An LLM-Driven Chatbot in Higher Education for Databases and Information Systems

Alexander Tobias Neumann, Yue Yin, Sulayman Sowe, Stefan Decker, and Matthias Jarke, *Life Senior Member, IEEE*

Abstract*—*Contribution: **This research explores the benefits and challenges of developing, deploying, and evaluating a large language model (LLM) chatbot, MoodleBot, in computer science classroom settings. It highlights the potential of integrating LLMs into LMSs like Moodle to support self-regulated learning (SRL) and help-seeking behavior.**

Background: **Computer science educators face immense challenges incorporating novel tools into LMSs to create a supportive and engaging learning environment. MoodleBot addresses this challenge by offering an interactive platform for both students and teachers.**

Research Questions: **Despite issues like bias, hallucinations, and teachers' and educators' resistance to embracing new (AI) technologies, this research investigates two questions: (RQ1) To what extent do students accept MoodleBot as a valuable tool for learning support? (RQ2) How accurately does MoodleBot churn out responses, and how congruent are these with the established course content?**

Methodology: **This study reviews pedagogical literature on AI-driven chatbots and adopts the retrieval-augmented generation (RAG) approach for MoodleBot's design and data processing. The technology acceptance model (TAM) evaluates user acceptance through constructs like perceived usefulness (PU) and Ease of Use. Forty-six students participated, with 30 completing the TAM questionnaire.**

Findings: **LLM-based chatbots like MoodleBot can significantly improve the teaching and learning process. This study revealed a high accuracy rate (88%) in providing course-related assistance. Positive responses from students attest to the efficacy and applicability of AI-driven educational tools. These findings indicate that educational chatbots are suitable for integration into courses to improve personalized learning and reduce teacher administrative burden, although improvements in automated fact-checking are needed.**

***Index Terms*—Chatbots, higher education, large language model (LLM), moodle, moodlebot.**

I. Introduction

In the rapidly evolving technological landscape, Artificial Intelligence (AI) has carved a niche in various sectors, including education. Chatbots, especially AI-driven chatbots, have become a prominent educational tool [1], [2]. They are software applications designed to simulate human conversation through predetermined scripts or, increasingly, more sophisticated natural language processing (NLP) algorithms [3]. Their capacity to generate human-like responses showcases the influence AI on the broader educational paradigm. The growing complexity and diversity of data within learning management systems (LMS) present unique challenges, hindering their effectiveness. Former iterations of AI-driven chatbots face challenges, gaining acceptance and achieving widespread success in educational settings. These challenges primarily revolve around their limited understanding of context, rigid script-based interactions, and suboptimal user experiences [2]. However, the advent of large language models (LLMs), including technologies, such as ChatGPT, marks a significant evolution from these earlier AI-driven bots [4]. This is where LLM-driven solutions promise transformative change. Throughout this article, AI-driven chatbots are defined as the AI solutions for chatbots before the advent of LLM, which are characterized as LLM-driven chatbots enabling more dynamic and contextually aware interactions. LLM-driven chatbots can fill information gaps and significantly improve user experiences in such environments. For instance, chatbots can serve as relentless academic aides when integrated into an LMS [5]. They can help students find pertinent course materials or resolve organizational issues by employing similarity searches over the course content. Their unique capacity to provide on-demand explanations and entertain follow-up queries positions them as virtual tutors. With rapid, anytime-available support, these systems potentially eliminate the need for students to repeatedly reinforce already understood materials, thus enhancing their learning progress and overall motivation [6].

This article presents MoodleBot, an open-source chatbot designed to provide feedback on lecture content and tasks tailored for self-regulated learning (SRL). Moodle,[1] the platform for which MoodleBot is developed, is a widely used LMS in

Received 13 November 2023; revised 18 May 2024 and 6 September 2024; accepted 20 September 2024. Date of publication 7 October 2024; date of current version 5 February 2025. This work was supported by the German Federal Ministry of Education and Research (BMBF) through the Project Tech4compKI under Grant 16DHB2213. *(Corresponding author: Alexander Tobias Neumann.)*

Alexander Tobias Neumann and Yue Yin are with the Department of Computer Science 5, RWTH Aachen University, 52074 Aachen, Germany (e-mail: neumann@dbis.rwth-aachen.de; yue.yin@rwth-aachen.de).

Sulayman Sowe is with the Department of Computer Science 5, RWTH Aachen University, 52074 Aachen, Germany, and also with the Data Science and Artificial Intelligence Department, Fraunhofer Institute for Applied Information Technology, 53757 Sankt Augustin, Germany (e-mail: sowe@dbis.rwth-aachen.de).

Stefan Decker is with the Department of Computer Science 5, RWTH Aachen University, 52074 Aachen, Germany, and also with the Fraunhofer Institute for Applied Information Technology, 53757 Sankt Augustin, Germany (e-mail: decker@dbis.rwth-aachen.de).

Matthias Jarke, deceased, was with the Department of Computer Science 5, RWTH Aachen University, 52074 Aachen, Germany, and also with Fraunhofer Institute for Applied Information Technology, 53757 Sankt Augustin, Germany (e-mail: jarke@dbis.rwth-aachen.de).

Digital Object Identifier 10.1109/TE.2024.3467912

[1]https://moodle.org

higher education and used by the RWTH Aachen University. The study focuses on its deployment among students from the lecture on "Databases and Information Systems" at the RWTH Aachen University, a mandatory bachelor's computer science lecture with over 700 participants. With its human-like conversational process, MoodleBot offers students an experience similar to interacting with a real tutor but with the advantages of immediate responses and round-the-clock availability. This article addresses the following research questions.

RQ1: To what extent do students accept MoodleBot as a valuable tool for learning support?

RQ2: How accurately does MoodleBot churn out responses, and how congruent are these with the established course content?

The implications of the findings extend beyond academic discourse, holding value for educators, administrators, and software developers by shedding light on the dynamics and potential of LLM-augmented tools in educational settings.

II. Background and Related Work

AI-driven chatbots in educational contexts have seen significant advances in their application within academia. These chatbots fulfill diverse functions, including providing instant feedback on assignments [2], [7], [8], assisting with course-related queries [2], [9], [10], streamlining enrollment processes [10], [11], and disseminating campus information [10], [12], [13]. Mentoring bots, a specialized subset, offer 24/7 instructional and support services [6], notably easing the workload of teachers and teaching assistant (TA) [14]. Conversational agents show promise for delivering personalized tutoring in educational contexts [15], [16]. Some studies demonstrate the efficacy of mentoring chatbots in providing individualized learning support [17]. Recent work explores using chatbots to tailor tutoring to each learner's needs and characteristics [18]. Such bots can be seamlessly integrated into online courses [5], [6]. They can enhance student engagement within the learning community [19]. The output accuracy of most educational chatbots highly depends on the input data [20]. Personalized hybrid e-learning models are suggested to consider a student's personality, tailoring the chatbot's interactions accordingly [21]. Another application of chatbots is in teaching programming languages, with findings indicating a significant uptick in user satisfaction when social dialogue is incorporated [22].

A. Self-Regulated Learning

SRL is a paradigm that enables learners to enhance their education by setting specific goals, monitoring performance, and adjusting behaviors through cognitive, metacognitive, and motivational strategies to optimize outcomes [23]. The three-layer model of SRL includes the regulation of processing modes, learning processes, and the self, emphasizing the dynamic interplay between these elements [24]. Self-regulation enhances academic performance by integrating cognitive strategies and metacognition [25], as well as motivation, engagement, and social support [26]. Integrating chatbots in e-learning platforms can reduce feelings of isolation and detachment while boosting intrinsic motivation and perceived competence [27], especially when combined with traditional teacher support [28]. Deploying AI-driven chatbots, guided by pedagogical strategies, such as goal setting and personalized feedback, can enhance learners' SRL skills [29]. These chatbots can support goal articulation [30] and pose strategic questions to facilitate SRL [31]. The application of LLM exemplifies this dynamic, as these tools deliver immediate, contextually appropriate information and support learners in reflective and self-regulatory practices [32]. LLM can simultaneously challenge and support the development of SRL skills. Learners can benefit from receiving alternative solutions, being exposed to diverse perspectives, and encouraging critical thinking. However, it is important to note that excessive use of LLM may diminish the learner's capacity for SRL. Therefore, educators must design interventions that balance the facilitation and disruption of self-regulation [33].

B. Students' Help-Seeking Behaviors

While engaging in SRL, students may encounter challenges or identify gaps in their knowledge and skills. Help-seeking is an essential academic strategy for self-regulation that facilitates learning [34]. It is considered an important form of behavioral self-regulation that can cognitively, behaviorally, and emotionally engage learners [35]. However, many students, particularly adolescents, avoid seeking help even when they need it [36]. This help avoidance behavior can be attributed to various factors, including competence concerns (fear of appearing incompetent) and autonomy concerns (desire to work independently) [36]. Students with low self-efficacy (SE) or those focused on performance goals are particularly prone to avoiding help-seeking due to concerns about negative judgments from teachers or peers [36]. Effective help-seeking is timely and context-dependent, with early help-seeking in problem-solving associated with better learning outcomes [37]. In online environments, it is beneficial to seek help on challenging steps; however, overuse can reduce learning outcomes [38]. Effective help-seeking behaviors include asking precise questions and persisting in seeking help, while effective help-giving involves providing detailed explanations and monitoring student understanding [39]. Chatbots, as AI-driven educational tools, have the potential to address these concerns and alleviate help avoidance behavior. They provide a private, nonjudgmental space for students to ask questions without fear of embarrassment in front of peers or teachers [36]. Regarding accessibility, office hours are essential for addressing students' queries about assignments [40], course content, and administrative issues. Integrating LLM can enhance these interactions by offering real-time, accessible solutions and promoting more proactive and effective engagement with learning resources.

C. LLM-Driven Education

In the domain of natural language generation (NLG), LLM represent a significant advancement, particularly those founded on the transformer architecture with self-attention mechanisms [41]. These models are distinguished by their capability to produce contextually relevant and coherent human-like text [42]. The generative pretrained transformer (GPT) series, a family of models, has gained recognition in this space. Among these, ChatGPT, a specialized chat model developed by OpenAI, is built on the architecture of specific GPT versions and tailored for interactive applications.[2] OpenAI offers specialized fine-tuning for the GPT models, facilitating deployment across diverse domains and applications.

LLM offer versatile support in education, facilitating tasks, such as providing comprehensive feedback on assignments, discussing intricate concepts, annotating code to highlight errors, and generating exercises along with sample solutions [43]. The research presented in this article builds upon these capabilities, providing comprehensive mentoring support to students, addressing their needs ranging from organizational queries to conceptual discussions and exercise generation for exam preparation. This holistic approach sets the presented work in this article apart from studies [44], [45], which focus primarily on content support and administrative queries.

However, challenges remain, including control over its responses [44], [46] and occasional "hallucinations," where it produces incorrect content [47], [48], [49], [50]. The presented solution in this article adopts an retrieval-augmented generation (RAG) approach [51] that can enhance the accuracy and correctness of responses [44], [52]. Similar works offer either administrative and content support [45], [52], exercise support [53], [54], [55] or exercise generation [56], [57], but not both in one package. Furthermore, instead of using a pool of tasks only [44], [50], this approach involves assessing students' acceptance of the tool using technology acceptance model (TAM). While existing studies often discuss the accuracy of such models [44], [45], [48], [52], [54], [56], [57], none of them conduct a TAM analysis. Despite students' awareness of these models [58], only a few explored their willingness to adopt them in their current state [52], [59]. Moreover, while the technical scalability is frequently mentioned [46], [50], [52], [60], [61], only one examined the associated costs [52]. Among the reviewed studies, one notable paper addresses a similar database course utilizing an RAG approach [44]. However, their validation process is limited to the course instructor's assessment, omitting a comprehensive evaluation of student acceptance and feedback. In contrast, the approach presented in this article extends the validation to include studies on student engagement and acceptance, providing a more holistic understanding of the educational impact and practical efficacy of integrating chatbots in academic settings.

III. MoodleBot

MoodleBot was developed by the authors as part of the *tech4compKI* project,[3] a German research project aimed at supporting personalized learning and skill development through hybrid AI mentoring. MoodleBot represents a pioneering effort in integrating an advanced LLM-driven chatbot into the RWTH academic course framework. The open-source integration[4] of the chatbot MoodleBot into Moodle targets two interrelated use cases relevant to enhancing pedagogical and operational support for academic environments. Today's academic courses often overwhelm students with a vast amount of information [62], sometimes causing them to lose track of essential organizational details [5], [63]. MoodleBot is a reference point that helps learners to find and reconcile this information quickly. Beyond this organizational assistance, the chatbot's pedagogical role is to serve as an always-available tutor [7], particularly during intense periods, such as exam preparation [63]. The primary goal of MoodleBot is to dynamically generate exercises and solve questions by deploying these chains and using Python-based functions. The bot aids students in clarifying doubts and comprehending complex lecture content. MoodleBot's response accuracy is paramount, as misinformation can significantly affect the learner's progress [64]. Concurrently, the benefits of MoodleBot extend to educators. Traditional Q&A forums, while beneficial, often become overloaded with repetitive and basic inquiries [5]. Educators can focus on more complicated, content-related questions by delegating most of these routine questions to MoodleBot. This enhances efficiency and elevates the quality of discourse within these forums. It is important to mention MoodleBot's fallibility. Traditional forums remain crucial, as they serve as channels for substantive discussions that may be beyond the scope of the chatbot and ensure a holistic educational experience.

A. Architecture

MoodleBot's design is not a generic, one-size-fits-all solution. Instead, it is a specialized, adaptive system engineered to meet a single course's specific needs and learning objectives. This approach requires an own instance of MoodleBot with its personalized agent for each course. It is fed with tailored materials, such as lecture slides, supplementary resources, and course-specific lecture notes. The architecture of MoodleBot, as illustrated in Fig. 1, integrates various tools and frameworks. OpenAI's pretrained models are utilized for optimal performance. They can be enhanced by compatible frameworks, such as LangChain[5] and LlamaIndex.[6] The following describes the core components and their functionalities:

Data Acquisition: MoodleBot can source its data from PDFs, including lecture notes, lecture slides, and exercise sheets.

Data Vectorization: Once acquired, the raw data undergoes a transformation process where it is converted into vectors. These vectors are then stored in a Weaviate database.[7]

Moodle Integration: MoodleBot seamlessly integrates with Moodle's native chat interface and forums. This ensures users experience a consistent interface while benefiting from the enhanced capabilities of MoodleBot.

LangChain Agent: MoodleBot utilizes LangChain agents equipped with modular tools to improve functionality. Initially, the chatbot uses two tools, a *Question Generator* and a *Answer Generator*. The *Question Generator* assists in formulating

[2] https://chat.openai.com/

[3] https://tech4comp.de

[4] https://github.com/rwth-acis/LMS-chatbot-service

[5] LangChain is an open-source framework for developing LLM-based applications. https://www.langchain.com/.

[6] LlamaIndex is a data framework for connecting external data to LLMs. https://www.llamaindex.ai/.

[7] Weaviate is a cloud-native vector database for AI applications to store and search vector embeddings of data. https://weaviate.io/.

relevant questions based on the content. In contrast, the *Answer Generator* provides accurate answers to user queries by tapping into the stored vectorized data. The LangChain agent is backed by advanced LLMs, specifically GPT-4 [4].

Cost and Chat History: A MongoDB[8] instance stores all associated costs and chat histories with MoodleBot to maintain transparency and aid in future optimizations.

B. Data Ingestion and Retrieval

The chatbot needs knowledge from a lecture to provide support. A Weaviate vector database is created with LangChain and LlamaIndex to take documents from various sources in various formats. A document loader transfers lecture notes to the vector database. This ensures systematic indexing of the information and its subsequent retrievability. The application utilizes LangChain's text splitter to break down longer text into manageable chunks. Each piece of text is then processed by the OpenAI Embedding model, which is accessible through LlamaIndex. This model converts the content into high-dimensional vectors, an essential step for efficient similarity search and information extraction.

C. Langchain Agent

The LangChain agent is a mediator bridging the gap between user input and the sophisticated functionalities of the LLM. Its role is to determine the actions following a user's interaction. The agent formulates a request that bundles the agent's prompt and the given user input, subsequently sending it to the LLM. The LLM then evaluates the information provided and checks whether the user's input, in conjunction with the agent's prompt, is sufficient for an immediate response. If it determines that it is sufficient, it returns a final response. If, on the other hand, it determines that additional context or data is required, the LLM responds with further instructions for the agent. Typically, these instructions cause the agent to leverage various tools to extract or derive additional context. After this information retrieval process, the LLM is invoked again with a more extensive dataset. At this point, the LLM will either generate a comprehensive response for the user or conclude that the existing context remains insufficient, thereby instructing the agent to utilize the tools again.

1) Agent Prompt: To ensure optimal results, effective prompts are needed. A prompt should clearly and concisely convey knowledge, incorporating four key components: 1) instruction; 2) context; 3) input data; and 4) output indicator [65]. Together, these components ensure that the LLM understands and processes the request to deliver optimal results. Below, more details are provided about these components' nuances and how they were integrated into the Moodle LangChain agent. The instructions should be expressed clearly and comprehensibly so that the LLM can provide accurate responses. The context, which should be brief and directly relevant to the task at hand, is shaped by the agent's role and goal. The agent is tailored to specific use cases, increasing its efficiency and relevance [65].

[8]MongoDB is a NoSQL document-oriented database management system. https://www.mongodb.com/.

A LangChain agent prompt consists of multiple components, including the agent's instruction, background context, user input, and an output indicator. It already adheres to the suggested prompt pattern. The instruction consists of the *personality*, the *goal* of the agent/tool, its *tasks* and the *limitations*. The background context is an additional source of information that can be manually inserted into the prompt or retrieved from the vector database. User input is a query submitted by the user, and the output indicator marks the beginning of the generated text. LangChain provides various prompt templates to use. The ChatPromptTemplate is tailored to the conversational context and is divided into four parts: 1) `System Prompt`; 2) `Chat History`; 3) `User Input`; and 4) `Agent Scratchpad`. The `System Prompt` section specifies the agent's personality and must be tailored to the specific case. For MoodleBot, the `System Prompt` is as follows.

Prompt 1: System Prompt

> You are a tutor for the lecture Databases and Information Systems at RWTH Aachen University. Your goal is to support the students during the lecture and to answer questions about the lecture by having a conversation with them. You can generate exercises for the students and correct their answers. Since the lecture is held in German, you should also answer in German. You can only answer questions about the lecture. You should refuse to answer any content not part of the lecture. Always be friendly, and if you cannot answer a question, admit it. In summary, the tutor is a powerful system that can help with various tasks and provide valuable insight and information on various topics. Whether you need help with a specific question or just want to have a conversation about a particular topic, Tutor is here to help.

The agent acts as a tool during a lecture, facilitating answering questions and automatically generating exercises. Therefore, the agent focuses on the lecture's content and is designed to ignore irrelevant or off-topic queries. Each sentence is tested for optimal outcomes using contextual information and a chain of thought [66]. User input is treated as an input variable, and the LLM temperature determines the randomness of the outputs. The temperature is set to zero to avoid incorrect terms in an educational context.

2) Agent Type: `BaseMultiActionAgent` was chosen as the most appropriate agent type for MoodleBot. This choice derives from the requirements that the bot should be able to use multiple tools with different functionalities and have interactive dialogues with students. Since the purpose is to chat, a chat model is necessary. During the implementation phase, two leading LLMs: 1) `gpt-3.5-turbo` and 2) `gpt-4` were tested. After manually comparing some generated answers, `gpt-4` showed superior ability in formulating answers and explaining lecture material. Unlike `gpt-3.5-turbo`, `gpt-4` can identify the need for a tool

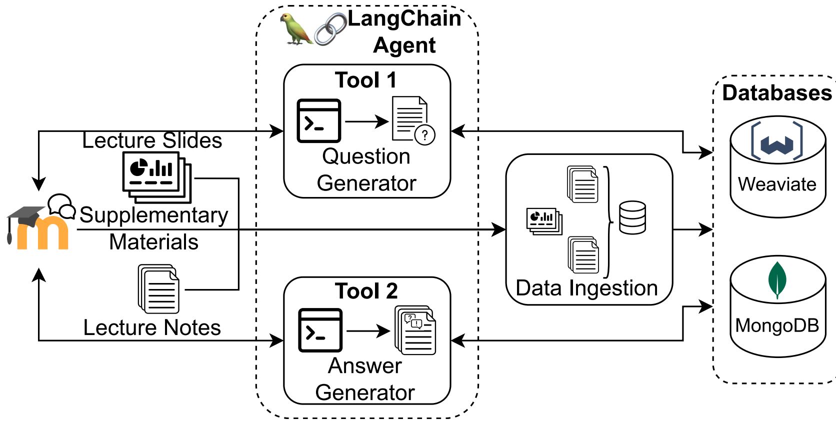

Fig. 1. System overview of MoodleBot's architecture.

with ease. Therefore, the chat model for the agent uses `gpt-4` as its LLM.

3) Tools: An agent uses tools to perform specific tasks efficiently. These tools can range from simple utilities to more complex entities, such as the composition of multiple chains.

a) Answer generation based on context: LangChain's `create-stuff-documents-chain` was adopted to generate answers in a particular lecture context. The retrieval mechanism relies on the Weaviate vector index containing embedded lecture data. The system searches the database using similarity and retrieves the most relevant documents. This search is facilitated using the same embedding model used when the database was initially set up. LangChain provides multiple chain types: *stuff*, *refine*, *reduce*, *map-reduce* and *map-rerank*. The *stuff* chain is designed to accept a list of collected documents, integrate them into a provided prompt, and then forward them to the target LLM. It is also faster in processing than its counterparts. However, to ensure both efficiency in similarity search and low costs, the number of documents fed into the process is limited to five.[9] The following custom prompt is configured to ensure that the `create-stuff-documents-chain` returns the best possible response.

Prompt 2: Answer Generation Tool

As a tutor for the lecture databases and information systems, your goal is to provide accurate and helpful information about the lecture. You should answer the user's inquiries as best as possible based on the context and chat history provided and avoid making up answers. If you do not know the answer, simply state that you do not know. Answer the question in German.

The LLM tool generates results derived from concise prompts and the context of data retrieval. It is capable of answering both organizational and content questions about lectures. The scenario depicted in Fig. 2 exemplifies the answering of an organizational question.

[9]https://python.langchain.com/docs/modules/chains/document/stuff

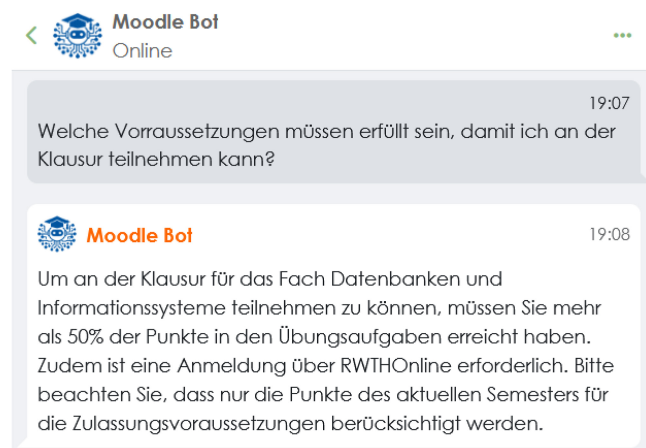

Fig. 2. User is inquiring about the process for obtaining admission to take an exam.

b) Question generation: To act like a tutor and support learners, the `QAGeneration-Chain` is used to generate rudimentary practice exercises for students. These can be useful to review the lecture content and have a similar effect as flashcards. The MongoDB database contains 170 question-answer pairs. Within this dataset, repetitions were identified, with ten pairs flagged as organizational and 20 as inaccurate. These incorrect entries were manually removed from the database. This tool is programmed to retrieve questions at a faster rate than generating them at runtime. As shown in Fig. 3, it responds with random exercise questions, which are helpful for students to learn from, and helps them comprehend the topic by providing an answer. While it ensures prompt delivery of answers, the basic nature of the generated questions could limit their applicability in advanced practice sheets or other specialized scenarios.

c) Specific question generation: The Question Generation tool can only generate random questions, and the quality of the generated questions is inferior to that produced by the `creates-stuff-documents-chain` using a customized prompt to generate specific questions. The prompt for the Specific Question Generation tool looks as follows.

Fig. 3. User has requested an exercise question, and MoodleBot is assisting with the solution.

Prompt 3: Question Generation Tool

You will assist students with practice tasks as a tutor for databases and information systems. The student will request a practice exercise on a particular subject. You will come up with a question that is pertinent to the subject. The questions should be in German.

Although the generation takes longer, the results are superior to those of the Question Generation tool mentioned above.

IV. Evaluation

In this section, the technology acceptance model (TAM) was utilized to assess how effectively the chatbot addresses the acceptance of MoodleBot as a valuable tool for learning support among students [67]. This is accomplished by examining hypotheses H1–H10 in Tables IV and V as part of answering *RQ1*. The technical evaluation focuses on output precision and a token-generation cost analysis for API calls to analyze the accuracy of MoodleBot's responses and its congruency with the established course content as stated in *RQ2*.

A. Conceptual Framework

The TAM, proposed by Davis [67], is a widely used theoretical framework for predicting users' acceptance and adoption of new technology. The individual's technology acceptance is determined by two variables, perceived usefulness (PU) and perceived ease of use (PEOU). PU is defined as the extent to which the user believes a particular system would enhance their job performance, and PEOU refers to the degree to which the user acknowledges that using a specific system would be free of effort [67]. Both affect the user's Attitude (AT) toward the technology directly and indirectly affect their behavioral intention (BI) [67], [68]. As external variables, SE refers to the individual's perception of their ability to fulfill a particular task and system accessibility (SA) to the perceived ease of accessibility and interaction with the specific system [67], [68]. TAM is a valuable tool for researchers and practitioners interested in studying students' acceptance of learning technologies [69], [70], [71]. It helps educators gain insights into improving the effectiveness of technology-based learning environments. This study assesses the users' perceptions of the usefulness and ease of use of MoodleBot. Measuring the six key areas, PU, PEOU, AT toward the MoodleBot, BI, SE, and SA, provide insights into the likelihood of adoption and potential areas for improvement to enhance users' acceptance toward LLM-driven chatbots in higher education. The validity of the constructs was assessed through an explorative factor analysis utilizing principal component analysis (PCA), reducing the dimensionality of the dataset while preserving the variance and ensuring that the constructs measured what they were intended to. The reliability of the constructs was evaluated using Cronbach's alpha, confirming the internal consistency of the measurement items.

B. Evaluation Setup

Since students may accept the answers of these tools as a general truth, leading to the propagation of misinformation [48], 46 students who already passed the course participated voluntarily in this study. The participants, ranging in age from 20 to 31, were granted access to a live system within a controlled environment and were supervised throughout the study to ensure high-quality feedback. Of these, 30 answered the questionnaire, and 28 identified as male, while one preferred not to disclose their gender. All participants are either enrolled in or have completed a computer science program. The selection criterion emphasized that students should be familiar with the chatbot's content, ensuring their capability to evaluate its content accuracy. Having undergone the lecture titled Databases and Information Systems, they possess the requisite background to evaluate the chatbot's content accuracy. While 28 of the participants had previous interactions with chatbots, only 17 have used them in an educational setting. Despite their familiarity, they were initially unaware of MoodleBot's capabilities. After a brief orientation to MoodleBot's functionalities, they navigated the chatbot interactions more confidently. Ultimately, they were asked to complete the TAM questionnaire.

C. Evaluation Results

Table I illustrates the average responses and standard deviations from participants ($N = 30$) based on a Likert scale of 1–5, with 1 being "Strongly Disagree" and 5 as "Strongly Agree." In general, participants indicated a favorable perspective toward interacting with the MoodleBot. From the PEOU category, the statement with the highest average score, $4.\overline{6}$, was $PEOU_3$, indicating participants found its operation straightforward. In terms of PU, the statement "The MoodleBot could make it easier to study course content" (PU_7) recorded the

TABLE I
RESULTS OF THE TAM QUESTIONNAIRE BASED ON PREVIOUS EDUCATIONAL CHATBOT EXPERIENCE. (N=30; LIKERT SCALE: 1 = "STRONGLY DISAGREE," 5 = "STRONGLY AGREE")

Variable	Statement	$\bar{x}_i$	σ_i
$PEOU_1$	It was easy to operate and interact with the bot.	4.5	(±0.93)
$PEOU_2$	I understood the functions of the bot well	$4.5\overline{6}$	(±0.62)
$PEOU_3$	The operation of the bot was straightforward and intuitive.	$4.\overline{6}$	(±0.54)
PU_1	The MoodleBot helped me to gain the knowledge I wanted.	$4.\overline{3}$	(±0.75)
PU_2	The MoodleBot understood my request correctly.	$4.3\overline{6}$	(±0.71)
PU_3	The answers to my questions were helpful and informative.	4.4	(±0.77)
PU_4	I was able to improve my learning performance through the MoodleBot.	$4.0\overline{6}$	(±0.82)
PU_5	The learning questions asked by the MoodleBot were always good.	$3.9\overline{3}$	(±0.94)
PU_6	The MoodleBot would increase my academic productivity.	4.1	(±0.84)
PU_7	The MoodleBot could make it easier to study course content.	$4.4\overline{6}$	(±0.73)
AT_1	Studying together with the MoodleBot is a good idea.	$4.3\overline{6}$	(±0.71)
AT_2	The bot is a good complement to traditional learning methods.	$4.7\overline{3}$	(±0.44)
AT_3	I am positive towards the MoodleBot.	4.7	(±0.53)
AT_4	I would recommend the bot to others.	$4.3\overline{6}$	(±0.61)
BI_1	I would prefer the interaction with the MoodleBot to a real tutor.	2.7	(±1.14)
BI_2	I intend to be a heavy user of the MoodleBot.	$3.\overline{3}$	(±1.09)
BI_3	I would continue to use the MoodleBot in the future as part of other courses.	$4.2\overline{3}$	(±0.85)
BI_4	I exhaust all learning materials that the particular lecture provides for me.	3.1	(±1.21)
SE_1	I feel confident finding information through the MoodleBot.	3.6	(±1.06)
SE_2	I have the necessary skills for using the MoodleBot.	$4.7\overline{6}$	(±0.43)
SE_3	I have no concerns about my data.	$3.8\overline{6}$	(±1.27)
SE_4	I was able to control the interactions with the bot well.	$4.3\overline{6}$	(±0.80)
SA_1	I have no difficulty accessing and using the Moodle Bots in a Moodle course.	$4.\overline{3}$	(±0.88)

highest mean of $4.4\overline{6}$, while "The learning questions asked by the MoodleBot were always good" (PU_5) received the lowest mean of $3.9\overline{3}$. Within the AT section, participants expressed the most positive sentiment with the statement "The bot is a good complement to traditional learning methods" (AT_2), yielding an average of $4.7\overline{3}$. When discussing BI to use MoodleBot, the statement "I would continue to use the MoodleBot in the future as part of other courses" (BI_3) garnered a high score of $4.2\overline{3}$. However, the average score for participants' preference for interacting with the MoodleBot instead of a real tutor was only 2.7 (BI_1). SE contained items with averages around the mid-range, suggesting that while participants generally felt confident using the MoodleBot and saw its societal value, there were still some reservations. Lastly, the SA category had a high mean value of $4.\overline{3}$ for SA_1, indicating that participants did not find difficulty accessing and using the MoodleBot within the Moodle course. Since no other factors affect the bot's accessibility, this single question captures the key aspect of SA and accurately reflects users' experience.

1) Reliability and Validity: The reliability of the variables gauged on a Likert scale was determined using Cronbach's alpha, and the results are presented in Table II. The Cronbach's alpha values range from 0.688 (for BI) to 0.802 (for PU) which can be described as reasonable [72]. Notably, the Cronbach's Alpha for AT was 0.800. Only the coefficients that exceed the 0.6 threshold were considered, suggesting that the questionnaire items for these coefficients present satisfactory reliability since values below are questionable [72]. Furthermore, the Kaiser–Meyer–Olkin (KMO) measure verified the good sampling adequacy for the analysis, with a KMO of 0.7041 [73]. It is noteworthy that values closer to 1 are more suitable for factor analysis. The Bartlett's test of sphericity was significant

TABLE II
CRONBACH'S ALPHA VALUES

Construct	Cronbach's Alpha
PU	0.802
AT	0.800
BI	0.688

with $\chi^2 = 227.780$, $df = 13$, and $p = 0.000$. This indicates that the correlations between items were sufficiently large for factor analysis.

Table III presents the outcomes of the explorative factor analysis, revealing the necessity to remove specific items based on a threshold of 0.6, which is acceptable for samples less than 100 [74]. Selecting components with eigenvalues greater than 1 resulted in three components with eigenvalues of 5.625, 1.72, and 1.469, explaining 67% of total variance [75]. The Cronbach's alpha coefficients for these components were 0.864, 0.806, and 0.751, suggesting satisfactory reliability. In terms of communalities, the items PU_1, PU_4, PU_6, PU_7, AT_3, and BI_3 were greater than 0.6 and the items AT_2 and BI_1 were slightly below the threshold. This indicates that most items retained in the components have a moderate degree of variance, underlining the trustworthiness of the measurement model.

2) Regression Analysis: Multiple linear regression analyses were conducted as shown in Fig. 4. This aimed to determine the strength and direction of the relationships between various internal and external variables and their impact on user behavior and perceptions. Tables IV and V summarize the tested hypotheses and their outcomes.

TABLE III
RESULTS OF THE EXPLORATIVE FACTOR ANALYSIS: LOADINGS, COMMUNALITIES, AND VARIANCES FOR VARIABLES

Variable	$\mathcal{L}_1$	$\mathcal{L}_2$	$\mathcal{L}_3$	**Communality**
PU_1		0.429	0.824	0.879
PU_3			0.694	0.492
PU_4		0.750		0.638
PU_6	0.639			0.608
PU_7	0.452	0.452	0.493	0.651
AT_1	0.624			0.532
AT_2		0.754		0.596
AT_3		0.661		0.622
AT_4	0.522			0.522
BI_1	0.747			0.596
BI_2	0.540	0.478		0.552
BI_3	0.886			0.857
BI_4				0.152
Cronbach's alpha	0.864	0.806	0.751	
Eigenvalues	5.625	1.72	1.469	
Variance accounted for (%)	43.272	13.23	11.301	

TABLE IV
RESEARCH HYPOTHESIS FOR THE INTERNAL VARIABLES

Hypothesis	Content	Result	R^2	$\sigma_{\overline{x}}$
H1	The perceived usefulness impacts the attitude towards the bot.	Hypothesis invalid.	0.282	0.140
H2	The perceived ease of use impacts the attitude towards the bot.	Hypothesis invalid.	0.150	0.152
H3	The Perceived Usefulness has an impact on the behavioral intention.	Hypothesis invalid	0.207	0.247
H4	The attitude towards the bot impacts the behavioral intention.	Hypothesis valid.	0.405	0.245

a) Internal variables: Table IV presents the internal variables and their correlation with user attitudes and behaviors. Among the variables assessed, AT and PU proved to be a significant predictor of BI, with an R^2 value of 1.0699 and a standard error of 0.245. The hypothesized relationship between PU and BI could be confirmed, as evidenced by a R^2 value of 0.6678 and a standard error of 0.247. A statistical correlation was noticed between PEOU and PU with p-values below the threshold of 0.05. Furthermore, the derived R^2 value of 0.4636 and the associated standard error ($\sigma_{\overline{x}} = 0.140$) underlines the positive effect of PU to AT. This is consistent with the expected influence of PEOU on AT as evidenced by an R^2 value of 0.3384 and an standard error $\sigma_{\overline{x}}$ of 0.152. The acceptable range for the t-value depends on both the p-value and the degrees of freedom [76].

b) External variables: In the context of the TAM, the PU and PEOU may be influenced by various external variables, including age, gender, academic degree, prior experience with educational chatbots, SE, and system accessibility, SA [77]. Due to the homogeneous nature of responses in the dataset, factors like age, gender, and academic degree were excluded from the analysis. From an in-depth examination of the data in Tablet V and as visualized in Fig. 4, it is evident that SA is a significant predictor of PEOU, registering an R^2 value of 0.3333 and a standard error of 0.095. Conversely, experience with educational chatbots showed no observable influence on either PU or PEOU. SE exhibited a positive correlation with PU (with $R^2 = 0.4052$ and $\sigma_{\overline{x}} = 0.136$) and PEOU (with $R^2 = 0.3506$ and $\sigma_{\overline{x}} = 0.141$). While SA had an impact on PEOU, it did not show a significant correlation with PU, as indicated by an R^2 value of 0.1723 and a standard error of 0.109. Interestingly, prior experience with educational chatbots provided minimal explanatory power for both PU (with $R^2 = -0.1658$ and $\sigma_{\overline{x}} = 0.196$) and PEOU (with $R^2 = -0.1220$ and $\sigma_{\overline{x}} = 0.197$).

TABLE V
RESEARCH HYPOTHESIS FOR THE EXTERNAL VARIABLES

Hypothesis	Content	Result	R^2	$\sigma_{\overline{x}}$
H5	The self-efficacy has an impact on perceived usefulness.	Hypothesis invalid.	0.241	0.136
H6	The self-efficacy has an impact on the perceived ease of use.	Hypothesis invalid.	0.181	0.141
H7	The system accessibility has an impact on the perceived usefulness.	Hypothesis valid	0.082	0.109
H8	The system accessibility has an impact on the perceived ease of use.	Hypothesis invalid.	0.307	0.095
H9	Experience with educational chatbots has an impact on the perceived usefulness.	Hypothesis invalid	0.025	0.196
H10	Experience with educational chatbots impacts the perceived ease of use.	Hypothesis invalid.	0.014	0.197

D. Correctness and Fact-Checking

Following the evaluation of the acceptance and utility of MoodleBot, the accuracy of its responses emerges as a critical factor. To scrutinize the congruence of MoodleBot's responses with the established course content, the results were manually verified by a TA and subjected to a LangChain-based fact-checker chain, the LLMSummarization-CheckerChain.

1) Correctness of Generated Outputs: To evaluate the user-perceived accuracy of MoodleBot's generated outputs, manual evaluations and student feedback were employed to assess the quality. Content answers were scrutinized for factual accuracy by the TA, while students provided feedback on each answer's utility and perceived correctness. It is important to note that feedback from students who completed the lecture over a year prior was excluded, particularly for answers they could not verify or those irrelevant to the lecture content. Of the total evaluations, a TA manually reviewed 65 responses and obtained feedback for each of them from the students. A confusion matrix, detailed in Table VI, classifies the results (including accuracy, precision, sensitivity, and specificity [78]). Using the confusion matrix, 53 out of 65 responses were accurate. However, some users were unhappy with the answers when looking at feedback from 17 chatbot responses. They said MoodleBot sometimes gave repetitive, unhelpful, or too much confusing information.

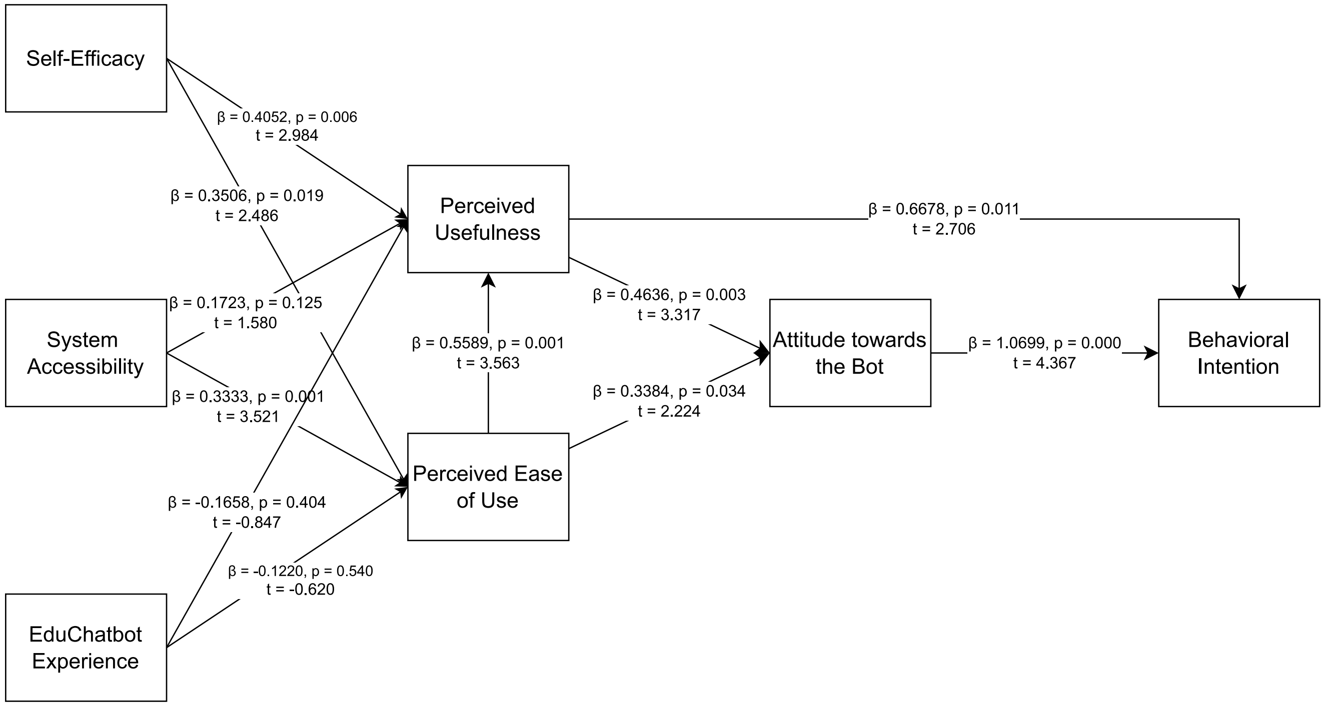

Fig. 4. Analysis result model of the hypothesis test (β = Standardized Coefficients; t = Test Statistics; p = Significance Probability (p-value)).

TABLE VI
USER-PERCEIVED ACCURACY: CONFUSION MATRIX OF MANUALLY EVALUATED CORRECTNESS OF GENERATED OUTPUTS

		User		
		Predicted Pos.	Predicted Neg.	
TA	Actually Pos.	36 (TP)	17 (FN)	67.92% (Sensitivity)
	Actually Neg.	3 (FP)	9 (TN)	75% (Specificity)
		92.31% (Precision)	34.62% (Neg. Predictive)	69.23% (Accuracy)

To assess the congruency of MoodleBot's responses, the TA conducted a thorough evaluation to verify the factual accuracy and relevance of the generated outputs. This process involved cross-referencing MoodleBot's responses with the course materials, including the syllabus and lecture notes, to ensure that the answers provided aligned with the course's intended content. The TA also examined the contextual appropriateness of MoodleBot's responses, ensuring they maintained consistency with the ongoing lecture topics and reflected the expected depth of understanding. A part of this evaluation included documenting any errors or discrepancies found in MoodleBot's responses. This documentation identifies common issues, such as the mix-up of terms like "Wertebereich (Domäne, Domain)" and "Wertedomäne," highlighting the need for more precise handling of specific terminology. In addition, student feedback played a crucial role in assessing congruency. Feedback from recent course completers was analyzed to understand areas where MoodleBot's responses might have been repetitive, overly complex, or unhelpful.

Another aspect of the evaluation involved assessing how MoodleBot handled incorrect student input. While MoodleBot sometimes missed specific errors, its feedback was still useful, as it often pointed out other mistakes, helping students learn through iterative feedback. For example, in an SQL task, MoodleBot caught a keyword error but missed an ID number mistake.

Upon examination, no further significant discrepancies or issues were found. As a result, the responses generated by MoodleBot can be inferred to be congruent with the course content. While MoodleBot is right about 81% of the time, there is still room for improvement.

2) Automated Fact-Checking: In the implementation phase, to avoid disrupting interaction flow, responses from the LMS Chatbot Service were stored in a MongoDB database for subsequent review. In addressing *RQ2*, the aim was to determine whether the quality and congruence of MoodleBot's responses could be measured automatically rather than relying solely on manual verification. Therefore, each was manually verified by a TA to ensure accuracy, with 88% of 100 responses deemed correct. Redundant answers, such as greetings and organizational queries, were excluded, resulting in approximately 160 content-specific questions. Each response was passed through a fact-checker chain to automate this process using the `gpt-3.5-turbo` model. The outputs of this fact-checker were compared against the TA's manual evaluations to assess the tool's utility and accuracy. Only the first 100 responses, which were also previously checked by the TA, underwent the fact-checker chain to maintain cost efficiency, detailed further in Section IV-E. As seen in Table VII similar to Section IV-D1, a confusion matrix was established to evaluate the bot's performance.

In the evaluation, the model achieved an accuracy of 82%, a precision of 88.04%, and high sensitivity values, suggesting proficient identification of true assertions while minimizing

TABLE VII
AUTOMATED CONGRUENCE ASSESSMENT: CONFUSION MATRIX OF TEACHING ASSISTANT VERSUS FACT-CHECKER CHAIN

		Fact-Check		
		Predicted Pos.	Predicted Neg.	
TA	Actually Pos.	81 (TP)	7 (FN)	92.05% (Sensitivity)
	Actually Neg.	11 (FP)	1 (TN)	8.33% (Specificity)
		88.04% (Precision)	12.5% (Neg. Predictive)	82% (Accuracy)

false positives. It should be noted that the same LLM handles both fact-checking and answer generation, which might bias it against detecting false statements, as evidenced by the Negative Predictive Value of 12.5%. The specificity, which measures the ratio of correctly identified negative instances to all negative instances, stands at 8%, indicating a deficiency in detecting negative instances. Moreover, the evaluation does not account for instances where the agent's generated answers deviate from lecture content, such as verbose responses to queries about briefly mentioned topics. While this could assist students in understanding broader content, it might also lead to misconceptions about exam relevance or misuse. Despite the 88% accuracy in generating correct answers, the fact-checking service falls short in identifying erroneous responses, which is crucial for a checker. Thus, the performance of the fact-checker chain combined with LLM is suboptimal, and relying on LLM for accuracy verification to produce precise responses is not recommended.

3) Summary of Fact-Checking: In relation to *RQ2*, the analysis indicates that 88 out of 100 GPT-generated responses, when measured against a predetermined confusion matrix, were accurate. Despite this significant accuracy rate, there is an evident need for refinement. Feedback from student evaluations provides valuable insight into GPT's reliability and user satisfaction. Notably, GPT exhibits limitations in identifying false assertions. The current fact-checking system, relying exclusively on LLMChain from LangChain, lacks integration with lecture slide data, potentially affecting response accuracy. In the educational domain, ensuring the integrity of information is paramount. Fact-checking enhancement could benefit from integrating LLMs trained in educational content or refining the overarching answer-generation process. Considering precision, employing multiple LLM agents and cross-referencing their outputs might optimize response accuracy.

E. Cost Calculation

OpenAI's API platform offers a suite of LLMs, each designed for specific tasks and associated with distinct pricing. The API usage fees can be calculated with the processed token amount and the selected model version.[10] Standard models can process up to 4096 tokens, which is approximately equivalent to 3000 words[11] (costs are assessed per 1000 tokens). For the cost calculation, the following equation is utilized:

$$C(n_\tau, c_m) = n_\tau * \frac{c_m}{1000}. \tag{1}$$

where C is the total cost, n_τ is the number of tokens, and c_m is the cost for the chosen model.

1) Database Setup: The `text-embedding-ada-002` model from OpenAI was employed for the Weaviate database setup. It required approximately 280 000 tokens to embed lecture slides and exercise texts

$$C_{\text{Embedding}} = C(280000, \$0.0001) = \$0.028. \tag{2}$$

2) Similarity Search Embeddings: A similarity search to retrieve information from the vector database requires embeddings for each query, which adds computational cost. During the evaluation phase, embeddings did not surpass 1000 tokens per query. Utilizing the `text-embedding-ada-002` model this results in \$0.0001 per query

$$C(1000, \$0.0001) = \$0.0001. \tag{3}$$

3) Chat: Answer generation based on context is similar to forwarding a prompt to the LLM, implying that costs involve more than just the embedding. During the evaluation, the token count per chat message ranged from 327 (319 In and 8 Out) to 71 981 (66 187 In and 5794 Out) tokens per student. Since only a limited student cohort was present, the `GPT-4` model with an 8K context was employed for the best user experience.[12] It is important to highlight that when using GPT models for conversations, both input and output tokens must be considered, which results in the following cost calculation per chat message:

$$\begin{aligned} &C(319, \$0.03) + C(8, \$0.06) \approx \$0.01 \\ &\quad \le C_{\text{message}} \\ &\quad \le C(66187, \$0.03) + C(5794, \$0.06) \approx \$2.33. \end{aligned} \tag{4}$$

The upper token bound for chat messages stems from the fact that, in such a conversational scenario, up to four prior turns (equating to 8 messages) were used to be retained for context. This history storage, the current message, and its corresponding response can lead to higher token values. Post-evaluation, the aggregate expenses were approximately \$41.15, averaging to ≈ \$1.65 per participant. The observed post-evaluation costs are notably lower than the projected range. This discrepancy can be attributed to the fact that, during the evaluation phase, most students did not fully utilize the stipulated 100-message cap with MoodleBot. Only a restricted amount of lecture questions were asked, and a specific number of exercises were generated, with a maximum of 34 conversation turns. For subsequent courses, employing a self-hosted LLM or utilizing a more cost-effective LLM, such as `GPT-3.5-turbo`, within a 16K context is recommended.

F. Discussion and Implications

This study aimed to investigate the effectiveness of a Moodle-integrated chatbot in an academic setting tailored

[10] https://openai.com/pricing
[11] https://platform.openai.com/tokenizer
[12] https://openai.com/gpt-4

explicitly for students enrolled in a course on databases and information systems. While other studies have effectively encouraged the use of similar tools by embedding them within course policies, thus indirectly increasing participation rates [52], the presented approach in this article was voluntary and contained only students who had completed the course. This decision prevented unforeseen misuse or misinformation propagation, as related studies highlighted [48], [58]. In contrast to the study by [57], this research utilized the TAM to assess student perceptions. The evaluation encompassed an in-depth questionnaire to gauge the quality, utility, and suitability of MoodleBot's responses while examining students' AT and BI. A comprehensive analysis of the quantitative feedback obtained from participants showcased high average scores across the assessed metrics, indicating a favorable acceptance of MoodleBot. As hypothesized in *RQ1*, students viewed MoodleBot as a helpful tool enhancing their educational journey. Nonetheless, certain constraints were evident. Despite MoodleBot's natural communication, students still prefer a human tutor when choosing, as reflected in the modest BI scores, indicating that chatbots should be used to complement rather than replace traditional teaching methods. Evaluation of the external variables revealed unexpected insights. While prior experience with chatbots had minimal impact on PU and PEOU, SA of the system and SE proved to be relevant factors for PEOU. This underscores the importance of easy access to LMS in shaping students' academic interactions. As outline in TAM [67], a positive correlation exists between PU, PEOU and AT. For users with technical expertise, PU and PEOU are not always the most important factors for adoption, as the significance of these factors can vary depending on the nature of the task [79]. Factors, such as SE, subjective norm, enjoyment, computer anxiety, and experience often shape students' PEOU, and PU, particularly in e-learning contexts [80]. The high acceptance rates and positive student perceptions of MoodleBot demonstrate the readiness to use LLM-based chatbot technologies. This suggests that integrating these technologies within educational platforms can increase student engagement by providing timely and accessible support outside regular instructional hours.

In efforts to affirm MoodleBot's accuracy (*RQ2*), the fact-checker chain function showed an accuracy rate of 88%, aligning with the results reported by other studies [52]. The accuracy evaluated in this study depends on both the context and usefulness of the answer, supporting the findings of similar research that focused solely on the answer's context [44]. TAs may consider the generated output more accurate and valuable. However, some students may be dissatisfied with the answer and require a more detailed explanation or description of the solution, resulting in a lower accuracy value of 69.23% in the assessment of both the TA and the students. However, it became clear that detecting inaccurate statements generated during the answer formulation requires improvement. Certain discrepancies where MoodleBot's responses differed from lecture notes underscored this requirement. Therefore, introducing a more sophisticated fact-checking mechanism may further optimize accuracy. When integrating chatbots into learning environments, educators should ensure accurate responses and contextually relevant information to maintain educational integrity. They should encourage students to verify information and critically assess the generated responses, emphasizing the role of chatbots as supplementary tools rather than definitive sources of knowledge.

Another consideration is the cumulative cost associated with using the API for MoodleBot responses increases with long-term use. The direct integration into an educational course should consider integrating a self-hosted LLM or exploring cost-effective alternatives that maintain high performance. It is noteworthy that the presented approach has demonstrated comparable cost-efficiency ($1.65 per student) to those reported by others ($1.90 per student) and is quite fair regarding costs per student [52]. A more extended observation period could offer deeper insights into users' perceptions. In addition, expanding the sample to include a broader range of genders and demographic backgrounds would provide a more holistic view of MoodleBot's utility.

Overall, this approach holds promise in addressing the shortage of specialists in higher education and providing valuable support. Both pupils and students will increasingly engage with chatbots, and the prevalence of pilot projects utilizing LLM for learning support is expected to rise.[13]

V. Limitations

These constraints imposed by course policies and German GDPR likely contributed to the lower participation rate. Consequently, the sample size was relatively small, with 46 students, out of which 30 participated in the survey, leading to a small number of responses for which feedback was received. The voluntary nature of participation may have introduced a selection bias, as those who chose to participate might have had a predisposition toward the technology, affecting the generalization of the results. Voluntary participation is a common challenge [81], yet it remains a significant limitation in drawing broad conclusions. Furthermore, as the students who have completed the course are already acquainted with the course content, they will likely possess a more profound grasp of the course material and terminology, enabling them to pose more precise and targeted questions. Conversely, students who have yet to complete the course may require assistance employing the correct terminology, which could potentially result in more general and ambiguous prompts, thus yielding less informative responses. These disparities bear substantial implications for educational chatbots' design and operational effectiveness, such as adaptive responses with learner stage recognition. Regarding feedback, the manual evaluation is done by a sole TA, which could lead to incorrect feedback for complex responses.

Another limitation of the study is using the `gpt-4` model for the chatbot implementation. The presented study does not compare its performance with other contemporary LLMs. While `gpt-4` demonstrated a high accuracy rate and provided valuable support to students, other models might offer different advantages regarding response accuracy, processing

[13]The Sabrewing Progamme. https://www.davidgamecollege.com/courses/courses-overview/item/102/gcse-ai-adaptive-learning-programme

speed, or cost-effectiveness. Future research should explore and compare various LLMs to determine the most effective and efficient options for educational chatbots.

VI. Conclusion

The pervasiveness of AI tools and associated technologies like LLMs continues to revolutionize how tutors teach and support students in their SRL and help-seeking behavior. The promise is to usher in an AI-enabled instructional environment and make serendipitous learning more enjoyable for students. However, the complicated setup of many modern LMSs in higher education, increasing student numbers, the size and variety of educational material, information overload, and the problems of recruiting competent teachers or lecturers present fundamental challenges when developing and deploying AI tools, especially in computer science education settings.

This article addressed some of these challenges by presenting an LLM-driven chatbot called MoodleBot. A discussion on the pedagogical justification for the development of MoodleBot was followed by the architectural design and the integration of MoodleBot into the Moodle LMS of RWTH Aachen University. Two research questions concerning the acceptance and efficacy of MoodleBot as a learning support tool (*RQ1*), and the accuracy of MoodleBot's responses concerning the established coursework (*RQ2*) were addressed.

Regarding *RQ1*, six measures of the TAM were used to evaluate the MoodleBot. These measures are PU, PEOU, users' Attitude (AT), BI, SE, and SA. The evaluation shed light on the beneficial impact of MoodleBot on the student's learning process, with the chatbot receiving favorable PU and PEOU feedback from the students. These results confirm MoodleBot's potential to enhance learning by providing immediate assistance and enabling SRL.

However, despite the natural-language communication offered by MoodleBot, students still preferred a human tutor, indicating room for enhancing MoodleBot's human-like interaction and tutoring capability. The results also emphasize the positive influence of LLMs on students' academic engagement. The overall student acceptance of MoodleBot signifies the potential of AI-driven chatbots in educational settings.

As for *RQ2*, while MoodleBot demonstrated an accuracy rate of 88% when using the fact-checker function and proven congruent to the course materials by TAs, it is challenging for MoodleBot to detect inaccurate statements generated during the answer formulation process.

Despite the limitations (course or teaching policies, small number and voluntary nature of participants, etc.) enumerated below, the research can conclude with a caution that LLMs in general, and LLM-driven chatbots in particular, can effectively be used to personalize and support students throughout their learning journeys and foster SRL.

For educators, including but not limited to electrical and electronic engineering, computer engineering, and computer science, AI tools, such as LLMs can reduce administrative burdens and provide supplementary assistance, allowing more focus on instruction. Educators should use AI tools or chatbots to complement, not replace, human instruction. Continuous monitoring and evaluation of AI-enabled teaching and learning technologies, such as the MoodleBot presented in this research has the potential to impact students' learning and engagement. Institutions can benefit from LLM-driven chatbots' scalable and cost-effective nature to make quality education more accessible and efficient.

VII. Future Work

However, future research could still optimize the performance and usability of MoodleBot. This includes improving its fact-checking ability and accuracy of responses and exploring other (self-hosted) LLMs to reduce its reliance on (expensive) proprietary APIs. Furthermore, employing a more extensive and diverse sample of users for evaluation would offer robust insights into MoodleBot's utility and acceptance. As with any pioneering technology, the ongoing refinement of these educational tools, as informed by qualitative and quantitative assessments, will be crucial in maximizing their potential value to learners and educators. In conclusion, LLM-powered educational chatbots, such as MoodleBot, promise to revolutionize traditional learning paradigms, equipping learners with accessible, instantaneous, and personalized learning support. Continuous improvements in AI capabilities will lead to even more effective and reliable educational chatbots, further integrating them into everyday learning and teaching processes.

Acknowledgment

The authors would like to express their sincere gratitude to the editor, associate editor, and reviewers for their valuable comments and constructive suggestions, which have significantly improved the quality of this work.

References

[1] A. Kerly, P. Hall, and S. Bull, "Bringing chatbots into education: Towards natural language negotiation of open learner models," *Knowl.-Based Syst.*, vol. 20, no. 2, pp. 177–185, 2007.

[2] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are we there yet?-A systematic literature review on chatbots in education," *Front. Artif. Intell.*, vol. 4, Jul. 2021, Art. no. 654924.

[3] R. Dale, "The return of the chatbots," *Natural Lang. Eng.*, vol. 22, no. 05, pp. 811–817, 2016.

[4] "GPT-4 technical report." OpenAI. 2023. [Online]. Available: https://cdn.openai.com/papers/gpt-4.pdf

[5] A. K. Goel and L. Polepeddi, *Jill Watson: A Virtual Teaching Assistant for Online Education*, Georgia Tech Library, Atlanta, GA, USA, 2016.

[6] A. T. Neumann, P. de Lange, R. Klamma, N. Pengel, and T. Arndt, "Intelligent mentoring bots in learning management systems: Concepts, realizations and evaluations," in *Proc. Int. Symp. Emerg. Technol. Educ.*, 2021, pp. 3–14.

[7] A. T. Neumann, A. D. Conrardy, and R. Klamma, "Supplemental mobile learner support through moodle-independent assessment bots," in *Proc. Int. Conf. Web-Based Learn.*, 2021, pp. 75–89.

[8] S. Ruan et al., "QuizBot: A dialogue-based adaptive learning system for factual knowledge," in *Proc. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, 2019, pp. 1–13.

[9] P. Smutny and P. Schreiberova, "Chatbots for learning: A review of educational chatbots for the Facebook messenger," *Comput. Educ.*, vol. 151, pp. 1–11, Jul. 2020.

[10] R. M. V. Wolff, J. Nörtemann, S. Hobert, and M. Schumann, "Chatbots for the information acquisition at universities–a student's view on the application area," in *Proc. Int. Workshop Chatbot Res. Design*, 2020, pp. 231–244.

[11] B. Vijayakumar, S. Höhn, and C. Schommer, "Quizbot: Exploring formative feedback with conversational interfaces," in *Technology Enhanced Assessment* (Communications in Computer and Information Science 1014), S. Draaijer, D. Joosten-ten Brinke, and E. Ras, Eds. Cham: Springer, 2019, pp. 102–120. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-25264-9_8

[12] D. Pérez-Marín, "A review of the practical applications of pedagogic conversational agents to be used in school and university classrooms," *Digital*, vol. 1, no. 1, pp. 18–33, 2021.

[13] R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," in *Proc. Acad. Manag. Annu. Meet. (AOM)*, 2018, Art. no. 15903.

[14] S. Yang and K. Stansfield, "AI Chatbot for educational service improvement in the post-pandemic era: A case study prototype for supporting digital reading list," in *Proc. 13th Int. Conf. E-Educ., E-Bus., E-Manage., E-Learn. (IC4E)*, New York, NY, USA, 2022, pp. 24–29.

[15] A. M. Latham, K. A. Crockett, D. A. McLean, B. Edmonds, and K. O'Shea, "Oscar: An intelligent conversational agent tutor to estimate learning styles," in *Proc. Int. Conf. Fuzzy Syst.*, 2010, pp. 1–8.

[16] S. Tegos, S. Demetriadis, and A. Karakostas, "MentorChat: Introducing a configurable conversational agent as a tool for adaptive online collaboration support," in *Proc. 15th Panhellenic Conf. Inform.*, 2011, pp. 13–17.

[17] K. Gaglo, B. M. Degboe, G. M. Kossingou, and S. Ouya, "Proposal of conversational chatbots for educational remediation in the context of COVID-19," in *Proc. 23rd Int. Conf. Adv. Commun. Technol. (ICACT)*, 2021, pp. 354–358.

[18] J. Maldonado-Mahauad, M. Pérez-Sanagustín, J. Carvallo-Vega, E. Narvaez, and M. Calle, "Miranda: A chatbot for supporting self-regulated learning," in *Proc. Eur. Conf. Technol. Enhanced Learn.*, 2022, pp. 455–462.

[19] A. Bozkurt, W. Kilgore, and M. Crosslin, "Bot-teachers in hybrid massive open online courses (MOOCs): A post-humanist experience," *Aust. J. Educ. Technol.*, vol. 34, no. 3, pp. 39–59, 2018.

[20] W. Boles, L. Trouton, and E. Margerison, "A review of chatbots in education: Practical steps forward," in *Proc. 30th Annu. Conf. Aust. Assoc. Eng. Educ. (AAEE)*, 2019, pp. 299–306.

[21] G. E. Acuna, L. A. Alvarez, J. Miraflores, and M. J. Samonte, "Towards the development of an adaptive E-learning system with chatbot using personalized E-learning model," in *Proc. 7th Int. Conf. Front. Educ. Technol.*, New York, NY, USA, 2021, pp. 120–125.

[22] M. Coronado, C. A. Iglesias, Á. Carrera, and A. Mardomingo, "A cognitive assistant for learning java featuring social dialogue," *Int. J. Human-Comput. Stud.*, vol. 117, pp. 55–67, 2018.

[23] T. Sitzmann and K. Ely, "A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go," *Psychol. Bull.*, vol. 137, no. 3, pp. 421–442, 2011.

[24] M. Boekaerts, "Self-regulated learning: Where we are today," *Int. J. Educ. Res.*, vol. 31, no. 6, pp. 445–457, 1999.

[25] B. J. Zimmerman, "Self-regulated learning and academic achievement: An overview," *Educ. Psychol.*, vol. 25, no. 1, pp. 3–17, 1990.

[26] S. G. Paris and A. H. Paris, "Classroom applications of research on self-regulated learning," *Educ. Psychol.*, vol. 36, no. 2, pp. 89–101, 2001.

[27] E. H.-K. Wu, C.-H. Lin, Y.-Y. Ou, C.-Z. Liu, W.-K. Wang, and C.-Y. Chao, "Advantages and constraints of a hybrid model K-12 E-learning assistant chatbot," *IEEE Access*, vol. 8, pp. 77788–77801, 2020.

[28] T. K. Chiu, B. L. Moorhouse, C. S. Chai, and M. Ismailov, "Teacher support and student motivation to learn with artificial intelligence (AI) based chatbot," *Interact. Learn. Environ.*, vol. 32, no. 7, pp. 1–17, 2023.

[29] D. H. Chang, M. P. C. Lin, S. Hajian, and Q. Q. Wang, "Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization," *Sustainability*, vol. 15, no. 17, pp. 1–15, 2023.

[30] J. Du, W. Huang, and K. F. Hew, "Supporting students goal setting process using chatbot: Implementation in a fully online course," in *Proc. IEEE Int. Conf. Eng., Technol. Educ. (TALE)*, 2021, pp. 35–41.

[31] G. R. M. Prondoza and J. F. D. Panoy, "Development of chatbot supplementary tool in science and the self-regulated learning skills among the grade 10 students," *Asia Pac. J. Adv. Educ. Technol.*, vol. 1, pp. 107–116, Sep. 2022.

[32] T.-T. Wu, H. Y. Lee, P. H. Li, C. N. Huang, and Y. M. Huang, "Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid," *J. Educ. Comput. Res.*, vol. 61, no. 8, pp. 3–31, 2024.

[33] P. Prasad and A. Sane, "A self-regulated learning framework using generative AI and its application in CS educational intervention design," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 1070–1076.

[34] V. Aleven, B. Mclaren, I. Roll, and K. Koedinger, "Toward Meta-cognitive tutoring: A model of help seeking with a cognitive tutor," *Int. J. Artif. Intell. Educ.*, vol. 16, no. 2, pp. 101–128, 2006.

[35] S. Järvelä, "How does help seeking help?–New prospects in a variety of contexts," *Learn. Instruct.*, vol. 21, no. 2, pp. 297–299, 2011.

[36] A. M. Ryan, P. R. Pintrich, and C. Midgley, "Avoiding seeking help in the classroom: Who and why?" *Educ. Psychol. Rev.*, vol. 13, no. 2, pp. 93–114, 2001.

[37] V. Almeda, R. Baker, and A. Corbett, "Help avoidance: When students should seek help, and the consequences of failing to do so," *Teachers College Rec.*, vol. 119, no. 3, pp. 1–24, 2017.

[38] I. Roll, R. S. J. Baker, V. Aleven, and K. R. Koedinger, "On the benefits of seeking (and avoiding) help in online problem-solving environments," *J. Learn. Sci.*, vol. 23, no. 4, pp. 537–560, 2014.

[39] N. M. Webb and A. Mastergeorge, "Promoting effective helping behavior in peer-directed groups," *Int. J. Educ. Res.*, vol. 39, nos. 1–2, pp. 73–97, 2003.

[40] K. Wang and R. Lawrence, "HelpMe: Student help seeking using office hours and Email," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 1388–1394.

[41] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[42] T. B. Brown et al., "Language models are few-shot learners," 2020, *arXiv:2005.14165*.

[43] D. Cambaz and X. Zhang, "Use of AI-driven code generation models in teaching and learning programming: A systematic literature review," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 172–178.

[44] K. Wang, J. Ramos, and R. Lawrence, "ChatEd: A Chatbot leveraging ChatGPT for an enhanced learning experience in higher education," 2023, *arXiv:2401.00052*.

[45] M. Liu and F. M'Hiri, "Beyond traditional teaching: Large language models as simulated teaching assistants in computer science," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 743–749.

[46] G. Pinto, I. Cardoso-Pereira, D. Monteiro, D. Lucena, A. Souza, and K. Gama, "Large language models for education: Grading open-ended questions using ChatGPT," in *Proc. 37th Brazil. Symp. Softw. Eng.*, New York, NY, USA, 2023, pp. 293–302.

[47] H. Alkaissi and S. I. McFarlane, "Artificial hallucinations in ChatGPT: Implications in scientific writing," *Cureus*, vol. 15, no. 2, 2023, Art. no. e35179.

[48] K. Malinka, M. Perešíni, A. Firc, O. Hujnák, and F. Januš, "On the educational impact of ChatGPT: Is artificial intelligence ready to obtain a university degree?" in *Proc. Conf. Innov. Technol. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2023, pp. 47–53.

[49] C. C. Tossell, N. L. Tenhundfeld, A. Momen, K. Cooley, and E. J. de Visser, "Student perceptions of ChatGPT use in a college essay assignment: Implications for learning, grading, and trust in artificial intelligence," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1069–1081, Jan. 2024.

[50] T. Phung et al., "Automating human tutor-style programming feedback: Leveraging GPT-4 tutor model for hint generation and GPT-3.5 student model for hint validation," in *Proc. 14th Learn. Anal. Knowl. Conf.*, New York, NY, USA, 2024, pp. 12–23.

[51] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, Red Hook, NY, USA, 2020.

[52] R. Liu et al., "Teaching CS50 with AI: Leveraging generative artificial intelligence in computer science education," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 750–756.

[53] M. Daun and J. Brings, "How ChatGPT will change software engineering education," in *Proc. Conf. Innov. Technol. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2023, pp. 110–116.

[54] Y. Niu and H. Xue, "Exercise generation and student cognitive ability research based on ChatGPT and Rasch model," *IEEE Access*, vol. 11, pp. 116695–116705, 2023.

[55] E. L. Ouh, B. K. S. Gan, K. J. Shim, and S. Wlodkowski, "ChatGPT, can you generate solutions for my coding exercises? An evaluation on its effectiveness in an undergraduate java programming course," in *Proc. Conf. Innov. Technol. Comput. Sci. Educ.*, New York, NY, USA, 2023, pp. 54–60.

[56] S. Sarsa, P. Denny, A. Hellas, and J. Leinonen, "Automatic generation of programming exercises and code explanations using large language models," in *Proc. ACM Conf. Int. Comput. Educ. Res.*, New York, NY, USA, 2022, pp. 27–43.

[57] C. Xiao, S. X. Xu, K. Zhang, Y. Wang, and L. Xia, "Evaluating reading comprehension exercises generated by LLMs: A showcase of ChatGPT in education applications," in *Proc. 18th Workshop Innovat. Use NLP Build. Educ. Appl. (BEA)*, 2023, pp. 610–625.

[58] M. P. Rogers, H. M. Hillberg, and C. L. Groves, "Attitudes towards the use (and misuse) of ChatGPT: A preliminary study," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 1147–1153.

[59] G. A. Katuka et al., "Integrating natural language processing in middle school science classrooms: An experience report," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 639–645.

[60] J. Leinonen et al., "Using large language models to enhance programming error messages," in *Proc. 54th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2023, pp. 563–569.

[61] E. Al-Hossami, R. Bunescu, J. Smith, and R. Teehan, "Can language models employ the socratic method? Experiments with code debugging," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ. V. 1*, New York, NY, USA, 2024, pp. 53–59.

[62] M. Dibitonto, K. Leszczynska, F. Tazzi, and C. M. Medaglia, "Chatbot in a campus environment: Design of LiSA, a virtual assistant to help students in their university life," in *Proc. Int. Conf. Human-Comput. Interact.*, 2018, pp. 103–116.

[63] A. T. Neumann et al., "Chatbots as a tool to scale mentoring processes: Individually supporting self-study in higher education," *Front. Artif. Intell.*, vol. 4, pp. 64–71, May 2021.

[64] J. Osborne and D. Pimentel, "Science, misinformation, and the role of education," *Science*, vol. 378, no. 6617, pp. 246–248, 2022.

[65] J. White et al., "A prompt pattern catalog to enhance prompt engineering with ChatGPT," 2023, *arXiv:2302.11382*.

[66] J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 24824–24837.

[67] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quart.*, vol. 13, no. 3, pp. 319–340, 1989.

[68] Ö. F. Ursavaş, "Technology acceptance model: History, theory, and application," in *Conducting Technology Acceptance Research in Education*. Cham, Switzerland: Springer, 2022, pp. 57–91. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-10846-4_4

[69] F. An, L. Xi, and J. Yu, "The relationship between technology acceptance and self-regulated learning: The mediation roles of intrinsic motivation and learning engagement," *Educ. Inf. Technol.*, vol. 29, no. 3, pp. 1–19, 2023.

[70] G. Liu and C. Ma, "Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model," *Innov. Lang. Learn. Teach.*, vol. 18, no. 2, pp. 125–138, 2024.

[71] Y. Lin and Z. Yu, "Extending technology acceptance model to higher-education students' use of digital academic reading tools on computers," *Int. J. Educ. Technol. Higher Educ.*, vol. 20, no. 1, pp. 1–24, 2023.

[72] K. S. Taber, "The use of Cronbach's alpha when developing and reporting research instruments in science education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, 2018.

[73] M. Naseer et al., "Critical issues at the upstream level in sustainable supply chain management of Agri-food industries: Evidence from Pakistan's citrus industry," *Sustainability*, vol. 11, no. 5, p. 1326, 2019.

[74] R. C. MacCallum, K. F. Widaman, S. Zhang, and S. Hong, "Sample size in factor analysis," *Psychol. Methods*, vol. 4, no. 1, pp. 84–99, 1999.

[75] M. Tavakol and A. Wetzel, "Factor analysis: A means for theory and instrument development in support of construct validity," *Int. J. Med. Educ.*, vol. 11, pp. 245–247, Nov. 2020.

[76] X. Tian, C. Li, and Y. Zhao, "Investigation on computational thinking of normal students based on technology acceptance model," in *Proc. 6th Int. Conf. Distance Educ. Learn.*, New York, NY, USA, 2021, pp. 303–308.

[77] I. A. Castiblanco Jimenez et al., "Commonly used external TAM variables in e-learning, agriculture and virtual reality applications," *Future Internet*, vol. 13, no. 1, p. 7, 2021.

[78] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, Informedness, markedness & correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.

[79] D. Gefen and D. W. Straub, "The relative importance of perceived ease of use in IS adoption: A study of E-commerce adoption," *J. Assoc. Inf. Syst.*, vol. 1, no. 1, pp. 1–30, 2000.

[80] F. Abdullah and R. Ward, "Developing a general extended technology acceptance model for E-learning (GETAMEL) by analysing commonly used external factors," *Comput. Human Behav.*, vol. 56, pp. 238–256, Mar. 2016.

[81] J. Hu, T. Zhang, H. Wang, Z. Chen, and L. Liu, "Intention patterns predicting college students' volunteer service participation," *Heliyon*, vol. 9, no. 11, 2023, Art. no. e21897.

Alexander Tobias Neumann received the master's degree in computer science from RWTH Aachen University, Aachen, Germany, in 2018, where he is currently pursuing the Ph.D. degree with the Advanced Community Information Systems Group, Information Systems and Databases Chair.

He was a Research Assistant with RWTH Aachen University. His research focuses on distributed architectures, chatbots, and scaling up digital mentoring processes.

Yue Yin received the bachelor's degree in computer science from RWTH Aachen University, Aachen, Germany, in 2023.

She has been a Member of the Advanced Community Information Systems group at the Chair of Information Systems and Databases, RWTH Aachen University, since 2023. Her research interests include social bots, machine learning, and legal tech.

Sulayman Sowe received the Ph.D. (summa cum laude) degree in informatics from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2007, the M.Sc. degree in computer science from Sichuan University, Chengdu, China, in 1997, the B.Ed. degree in science (physics) education from Bristol University, Bristol, U.K., in 1991, and the Higher Teachers Certificate in physics and chemistry from The Gambia College, Brikama, The Gambia, in 1988.

He is a Research Associate/Assistant with the Faculty of Computer Science 5, RWTH Aachen University, Aachen, Germany. His scholarly publications include two edited books: *Emerging Free and Open Source Software Practices*, IGI Global, 2008, and *Free and Open Source Software and Technology for Sustainable Development*, UNU Press, 2012.

Dr. Sowe is an Academic Reviewer for several scientific journals and a Program Committee Member for multiple international organizations. He received the Higher Teachers Certificate in Physics and Chemistry from The Gambia College (The Gambia). His international work experience in education, the IT Industry, and think-tank institutions took him to several countries in Africa, Asia, Europe, and North America.

Stefan Decker received the Diplom-Informatiker degree (equivalent to master's degree) in computer science from the University of Kaiserslautern, Kaiserslautern, Germany, in 1995, and the Ph.D. degree in computer science from the University of Karlsruhe (currently, Karlsruhe Institute of Technology), Karlsruhe, Germany, in 2002.

He is a Full Professor Heading the Chair of Databases and Information Systems, RWTH Aachen University, Aachen, Germany. Further, he is the Director of the Fraunhofer Institute for Applied Information Technology, Sankt Augustin, Germany. He was formerly the Director of the Insight Centre for Data Analytics, and a Professor of Informatics with the University of Galway, Galway, Ireland. Before that, he was an Assistant Professor with the University of Southern California, Los Angeles, CA, USA, and as a Postdoctoral Researcher with Stanford University, Stanford, CA, USA. His research interests include semantic Web, metadata, ontologies and semistructured data, Web services, applications for digital libraries, knowledge management, information integration, and peer-to-peer technology.

Matthias Jarke (Life Senior Member, IEEE) received the Dipl.-Kfm. degree (equivalent to MBA), the Dipl.-Inform. degree (equivalent to M.Sc.) in computer science, and the Dr. rer. pol. degree (Ph.D.) in business administration from Hamburg University, Hamburg, Germany, in 1977, 1979, and 1980, respectively.

He was a Professor Emeritus of Databases and Information Systems with RWTH Aachen University, Aachen, Germany, in 2024, and a Past Director of the Fraunhofer FIT Institute for Applied IT. His research was focused on information systems support for cooperative tasks in engineering, business, and culture. Major contributions included query optimization, conceptual modeling, and requirements management.

Dr. Jarke was the recipient of the 2020 Peter Chen Award. He served as a Chief Editor of Information Systems and on the Editorial Board of numerous journals, including IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, and a Program Chair of prestigious database conferences, such as VLDB, EDBT, CAiSE, and SSDBM. He was an ACM Fellow, a GI Fellow, and a member of Germany's acatech National Academy of Science and Engineering.