Personalized Interaction	Supports analytical-thinking via adaptive explanations (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Cai et al., 2025; Singh et al., 2022)	Enhances individual learning and mastery (Gao, 2025; Rizwan et al., 2025; Sajja et al., 2025; Younas et al., 2025)	Tutors using adaptive strategy based on learner performance (Nazempour & Darabi, 2023; Sayed et al., 2023)	Personalization assessed through learner surveys (Albornoz-De Luise et al., 2023; Love et al., 2025)
Emotion and Sentiment-Based Interaction	Supports critical-thinking stability by reducing frustration and cognitive overload (Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025; Suresh Babu & Dhakshina Moorthy, 2024)	Improves engagement-driven performance consistency (Javaid et al., 2024; Stein & Pollak, 2024; Wah, 2025; Zhai et al., 2024)	Provides motivational and emotional tutoring support (Întorsureanu et al., 2025; Ortega-Ochoa et al., 2024)	Emotional response satisfaction measured using affective surveys (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Parker et al., 2025)
Interactive Prompts and Reminders	Supports decision-making ability by prompting learners (Albornoz-De Luise et al., 2023; Dai et al., 2024; Suresh Babu & Dhakshina Moorthy, 2024)	Improves task completion and learning retention (Hu et al., 2024; Javaid et al., 2024; Oanh & Nguyen, 2025; Sajja et al., 2025)	Encourages positive learning behaviour and task management (Akpan et al., 2025; Champaneria, 2025; Yildirim-Erbasli et al., 2025)	Reminder effectiveness evaluated using usability surveys (Albornoz-De Luise et al., 2023; Love et al., 2025)
Multimodal Interaction (Text, Voice, Visual)	Supports critical-thinking and decision-making by enriching communication using text, voice, and visuals (Al-Abri, 2025; Bozkurt, 2023; Maphalala et al., 2025; Milana et al., 2024)	Enhances comprehension, retention, and task success (Choi & Kim, 2025; Teng et al., 2025; Yusuf et al., 2025)	Provides various tutoring strategies (text, voice, and visual) (Kim et al., 2025)	Multimodal interaction satisfaction measured using affective surveys (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Parker et al., 2025)

In summary, the studies reviewed here show that conversational agents engage with users through various interaction methods, such as natural language interaction, multi-turn conversation, real-time feedback and support, personalization, emotional sensitivity and proactive prompting. These mechanisms allow conversational agents to offer ongoing support, minimize learning time, and keep learners engaged. Thus, these results answer RQ2, which explores the interaction methods of conversational agents to provide real-time support and enhance engagement in online learning systems.

3.3 RQ3: How is the Generative AI incorporated into the interaction mechanisms of conversational agents?

The chronological breakdown of GenAI based studies reveals a constant and gradual growth from 2022 to 2025. Three studies (Aperstein et al., 2025; Cai et al., 2025; Yin et al., 2024) were found in 2022, which marked the initial use of GenAI techniques in educational conversational agents. In 2023, four studies (Arguedas et al., 2024; Bozkurt, 2023; Maphalala et al., 2025; Reddig et al., 2025) were identified, suggesting an increased interest in LLM-based systems. Four studies (Choi & Kim, 2025; Love et al., 2025; Mariani et al., 2023; Wah, 2025) were also published in 2024, showing the growth and integration of GenAI techniques in education. Two studies (Casheekar et al., 2024; Milana et al., 2024) in 2025 were found, continuing the development of research. A total of 13 of these GenAI-related studies were included, most of which were published between 2023 and 2024, which demonstrates significant growth in the use of GenAI technologies in conversational educational agents in this period. Recent studies demonstrate the use of Natural Language Understanding (NLU), Natural Language Generation (NLG) and Personalization & Adaptivity, Emotional & Empathic Response, Knowledge Reasoning & Context Management, Assessment & Feedback Generation, and System Integration & Deployment to support dynamic content generation, reasoning and tutoring. These technologies enable the system to provide explanations, assessment, and feedback in real time, enhancing its adaptability and scalability for educational use (Belda-Medina & Calvo-Ferrer, 2022; Casheekar et al., 2024; Javaid et al., 2024; Mariani et al., 2023; Parker et al., 2025; Rizwan et al., 2025; Yin et al., 2024).

3.3.1 Natural Language Understanding (NLU)

Natural Language Understanding (NLU) is a key component where Generative AI is incorporated to facilitate the ability of conversational agents to understand and detect learner queries and intentions. A number of studies have used transformer-based models and conversational AI algorithms to interpret learner requests and identify user intent to support engaging conversational interactions. The study by Belda-Medina and Calvo-Ferrer (2022) developed a chatbot-based conversational learning system through the use of Natural Language Processing (NLP) algorithms and intent prediction. Similarly, Mariani et al. (2023) provided a systematic review of conversational AI-powered chatbots and identified the use of Transformer-based language models for semantic understanding and intent recognition. Likewise, Casheekar et al. (2024) conducted a survey of recent generative AI technologies and pointed out the use of deep learning-based Natural Language Understanding models, such as Bidirectional Encoder Representations from Transformers (BERT)-like models, for identifying learner requests and confusion. Also, Maphalala et al. (2025) implemented an AI-based e-tutoring system with conversational AI models, semantic parsing and intent classification algorithms. These techniques and models help conversational agents detect learner needs, contextual meanings, and support meaningful dialogue interactions, which are the essential components of the future of conversational learning environments.

3.3.2 Natural Language Generation (NLG)

NLG is an important aspect where GenAI is applied to generate useful responses. The study by Belda-Medina and Calvo-Ferrer (2022) used chatbots as conversational partners for language learning that used NLG algorithms based on template-based response generation and sequence-to-sequence (Seq2Seq) models. The study by Alfahaid and Hammami (2023) surveyed the use of conversational agents for education and noted the use of Transformer-based language generation models such as Generative Pre-trained Transformer (GPT)-like models to automatically generate explanations and examples. The study by Sayed et al. (2023) designed adaptive e-learning systems which implemented automated content generation algorithms integrated with adaptive

learning models. Further, Wang and Luo (2024) developed predictive analytics models integrated with report generation. The study by Maphalala et al. (2025) built AI-powered e-tutoring systems that used Transformer-based language models such as LLM response generation models to generate learning explanations and guidance. These models allow conversational agents to dynamically generate explanations, feedback, and provide ongoing teaching and learning support, improving engagement and learning in conversational learning environments.

3.3.3 Knowledge Reasoning and Context Management

Knowledge reasoning and context management are critical elements where GenAI supports conversational agents to keep track of discourse and provide contextually and logically coherent explanations. The study by Kusal et al. (2022) reviewed the use of AI-based conversational agents and found the use of knowledge representation models and rule-based reasoning algorithms combined with structured knowledge base systems. The study by Yin et al. (2024) developed chatbot-based formative feedback systems using context-specific response generation algorithms and interaction history models. The study by Maphalala et al. (2025) implemented AI-based e-tutoring systems using context-aware contextual models combined with semantic conceptual reasoning algorithms and learner profile databases. The study by Teng et al. (2025) explored Large Language Model (LLM) based teaching assistants using Transformer-based reasoning models integrated with long context memory. The study by Younas et al. (2025) reviewed the research on AI-powered self-directed learning systems and reported the use of Retrieval-Augmented Generation (RAG) models and knowledge management systems. These studies show that knowledge reasoning and context management mainly rely on knowledge base models, semantic reasoning algorithms, session memory modules, context window, and Retrieval-Augmented

Generation (RAG) models. These algorithms allow conversational agents to remember conversation history, access and retrieve relevant instructional information and generate contextual responses, supporting multi-turn learning interactions and enhancing learner's understanding.

3.3.4 System Integration and Deployment

System integration and deployment is the deployment layer in which GenAI conversational agents are integrated with learning platforms such as LMS, mobile apps and web-based educational software. The paper by Casheekar et al. (2024) reviewed the contemporary use of chatbots and AI-based conversational agents and presented the use of Application Programming Interfaces (API) to integrate conversational agents into educational platforms. The study by Choi and Kim (2025) built a LLM-based programming learning environment integrated with a learning platform using cloud-based deployment strategies and interactive programming platforms. The study by Love et al. (2025) explored conversational agents within structured learning environments using dialogue management systems and platform-based conversational systems. The study by Yildirim-Erbasli et al. (2025) created conversational agent-based assessment and feedback tools integrated with institutional learning platforms with automated grading engines and learning analytics dashboards. This technology allows conversational agents to act as in-built learning assistants in online education platforms, providing real-time reminders, feedback and instruction. Through the integration of GenAI models with LMS and other educational platforms and apps, conversational agents provide constant support to learners and enhance engagement in online learning.

Table 6 illustrates how GenAI converts traditional educational systems into highly personalized, empathetic and contextual tutors.

Table 6. Integration of Generative AI in Conversational Agent Components

Category	GenAI Techniques	Description	References
NLU	LLM (transformer-based) response generation, intent detection, semantic	Parses inputs (typed or voice questions, short/ambiguous sentences) to identify intent (seeking help, example, confusion)	(Belda-Medina & Calvo-Ferrer, 2022; Casheekar et al., 2024; Maphalala et al., 2025; Mariani et al., 2023; Naveed et al., 2025)
NLG	Text generation (prompt engineering, RAG)	LLM outputs (explanations, step-by-step solutions, hints, examples, follow-up questions) in natural language	(Alfehaid & Hammami, 2023; Belda-Medina & Calvo-Ferrer, 2022; Maphalala et al., 2025; Mariani et al., 2023; Sayed et al., 2023; Wang & Luo, 2024)
Knowledge Reasoning & Context Management	Chat history, memory, RAG, knowledge graph	Uses inputs (past questions, learner profile, topic history) to give outputs (context-aware answers, connected explanations)	(Kusal et al., 2022; Maphalala et al., 2025; Teng et al., 2025; Yin et al., 2024; Younas et al., 2025)
System Integration & Deployment	LLM API services, chatbot platforms, linking to LMS	Connects inputs (LMS data, assignments, deadlines) to outputs (reminders, alerts, guidance in learning platforms)	(Casheekar et al., 2024; Choi & Kim, 2025; Love et al., 2025; Stein & Pollak, 2024; Yildirim-Erbasli et al., 2025)

The results suggest that RQ3 indicates the shift from rule-based systems to integration of GenAI-powered systems that includes language processing, response generation, knowledge reasoning and system integration. But some studies continue to use early-generation models or lack pedagogical integration, suggesting the need for further research on advanced LLM capabilities, system deployment and design for education.

3.4 RQ4: What are the key metrics for evaluating the performance of conversational agents?

A comprehensive approach is needed for evaluating conversational agents in an educational context. Existing studies suggest the success of the system is not only related to the accuracy of responses but also to the improvement of learning outcomes, motivation and engagement (Albornoz-De Luise et al., 2023; Baabdullah et al., 2022; Rodrigues et al., 2022). The main evaluation metrics are summarised in Table 7. The metrics are grouped into six categories: learning outcomes, engagement, system usability, response accuracy, and emotions. A number of studies evaluated learning outcomes by using learning gain, improvement in scores, skill acquisition,

knowledge retention and achievement of learning goals, which were often measured by pre- and post-tests and long-term performance monitoring (Casheekar et al., 2024; García-Méndez et al., 2025; Hu et al., 2024; Javaid et al., 2024; Rizwan et al., 2025; Yousef et al., 2025).

Other metrics such as course completion and self-learning progress were also used to measure academic performance (Gao, 2025; Întorsureanu et al., 2025; Oanh & Nguyen, 2025; Sajja et al., 2025). Student engagement was assessed using interaction-based indicators including frequency of interaction, dialogue success rate, turn-taking fluency, and time to complete tasks, which reflect active engagement with conversational agents (Akpan et al., 2025; Albornoz-De Luise et al., 2023; Bandi et al., 2025; Milana et al., 2024; Yusuf et al., 2025). Ease of use was evaluated with user satisfaction questionnaires, System Usability Scale (SUS), response time, and system reliability, to ensure conversational systems are accessible and efficient for students (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Love et al., 2025; Parker et al., 2025). In terms of response accuracy, studies examined correctness, relevance, intention accuracy, and personalization, which are critical for

providing effective feedback and responses (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Gao, 2025; Javaid et al., 2024; Kusal et al., 2022; Milana et al., 2024; Rizwan et al., 2025; Younas et al., 2025; Yusuf et al., 2025). Furthermore, several studies measured emotional

impact through sentiment scores, motivation and persistence, underscoring the value of emotional engagement in keeping learners motivated (Akpan et al., 2025; Albornoz-De Luise et al., 2023; Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025).

Table 7. Metrics for Measuring the Performance of Conversational Agents

Category	Metric Description	References
Learning Outcomes	Learning gain, score improvement, skill mastery	(Casheekar et al., 2024; García-Méndez et al., 2025; Hu et al., 2024; Javaid et al., 2024; Rizwan et al., 2025; Yousef et al., 2025)
	Knowledge retention over time	(García-Méndez et al., 2025; Hu et al., 2024)
	Self-directed learning improvement	(Intorsureanu et al., 2025)
	Pre/post testing (learning gain measurement)	(Casheekar et al., 2024; García-Méndez et al., 2025; Hu et al., 2024; Javaid et al., 2024)
	Achievement of learning objectives	(Arguedas et al., 2024; Gao, 2025; Hu et al., 2024)
	Instructional alignment and scaffolding	(García-Méndez et al., 2025)
	Course completion rate	(Intorsureanu et al., 2025; Oanh & Nguyen, 2025)
	Long-term performance trends	(Sajja et al., 2025; Yousef et al., 2025)
Student Engagement	Interaction frequency and duration	(Akpan et al., 2025; Albornoz-De Luise et al., 2023)
	Dialogue success rate	(Bandi et al., 2025; Milana et al., 2024; Yusuf et al., 2025)
	Turn-taking smoothness	(Bandi et al., 2025; Milana et al., 2024)
	Repair and clarification success	(Kusal et al., 2022; Yusuf et al., 2025)
	Time to complete tasks	(Hu et al., 2024; Yusuf et al., 2025)
	Task completion rate	(Albornoz-De Luise et al., 2023; García-Méndez et al., 2025)
System Usability	Likert-scale user satisfaction surveys	(Albornoz-De Luise et al., 2023; Bozkurt, 2023; Love et al., 2025)
	System Usability Scale (SUS)	(Bozkurt, 2023; Parker et al., 2025)
	Latency and response time	(Bandi et al., 2025; Kusal et al., 2022)
	System reliability and error rate	(Bandi et al., 2025)
	Perceived trust and transparency	(Bozkurt, 2023; Love et al., 2025)

Category	Metric Description	References
	Privacy and data protection perception	(Bozkurt, 2023; Naveed et al., 2025)
Response Accuracy	Correctness of answers provided by agent	(Bozkurt, 2023; Kusal et al., 2022; Yusuf et al., 2025)
	Relevance and contextual appropriateness	(Javaid et al., 2024; Rizwan et al., 2025)
	Intent recognition accuracy	(Kusal et al., 2022; Milana et al., 2024)
	Accuracy of content personalization	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Gao, 2025; Younas et al., 2025)
	Learner-model update effectiveness	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Stein & Pollak, 2024)
Emotional Impact	Emotional valence and sentiment scores	(Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025)
	Motivation and affective engagement indicators	(Kim et al., 2025; Krstić et al., 2022; Parker et al., 2025)
	Behavioural persistence and engagement continuity	(Akpan et al., 2025; Albornoz-De Luise et al., 2023)

4. Discussion

4.1 Research Gap

The study of the 54 papers selected highlights that, despite the rapid increase in research on conversational agents and AI in education, there are common limitations in many studies. A large number of papers focus on a narrow set of data such as behavioral logs (e.g., click, time-on-task or message counts), and do not consider other emotional, contextual, and multimodal aspects to a sufficient extent. The papers show some common limitations across various research fields. These issues reveal gaps in methodological approaches, evaluation approaches, deployment strategies, and ethical issues that need to be considered in future studies. A key limitation is the lack of multimodal and contextual data, where most studies primarily used interaction logs (e.g., clicks and interactions) and ignored other data such as vocal signals, facial features and contextual environment information (Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Albornoz-De Luise et al., 2023; Wah, 2025; Younas et al., 2025). A further limitation is the absence of real classroom validation, where some systems were tested in laboratory or simulated classrooms rather than real classrooms, limiting their potential to be used in real classroom environments (Bandi et al., 2025; Casheekar et al.,

2024; García-Méndez et al., 2025; Hu et al., 2024; Khosrawi-Rad et al., 2022; Kusal et al., 2022; Teng et al., 2025; Yusuf et al., 2025).

Table 8 summarises the identified research gaps. Short evaluation periods were also identified as a limitation, where studies were limited to short periods of experimentation, and it's hard to track learning outcomes and sustained learner interest (Javaid et al., 2024; Krstić et al., 2022; Nazempour & Darabi, 2023; Rizwan et al., 2025; Yousef et al., 2025). Also, many studies demonstrate limited use of recent LLMs and continue to use older rule-based or early Natural Language Processing techniques that do not leverage the power of recent advances in GenAI technologies (Arguedas et al., 2024; Baabdullah et al., 2022; Champaneria, 2025; Maphalala et al., 2025; Reddig et al., 2025; Sayed et al., 2023; Yin et al., 2024).

Further, we observe an area of bias in subject domains where research is primarily focused on specific learning areas (such as language learning or programming) (Bozkurt, 2023; Choi & Kim, 2025; Javaid et al., 2024; Milana et al., 2024; Sajja et al., 2025; Singh et al., 2022; Yusuf et al., 2025). Likewise, poor personalization models were found in some studies, where personalization was mainly based on behavioral data, without extensive learner profiling and adaptive learning models

(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Alrajhi, 2024; Kim et al., 2025; Wah, 2025). There is also a lack of explainable feedback, suggesting many AI-based feedback models operate as black-box models without explanations of reasoning for the provided feedback (Krstić et al., 2022; Love et al., 2025; Nazempour & Darabi, 2023; Rizwan et al., 2025; Yousef et al., 2025).

In addition, low emotional intelligence is also observed, with emotional and affective learning factors such as motivation, frustration, and engagement often underused in conversational agent design and testing (Întorsureanu et al., 2025; Krstić et al., 2022; Ortega-Ochoa et al., 2024;

Parker et al., 2025; Wah, 2025). Finally, ethical aspects are also highlighted in the ethical and trust issues category, where ethical principles such as transparency, fairness, privacy, and ethical use of AI systems are often mentioned but less empirically tested (Albornoz-De Luise et al., 2023; Bozkurt, 2023; Dai et al., 2024; Oanh & Nguyen, 2025; Parker et al., 2025; Zhai et al., 2024). Finally, there was a lack of scalability testing, where many systems were not validated in large-scale deployment scenarios or with different types of learners, raising concerns about their scalability and generalizability (García-Méndez et al., 2025; Nazempour & Darabi, 2023; Teng et al., 2025; Yildirim-Erbaşlı et al., 2025).

Table 8. Identified Research Gaps

Research Gap	Description	Citations
Lack of Exploring Multimodal and Contextual Data	Mostly focused on behavioural logs without considering emotional, contextual and multimodal features such as voice, facial expressions, and environmental factors.	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Albornoz-De Luise et al., 2023; Wah, 2025; Younas et al., 2025)
Lack of Classroom Validation	Tested in laboratory or simulated classroom settings, rather than in real classrooms, making it difficult for practical application.	(Bandi et al., 2025; Casheekar et al., 2024; García-Méndez et al., 2025; Hu et al., 2024; Khosrawi-Rad et al., 2022; Kusal et al., 2022; Teng et al., 2025; Yusuf et al., 2025)
Limited Time in Experiments	Perform short-term experiments, making it hard to measure long-term learning and engagement.	(Javaid et al., 2024; Krstić et al., 2022; Nazempour & Darabi, 2023; Rizwan et al., 2025; Yousef et al., 2025)
Lack of Integration of LLMs	Use rule-based or conventional NLP approaches without considering recent large language model technologies.	(Arguedas et al., 2024; Baabdullah et al., 2022; Champaneria, 2025; Maphalala et al., 2025; Reddig et al., 2025; Sayed et al., 2023; Yin et al., 2024)
Bias towards Specific Domains	Primarily focus on domains like language learning or computer science, neglecting others.	(Bozkurt, 2023; Choi & Kim, 2025; Javaid et al., 2024; Milana et al., 2024; Sajja et al., 2025; Singh et al., 2022; Yusuf et al., 2025)
Poor Personalization Models	Personalization strategies rely only on learners' interaction data without detailed profiling techniques.	(Akhuseyinoglu & Brusilovsky, 2022; Al-Abri, 2025; Alrajhi, 2024; Kim et al., 2025; Wah, 2025)
Weak Explanation of Feedback	AI-driven feedback is unexplained and provided as a black box with no rationale for the generated feedback.	(Krstić et al., 2022; Love et al., 2025; Naveed et al., 2025; Nazempour & Darabi, 2023; Rizwan et al., 2025; Yousef et al., 2025)
Low Emotional Intelligence	Analytical models lack consideration of emotional and motivational learning aspects in conversational agent design and evaluation.	(Întorsureanu et al., 2025; Krstić et al., 2022; Ortega-Ochoa et al., 2024; Parker et al., 2025; Wah, 2025)

Research Gap	Description	Citations
Ethical and Trust Issues	Ethical aspects such as transparency, fairness, privacy protection, and responsible AI use are often discussed but rarely validated experimentally.	(Albornoz-De Luise et al., 2023; Bozkurt, 2023; Dai et al., 2024; Oanh & Nguyen, 2025; Parker et al., 2025; Zhai et al., 2024)
Lack of Scalability Testing	Systems are not evaluated in large-scale settings or with a variety of different learners.	(García-Méndez et al., 2025; Nazempour & Darabi, 2023; Teng et al., 2025; Yildirim-Erbasli et al., 2025)

Table 8 shows that although conversational agent research has evolved technologically, there are still many research and application gaps. Closing these research gaps, especially in the areas of multimodal integration, practical validation, longitudinal assessment, explainability, emotional intelligence, ethical considerations and scalability, will be crucial for future research to deliver effective, reliable and user-friendly conversational educational systems.

5. Conclusion

This Systematic Literature Review (SLR) examined 54 research papers from 2022 to 2025 exploring the use of conversational agents and GenAI in education. This research shows a clear evolution of conversational educational systems from rule-based and ML systems to DL systems, Large Language Models (LLMs), and Generative AI-based approaches. For instance, recent years have seen a significant shift towards using transformer-based models and retrieval-augmented systems, allowing conversational agents to provide dynamic responses, adaptive content, and tailored feedback. The evolution analysis shows a decrease in the use of rule-based systems and a sharp rise in the use of LLM-based and GenAI-based systems, revealing the growing trend of intelligent content generation and reasoning.

In terms of input data and AI techniques, the systematic review demonstrates a growth in the diversity of data used for learning and engagement analysis. The majority of systems use structured input data sources like learning logs, chatbot dialogues, performance records and feedback responses. Recently, sentiment and emotional data from textual interactions are also being used to enable adaptive learning. But according to the findings, multimodal data such as voice, facial expressions and environmental factors have limited use in these systems, suggesting a potential avenue for future research to improve the diversity

and contextual relevance of the inputs used in conversational learning systems.

The results for RQ1 (Inputs, Outputs and Methods for Low Engagement) indicate that conversational agents consistently use interaction logs, history and surveys to detect learner engagement. A range of AI techniques including ML classifiers, DL approaches, and sentiment analysis are commonly employed to identify patterns of low engagement and produce adaptive responses such as motivational feedback and learning recommendations. But many studies use self-reported data and limited behavioural cues, and do not adequately account for emotional and contextual influences.

The findings for RQ2 (Real-Time Support and Engagement) show that conversational agents engage with students through multi-turn dialogue systems, real-time feedback, personalised interaction approaches, and sentiment-driven interaction models. These approaches contribute to cognitive development through improved decision-making, analysis and reasoning. However, the review reveals the lack of integrated learner models and the lack of contextual and emotional information in real-time decision-making.

The findings for RQ3 (Integration of GenAI) reveal a high and growing dependence on transformer-based LLM models for Natural Language Understanding (NLU) and Natural Language Generation (NLG), context management and adaptation, and dynamic content delivery. GenAI models are extensively applied to generate explanations, hints, examples and feedback using human language, allowing conversational agents to mimic intelligent tutoring. Our trend analysis shows GenAI and LLM-based conversational agents have emerged in the years 2023–2025 and have replaced previous rule-based systems. But many studies still lack pedagogical and multimodal reasoning, suggesting that more

comprehensive educational design principles and frameworks are needed.

Finally, the results for RQ4 (Performance Metrics and Evaluation) reveal that the success of conversational agents is assessed through a range of metrics such as learning gain, engagement, system usability, response accuracy, emotional effect and ethical trust. The most common metrics for evaluating conversational agents include learning gain, task completion, dialogue success, and system usability. However, the review reveals a lack of long-term evaluation evidence, explainable feedback, and classroom deployment testing with large-scale user data, which calls for more comprehensive and longitudinal approaches to evaluation.

In conclusion, this review shows that the use of conversational agents with GenAI and LLM technologies has the potential to radically enhance learning systems by providing personalised learning, intelligent feedback and real-time interaction. The broadening of AI approaches and data inputs aligns with the advancement of educational technologies. But future studies need to focus on multimodal data sources, explainable and ethical AI systems, longitudinal classroom testing and integrated models that combine technical and pedagogic aspects. Overcoming these issues will be crucial to developing robust, scalable and learner-focused conversational educational systems to support diverse learning contexts.

References

- Akhuseyinoglu, K., & Brusilovsky, P. (2022). Exploring behavioural patterns for data-driven modeling of learners' individual differences. *Frontiers in Artificial Intelligence*, 5, Article 807320. <https://doi.org/10.3389/frai.2022.807320>
- Akpan, I. J., Kobara, Y. M., Owolabi, J., Akpan, A. A., & Offodile, O. F. (2025). Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *International Transactions in Operational Research*, 32(3), 1251–1281. <https://doi.org/10.1111/itor.13522>
- Al-Abri, A. (2025). Exploring ChatGPT as a virtual tutor: A multi-dimensional analysis of large language models in academic support. *Education and Information Technologies*. Advance online publication. <https://doi.org/10.1007/s1338-367-13484-x>
- Albornoz-De Luise, R. S., Arevalillo-Herráez, M., & Arnau, D. (2023). On using conversational frameworks to support natural language interaction in intelligent tutoring systems. *IEEE Transactions on Learning Technologies*, 16(5), 722–735. <https://doi.org/10.1109/TLT.2023.3245121>
- Alfehaid, A., & Hammami, M. A. (2023). Artificial intelligence in education: Literature review on the role of conversational agents in improving learning experience. *International Journal*, 10(3), 3121–3129. <https://doi.org/10.15379/ijmst.v10i3.3045>
- Alrajhi, A. S. (2024). Artificial intelligence pedagogical chatbots as L2 conversational agents. *Cogent Education*, 11(1), Article 2327789. <https://doi.org/10.1080/2331186X.2024.2327789>
- Aperstein, Y., Cohen, Y., & Apartsin, A. (2025). Generative AI-based platform for deliberate teaching practice: A review and a suggested framework. *Education Sciences*, 15(4), Article 405. <https://doi.org/10.3390/educsci15040405>
- Arguedas, M., Daradoumis, T., & Caballé, S. (2024). Measuring the effects of pedagogical agent cognitive and affective feedback on students' academic performance. *Frontiers in Artificial Intelligence*, 7, Article 1495342. <https://doi.org/10.3389/frai.2024.1495342>
- Baabdullah, A. M., Alalwan, A. A., Algharabat, R. S., Metri, B., & Rana, N. P. (2022). Virtual agents and flow experience: An empirical examination of AI-powered chatbots. *Technological Forecasting and Social Change*, 181, Article 121772. <https://doi.org/10.1016/j.techfore.2022.121772>
- Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., & Vilipala, S. V. (2025). The rise of agentic AI: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges. *Future Internet*, 17(9), Article 404. <https://doi.org/10.3390/fi17090404>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), Article 8427. <https://doi.org/10.3390/app12178427>
- Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, 18(1). <https://doi.org/10.5281/zenodo.7716416>

- Cai, C., Hong, S., Ma, M., Feng, H., Du, S., Chow, M., Teo, W. L.-L., Liu, S., & Fan, X. (2025). Analyzing the teaching and learning environments through student feedback at scale: A multi-agent LLMs framework. *Education and Information Technologies*, 30(15), 21815–21847. <https://doi.org/10.1007/s1338-367-1338-367>
- Casheekar, A., Lahiri, A., Rath, K., Prabhakar, K. S., & Srinivasan, K. (2024). A contemporary review on chatbots, AI-powered virtual conversational agents, ChatGPT: Applications, open challenges, and future research directions. *Computer Science Review*, 52, Article 100632. <https://doi.org/10.1016/j.cosrev.2024.100632>
- Champaneria, H. K. (2025). From insight to impact: Architecting AI-driven learning ecosystems for personalized, predictive, and proactive education. *Journal of Applied Sciences*, 5(12), 10–20. <https://doi.org/10.5281/zenodo.17806013>
- Choi, S., & Kim, H. (2025). The impact of a large language model-based programming learning environment on students' motivation and programming ability. *Education and Information Technologies*, 30(6), 8109–8138. <https://doi.org/10.1007/s1338-367-13107-x>
- Dai, C. P., Ke, F., Pan, Y., Moon, J., & Liu, Z. (2024). Effects of artificial intelligence-powered virtual agents on learning outcomes in computer-based simulations: A meta-analysis. *Educational Psychology Review*, 36(1), Article 31. <https://doi.org/10.1007/s1338-367-0338-367>
- Estévez-Ayres, I., Callejo, P., Hombrados-Herrera, M. Á., Alario-Hoyos, C., & Delgado Kloos, C. (2025). Evaluation of LLM tools for feedback generation in a course on concurrent programming. *International Journal of Artificial Intelligence in Education*, 35(2), 774–790. <https://doi.org/10.1007/s4338-367-0338-367>
- Gao, Y. (2025). Deep learning-based strategies for evaluating and enhancing university teaching quality. *Computers and Education: Artificial Intelligence*, 8, Article 100362. <https://doi.org/10.1016/j.caeai.2025.100362>
- García-Méndez, S., de Arriba-Pérez, F., & Somoza-López, M. D. C. (2025). A review on the use of large language models as virtual tutors. *Science & Education*, 34(2), 877–892. <https://doi.org/10.1007/s1338-367-0338-367>
- Hu, J., Huang, Z., Li, J., Xu, L., & Zou, Y. (2024). Real-time classroom behaviour analysis for enhanced engineering education: An AI-assisted approach. *International Journal of Computational Intelligence Systems*, 17(1), Article 167. <https://doi.org/10.1007/s4338-367-00572-y>
- Întorsureanu, I., Oprea, S. V., Bâra, A., & Vespan, D. (2025). Generative AI in education: Perspectives through an academic lens. *Electronics*, 14(5), Article 1053. <https://doi.org/10.3390/electronics14051053>
- Javaid, Z. K., Ijaz, S., Latif, E., & Azam, S. (2024). A qualitative study on student–teacher interaction: Student engagement and teacher feedback. *International Journal of Contemporary Issues in Social Sciences*, 3(2), 1670–1683.
- Khosrawi-Rad, B., Rinn, H., Schlimbach, R., Gebbing, P., Yang, X., Lattemann, C., Markgraf, D., & Robra-Bissantz, S. (2022). Conversational agents in education – A systematic literature review. In *Proceedings of the 30th European Conference on Information Systems (ECIS 2022)*. AIS Electronic Library. https://aisel.aisnet.org/ecis2022_rp/18
- Kim, J., Yu, S., Detrick, R., Lin, X., & Li, N. (2025). Designing AI-powered learning: Adult learners' expectations for curriculum and human–AI interaction. *Educational Technology Research and Development*, 73(6), 3397–3421. <https://doi.org/10.1007/s1338-367-10549-z>
- Krstić, L., Aleksić, V., & Krstić, M. (2022). *Artificial intelligence in education: A review*. <https://doi.org/10.46793/TIE22.223K>
- Kusal, S., Patil, S., Choudrie, J., Kotecha, K., Mishra, S., & Abraham, A. (2022). AI-based conversational agents: A scoping review from technologies to future directions. *IEEE Access*, 10, 92337–92356. <https://doi.org/10.1109/ACCESS.2022.3201144>
- Love, R., Law, E., Cohen, P. R., & Kulić, D. (2025). Teaching a conversational agent using natural language: Effect on learning and engagement. *International Journal of Artificial Intelligence in Education*. Advance online publication. <https://doi.org/10.1007/s4338-367-0338-367>
- Maphalala, M. C., Mkhasebe, R. G., & Mncube, D. W. (2025). Exploring the roles of AI-powered e-tutors in enhancing self-directed learning in open distance e-learning courses. *Interdisciplinary Journal of Education Research*, 7(1), Article a12. <https://doi.org/10.38140/ijer-2025.vol7.1.12>

- Mariani, M. M., Hashemi, N., & Wirtz, J. (2023). Artificial intelligence-empowered conversational agents: A systematic literature review and research agenda. *Journal of Business Research*, 161, Article 113838. <https://doi.org/10.1016/j.jbusres.2023.113838>
- Milana, M., Brandi, U., Hodge, S., & Hoggan-Kloubert, T. (2024). Artificial intelligence (AI), conversational agents, and generative AI: Implications for adult education practice and research. *International Journal of Lifelong Education*, 43(1), 1–7. <https://doi.org/10.1080/02601370.2024.2310448>
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., . . . Mian, A. (2025). A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology*, 16(5), 1–72. <https://doi.org/10.1145/3744746>
- Nazempour, R., & Darabi, H. (2023). Personalized learning in virtual learning environments using students' behaviour analysis. *Education Sciences*, 13(5), Article 457. <https://doi.org/10.3390/educsci13050457>
- Oanh, D. T. K., & Nguyen, T. N. T. (2025). Investigating the effect of lecturer–student interaction on student engagement in class. *International Journal of Instruction*, 18(3), 239–262. <https://doi.org/10.29333/iji.2025.18313a>
- Ortega-Ochoa, E., Arguedas, M., & Daradoumis, T. (2024). Empathic pedagogical conversational agents: A systematic literature review. *British Journal of Educational Technology*, 55(3), 886–909. <https://doi.org/10.1111/bjet.13413>
- Parker, M. J., Anderson, C., Stone, C., & Oh, Y. (2025). A large language model approach to educational survey feedback analysis. *International Journal of Artificial Intelligence in Education*, 35(2), 444–481. <https://doi.org/10.1007/s4338-367-0338-367>
- Prathigadapa, S., & Daud, S. B. M. (2025). A review of virtual tutoring systems and student performance analysis using GPT-3. *Journal of Learning for Development*, 12(1), 168–181. <https://doi.org/10.56059/jl4d.v12i1.1367>
- Reddig, J. M., Arora, A., & MacLellan, C. J. (2025). Generating in-context, personalized feedback for intelligent tutors with large language models. *International Journal of Artificial Intelligence in Education*. Advance online publication. <https://doi.org/10.1007/s4338-367-0338-367>
- Rizwan, S., Nee, C. K., & Garfan, S. (2025). Identifying the factors affecting student academic performance and engagement prediction in MOOC using deep learning: A systematic literature review. *IEEE Access*, 13. <https://doi.org/10.1109/ACCESS.2025.3533915>
- Rodrigues, C., Reis, A., Pereira, R., Martins, P., Sousa, J., & Pinto, T. (2022). A review of conversational agents in education. In *International Conference on Technology and Innovation in Learning, Teaching and Education* (pp. 461–467). Springer Nature Switzerland. https://doi.org/10.1007/338-367-338-367338-367_37
- Sajja, R., Sermet, Y., Cwiertny, D., & Demir, I. (2025). Integrating AI and learning analytics for data-driven pedagogical decisions and personalized interventions in education. *Technology, Knowledge and Learning*. Advance online publication. <https://doi.org/10.1007/s1338-367-0338-367>
- Sayed, W. S., Noeman, A. M., Abdellatif, A., Abdelrazek, M., Badawy, M. G., Hamed, A., & El-Tantawy, S. (2023). AI-based adaptive personalized content presentation and exercises navigation for an effective and engaging e-learning platform. *Multimedia Tools and Applications*, 82(3), 3303–3333. <https://doi.org/10.1007/s1338-367-1338-367>
- Singh, N., Gunjan, V. K., & Nasralla, M. M. (2022). A parameterized comparative analysis of performance between the proposed adaptive and personalized tutoring system "Seis Tutor" with existing online tutoring systems. *IEEE Access*, 10, 39376–39386. <https://doi.org/10.1109/ACCESS.2022.3166261>
- Stein, A. G., & Pollak, S. D. (2024). Computational modeling approaches to emotional development. *Developmental Psychology*. Advance online publication. <https://doi.org/10.1037/dev0001830>
- Suresh Babu, S., & Dhakshina Moorthy, A. (2024). Application of artificial intelligence in adaptation of gamification in education: A literature review. *Computer Applications in Engineering Education*, 32(1), Article e22683. <https://doi.org/10.1002/cae.22683>

- Teng, D., Wang, X., Xia, Y., Zhang, Y., Tang, L., Chen, Q., Zhang, R., Xie, S., & Yu, W. (2025). Investigating the utilization and impact of large language model-based intelligent teaching assistants in flipped classrooms. *Education and Information Technologies*, 30(8), 10777–10810. <https://doi.org/10.1007/s1338-367-13264-z>
- Wah, J. N. (2025). Artificial intelligence in language learning: A systematic review of personalization and learner engagement. *Forum for Linguistic Studies*, 7(9), 327–341.
- Wang, S., & Luo, B. (2024). Academic achievement prediction in higher education through interpretable modeling. *PLOS ONE*, 19(9), Article e0309838. <https://doi.org/10.1371/journal.pone.0309838>
- Yildirim-Erbasli, S., Bulut, O., Epp, C. D., & Cui, Y. (2025). Advancing higher education students' assessment experiences with conversational agents. *Educational Technology Research and Development*. Advance online publication. <https://doi.org/10.1007/s1338-367-1338-367>
- Yin, J., Goh, T. T., & Hu, Y. (2024). Using a chatbot to provide formative feedback: A longitudinal study of intrinsic motivation, cognitive load, and learning performance. *IEEE Transactions on Learning Technologies*, 17, 1404–1415. <https://doi.org/10.1109/TLT.2024.3364015>
- Younas, M., El-Dakhs, D. A. S., & Jiang, Y. (2025). A comprehensive systematic review of AI-driven approaches to self-directed learning. *IEEE Access*, 13. <https://doi.org/10.1109/ACCESS.2025.3546319>
- Yousef, M., Mohamed, K., Medhat, W., Mohamed, E. H., Khoriba, G., & Arafa, T. (2025). BeGrading: Large language models for enhanced feedback in programming education. *Neural Computing and Applications*, 37(2), 1027–1040. <https://doi.org/10.1007/s0338-367-10449-y>
- Yusuf, H., Money, A., & Daylamani-Zad, D. (2025). Pedagogical AI conversational agents in higher education: A conceptual framework and survey of the state of the art. *Educational Technology Research and Development*, 73(2), 815–874. <https://doi.org/10.1007/s1338-367-1338-367>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11(1), Article 28. <https://doi.org/10.1186/s4338-367-0338-367>