2025 3rd International Conference on Foundation and Large Language Models (FLLM) | 979-8-3315-9409-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/FLLM67465.2025.11390928

Cipherbot: An AI-Powered Educational Assistant for Conversational Q&A About Course Material

Jinan Azem[1], Soon-Gyo Jung[1], Johanne Medina[1], Kholoud K. Aldous[1], Joni Salminen[2], Bernard J Jansen[1]
[1] *Qatar Computing Research Institute, Hamad Bin Khalifa University*, Doha, Qatar
Email: jazem@hbku.edu.qa, sjung@hbku.edu.qa, khulood.k.2011@gmail.com, jjansen@acm.org
[2] *School of Marketing and Communication, University of Vaasa*, *Vaasa*, Finland
Email: jonisalm@uwasa.fi

Education is one of the areas where artificial intelligence (AI) is having a dramatic impact. In this paper, we present Cipherbot, an AI-powered educational assistant designed to deliver scalable, personalized learning experiences through conversational interfaces as an exemplar for the design of such systems. This paper presents the technical architecture, development methodology, and dual-interface functionality of Cipherbot for educators and students. We evaluate Cipherbot in an eight-week deployment with sixteen undergraduate participants. Cipherbot demonstrated strong performance in enhancing educational interactions, achieving a 95.8% task completion rate. Usability assessments yielded high satisfaction scores, 78% on the System Usability Scale and 76% on the Chatbot Usability Questionnaire. Despite these positive outcomes, a gradual decline in student engagement over time highlights the need for dynamic and immersive learning strategies. Findings emphasize the critical role of sustaining engagement and enhancing learning outcomes within emerging educational technologies that rely on AI.

Keywords—LLMs, Chatbot, Dialogue learning, Q&A

I. Introduction

Education faces critical challenges in scaling instructional resources to meet rising demands for high-quality learning. Disproportionate student-to-educator ratios, especially in online settings, where one instructor may oversee thousands of learners [1], [2], make personalized support increasingly inaccessible [3]. This lack of individualized attention contributes to poor academic performance, increased dropout rates, and reduced learner satisfaction [3], [4], [5]. Personalized guidance typically occurs through question-answering exchanges between educators and students, a practice that improves problem-solving skills and knowledge retention [6]. However, many students hesitate to ask questions publicly due to fear of judgment from peers or instructors [7], [8]. When they do, learners often encounter delays in receiving answers [8], [9], while instructors are often swamped by repetitive queries [9]. These dynamics highlight the need for innovative, scalable solutions, such as AI-driven educational assistants [10] that can simulate responsive, personalized interactions at scale [11], ultimately improving learning outcomes and sustained student engagement.

To meet the growing demand for scalable and personalized student support, educational research has increasingly explored AI-driven conversational agents [12], [13], particularly for managing student queries [14], [15], [16]. These systems leverage advances in natural language processing to enable rich human computer interaction (HCI) [17], [18] and are rapidly becoming an integral part of the educational technology ecosystem. The global AI educational market is projected to reach $1.23 billion by 2025, reaching $32 billion by 2030 [19], [20]. However, general-purpose large language model (LLM) chatbots such as ChatGPT raise concerns due to reliance on opaque training data, which can introduce inaccuracies and compromise the credibility of educational responses [21], [22]. In contrast, domain-specific educational chatbots restrict responses to instructor-approved content, thereby mitigating misinformation risks and enhancing reliability in learning contexts. As such systems gain adoption, rigorous evaluation is critical to assess their ability to support sustained engagement, uphold academic standards, support access, offer usability, and deliver effective, personalized learning experiences at scale.

To address the dual challenges of scalable personalization and factual accuracy in AI-mediated education, we present *Cipherbot,* an instructor-grounded AI educational assistance system designed for question-answering within course-specific contexts for students and actionable analytics for both students and teachers. Cipherbot (a play on the Arabic word for zero, صفر, and pronounced “say-fur"-bot) operates exclusively on educator-curated content [23], generating context-aware responses that align directly with approved instructional content [24]. Through a web-based interface, students engage in natural language dialogue centered on their coursework, while educators use a streamlined content management system to upload, structure, and annotate course materials. Unlike general-purpose chatbots, Cipherbot limits its responses to verified sources, reducing the risk of misinformation and enhancing trust in educational interactions [1]. The platform also supports metadata tagging and structured content organization, offering capabilities beyond standard LLM systems like ChatGPT (see Fig. 1).

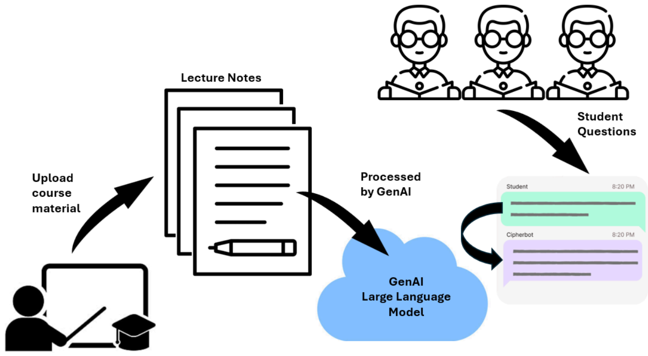


Fig. 1. The system overview of Cipherbot. Teachers upload course material to Cipherbot. Students ask questions or seek learning assistance via dialogues, and Cipherbot answers based on the uploaded material.

This targeted and bounded architecture enables Cipherbot to deliver reliable, course-aligned support at scale. In this paper, we describe Cipherbot’s system design, pedagogical motivations, and interaction model, illustrating its potential to advance personalized learning rooted in factual accuracy.

While domain-specific AI educational assistants have been explored across domains such as healthcare, customer service, and education, Cipherbot contributes a distinctive configuration to the educational AI landscape. Prior systems frequently rely on rule-based logic or tightly constrained domains [25], limiting adaptability and scalability, or general-purpose LLM models, leading to unvetted content. In contrast, Cipherbot employs a Retrieval-Augmented Generation (RAG) framework [26] paired with a LLM to generate responses and conversations grounded exclusively in educator-vetted content. This approach enables broad deployment across varied course contexts without requiring handcrafted rules or domain-specific fine-tuning. Beyond its response generation capabilities, Cipherbot integrates automated question categorization, response accuracy classification, and detailed engagement analytics [23], providing instructors (and students) with pedagogically relevant insights into student interaction patterns. Together, these features position Cipherbot not merely as a content-aware assistant but as an instructor-controlled, curriculum-aligned companion designed to support scalable, personalized learning across diverse educational settings.

Exploring Cipherbot, Section 2 reviews key related work on conversational agents. Section 3 discusses Cipherbot's design principles of LLM and RAG integration. Section 4 presents the results of an eight-week user study with sixteen undergraduate students, highlighting the positives of AI assistants in the classroom and potential challenges. We end with future work, especially creating sustained engagement.

II. Related Work

Conversational agents (CAs) are, with the rise of LLMs, AI systems designed to enable natural language interaction between users and machines [27], [28], [29]. These include both general-purpose platforms, such as ChatGPT, and domain-specific solutions like Cipherbot, which are tailored to specialized use cases. The integration of LLMs into education represents a rapidly expanding research frontier, with ongoing exploration into how best to harness their capabilities [28], [30], [31], [32] for learning support [33]. LLMs offer significant potential to personalize learning by rapidly contextualizing large bodies of educational content, enhancing information retrieval [34], [35], and recommending resources aligned with learners' goals and progress [36], [37]. Prior studies have further demonstrated the utility of AI in refining the cognitive structure of educational questions, facilitating deeper engagement with course material [38], [39].

Cipherbot serves as an alternative to both general-purpose RAG-based systems like Phind and domain-specific educational agents such as Khanmigo. While tools like Khanmigo integrate curated content with conversational AI for personalized learning, they often lack transparency in response sourcing, offer limited instructor control over the content, and provide minimal insight into student interaction patterns. Cipherbot addresses these limitations by enabling educators to upload and organize their course materials, ensuring that the AI assistant's responses are grounded in instructor-approved content [23]. Additionally, Cipherbot supports fine-grained categorization of student queries and delivers engagement analytics tailored for pedagogical reflection. These capabilities offer educators a more transparent and controllable interaction environment.

General-purpose LLMs like ChatGPT have seen widespread use in educational contexts, but persistent concerns about factual accuracy and alignment with course-specific content have been well-documented in both research and practice [21], [40]. While commercial platforms increasingly embed LLMs into learning environments, these systems often rely on unknown training data and lack mechanisms for instructor-level content control. Cipherbot addresses these gaps by offering an educator-configurable, domain-specific alternative that constrains responses strictly to uploaded, instructor-verified materials. This design ensures pedagogical alignment, content transparency, and traceability in student interactions. In addition, Cipherbot provides granular analytics on student engagement [23]and query types, features typically absent in general-purpose systems. Rather than competing with broad-use chatbots, Cipherbot complements them as a purpose-built solution for structured, curriculum-aligned support.

In sum, Cipherbot exemplifies the growing role of LLMs in reshaping education through personalized, context-aware interactions between learners and intelligent systems.

III. Design Principles of Cipherbot

A. Design Principles of Cipherbot

Cipherbot is an educational conversational agent integrating a RAG framework [41] with structured prompt engineering to produce accurate, context-sensitive responses. Powered by GPT-4o-mini and accessed via API to support future model upgrades, Cipherbot retrieves information exclusively from educator-uploaded content to ensure precision and reliability in its outputs [42], [43]. This architecture grounds all responses in verified instructional material, enhancing trust and transparency in AI-mediated learning. In contrast to general-purpose chatbots, Cipherbot is purposefully constrained to instructor-defined content, enabling course-specific, pedagogically aligned interactions that foster personalized and meaningful student engagement. By restricting outputs to educator-uploaded content, Cipherbot significantly reduces the risk of misinformation [21], ensuring that responses are accurate, transparent, and pedagogically aligned with the instructional context.

Cipherbot offers educators an intuitive interface for uploading and organizing diverse instructional materials, including lecture notes, slides, book chapters, and academic articles, enabling high-precision question answering grounded in course-specific materials. Once content is added, the underlying LLM retrieves relevant segments to generate accurate, contextually grounded responses to student queries. Educators can enrich materials with metadata, such as lecture dates and thematic tags, to enhance structure and navigability for themselves and their students. Classes can be created and shared via invitation links or direct email enrollment, and the system supports continuous updates to accommodate evolving course content. Upon joining, students immediately access a conversational interface that enables question-answering interactions strictly grounded in instructor-curated resources [44].

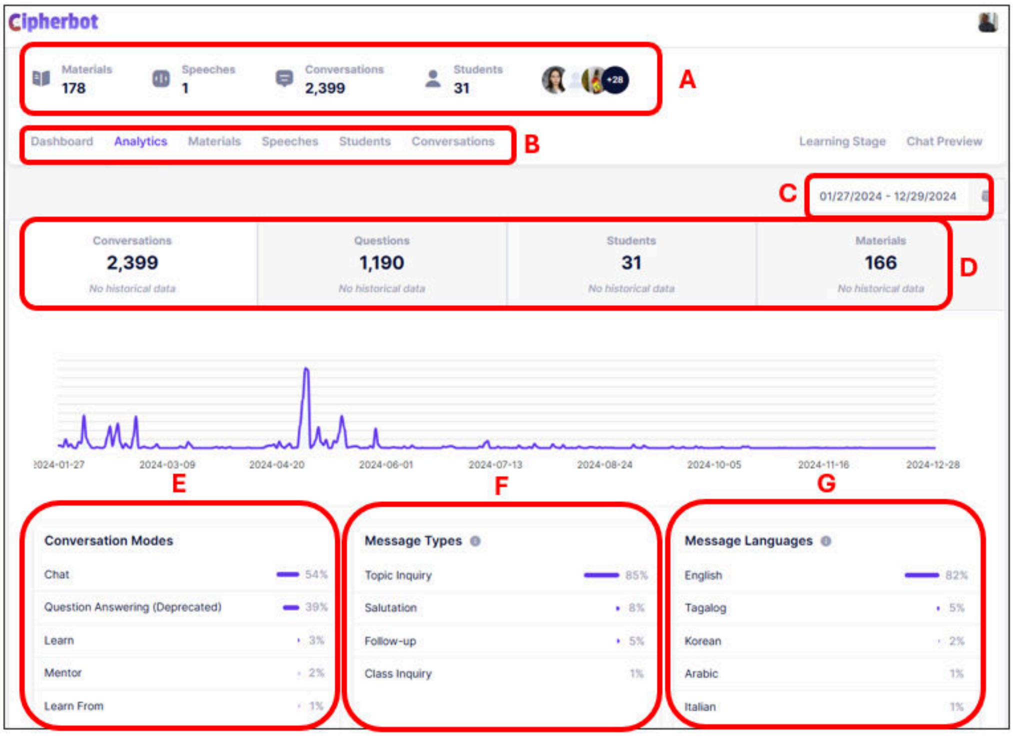


Fig. 2. Educators' interface where teachers can view existing class information, including **A**-course overview, **B**-instructor tab, **C**-reporting period, **D**-student engagement metrics, **E**-dialogue modes, **F**-message types, and **G**-language used, among many other instructor features and analytics.

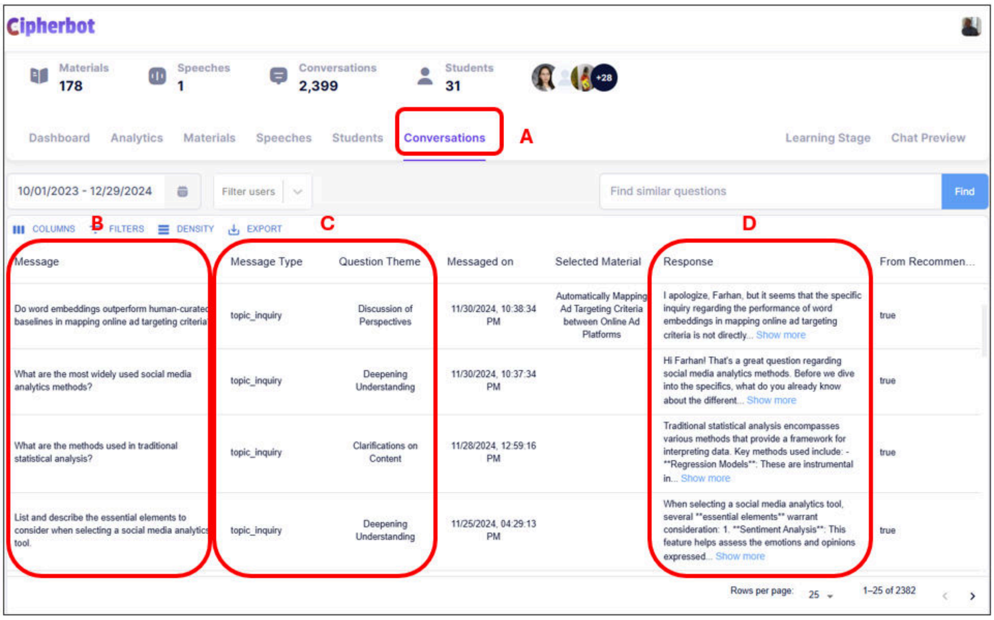


Fig. 3. Conversations tab for educators **A**-where teachers can view conversations that have occurred within the course, including **B**-student questions, **C**-message types and themes, and **D**-Cipherbot responses, among many other conversation attributes, including questions not answered by the system.

B. Class Management by Educators

Cipherbot simplifies class management through a dedicated educator interface. As illustrated in Fig. 2, the dashboard presents key course metrics, including student enrollment, uploaded materials, and the volume of student interactions. Educators can easily create new classes and upload instructional content in various formats, such as PDF, text, video, audio, PowerPoint, and Word, supporting broad compatibility and ensuring accessibility across various teaching resources. Cipherbot can also generate audio from textual files, including translation (e.g., from an English document generate an audio lecture in French).

The Conversations tab (Fig. 3) provides educators with a comprehensive view of all interactions between students and Cipherbot. The system logs each student's chat history in full, enabling students to revisit prior exchanges and allowing instructors to monitor dialogue trends over time. Educators can access complete transcripts, including student queries and corresponding chatbot responses, offering transparency into how learners interact with the system. To support instructional decision-making, Cipherbot automatically categorizes each exchange into pedagogically meaningful themes, such as *content clarification*, *application of knowledge*, *assignment support*, *conceptual connections*, *exam preparation*, and *critical thinking*, among others (see Table 1). This structured classification enables more targeted, data-driven teaching strategies personalized to each student within the course.

TABLE I. THE CATEGORIES OF STUDENT QUESTION TYPES ON CIPHERBOT, ALONG WITH THEIR DESCRIPTIONS.

Category	Description
Clarifications	Requests for clarification on specific concepts or topics.
Application of knowledge	Inquiries about applying class concepts in real-world scenarios.
Assignment assistance	Seeking help with assignments or homework.
Connections to other topics	Questions about how the current material relates to other topics.
Discussion of perspectives	Seeking different perspectives on a concept or theory.
Exam preparation	Asking for guidance on exam focus and format.
Deepening understanding	Requests for additional resources to explore a topic further.
Feedback	Providing feedback on lecture clarity and effectiveness.
Current events	Inquiries about how current events relate to the course.
Critical thinking	Asking thought-provoking questions for critical thinking.
Practical challenges	Seeking advice on overcoming challenges in understanding.
Greetings	Simple greetings such as ‘Hi’ or ‘Hey.’

Each category reflects a specific facet of student engagement, providing educators with granular insight into inquiries and communication patterns emerging during chatbot interactions. Cipherbot performs automatic classification of conversation, enabling it to generate contextually adaptive and personalized responses. Additionally, Cipherbot labels each exchange as “answered” or “not answered” based on whether the response is grounded in the uploaded instructional content. This labeling supports instructor evaluation of response accuracy and relevance. By organizing interactions into structured categories, Cipherbot equips educators with actionable data to tailor instruction, identify learning gaps, and refine pedagogical strategies—ultimately enhancing the overall learning experience.

Cipherbot’s class management features are central to its goal of delivering scalable, personalized learning while minimizing hallucination risk. The system ensures that all chatbot responses are based on verified, course-specific materials by enabling educators to upload, structure, and annotate instructional content with metadata, such as lecture dates and thematic tags. This functionality allows instructors to align the AI assistant’s knowledge base with the pedagogical goals of each course, increasing both the precision and contextual relevance of generated responses. Moreover, the centralized logging of student interaction data, including question categories and engagement levels, supports large-scale monitoring of learning behaviors. These capabilities make the class management interface a foundational element of Cipherbot’s architecture, directly supporting the system’s design emphasis on personalization, scalability, and content alignment.

C. Dialogues with Cipherbot by Students

As illustrated in Fig. 4, students interact with the system via smartphone, asking questions related to specific resources, themes, or spanning multiple instructional documents. Cipherbot streamlines the question-answering process by grounding all responses in educator-uploaded materials. To preserve contextual coherence, each enrolled class operates within its own dedicated chat session. All student interactions are automatically logged, allowing learners and instructors to revisit previous conversations as needed. Beyond receiving AI-generated responses, students can directly access linked course materials and explore referenced content in greater depth (see Fig. 5), supporting deeper engagement with the curriculum.

Fig. 4. Student conversation interface. **A**-student query, **B**-Cipherbot response with citations to course material, **C**-student reactions, copy, translate, etc., **D**-clickable course material, **E**-follow-up questions, and **F**-conversation history.

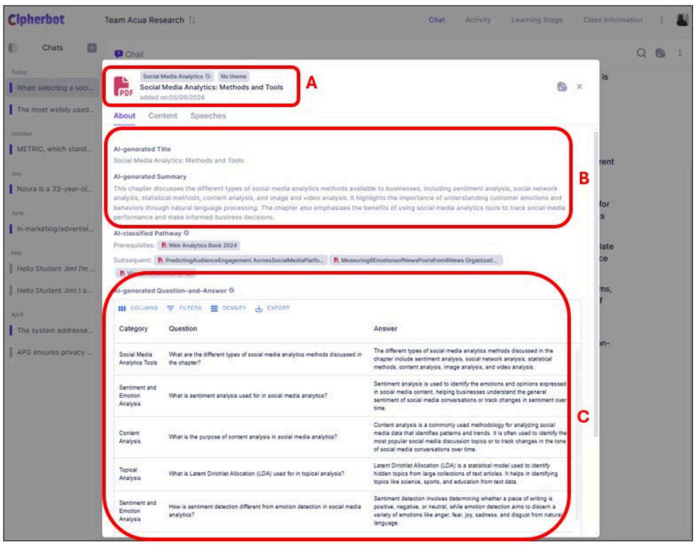


Fig. 5. Course material interface. **A**-material name, category, and theme, **B**-LLM generated title and summary based on the content, and **C**-LLM generated question and answer pairs. The Cipherbot responses are based on the course material vetted by the instructor.

IV. Student Evaluation of Cipherbot

A. Pilot Testing With Educators and Students

A pilot evaluation was conducted with four students and four educators to assess Cipherbot's core functionality across device types. Participants were evenly distributed between smartphones and desktop computers to ensure balanced testing. Educators completed 19 tasks, including account creation, class configuration, and content management, with success rates of 97.3% on smartphones and 100% on computers. Students completed 11 tasks, such as registration, initiating class chats, and navigating learning materials, with success rates of 95.4% on smartphones and 86.3% on computers. Insights from the pilot informed iterative design improvements. For example, participant P3 reported confusion when selecting materials during a chat session on a desktop, stating, "Something is up with the flow. You have to click a profile to get back to the main page. I could not figure out how to select the material." In response, the system interface was refined to include a prominent "ask about specific material" button beneath the input field, enhancing usability and streamlining task flow.

B. Usability and Functionality Assessment by Undergraduate Students

Building on insights from the pilot phase, we conducted a longitudinal user study to evaluate Cipherbot's usability and functional performance in a real-world educational context. Sixteen undergraduate students used Cipherbot over eight weeks, with all system interactions continuously logged for analysis. Following the deployment, we conducted a comprehensive usability assessment to evaluate Cipherbot's effectiveness in supporting educational engagement and delivering personalized, course-aligned learning support.

Participants and Procedure. Sixteen undergraduate students, enrolled in an introductory in-person HCI course, participated in the study. Participants, aged 18 to 22, were recruited from six universities, five located in Qatar and one in the United States, and represented diverse academic backgrounds, including Computer Science, Business, Political Science, and Digital Media. Of the cohort, 12 participants (75%) identified as female and 4 (25%) as male, with no additional gender identities reported. The study was conducted in two phases, designed to evaluate both the usability and functional performance of Cipherbot in an authentic learning environment with real course content in an actual learning environment. As this was a real-world deployment, this is a naturalistic rather than a controlled lab study with control and test groups.

Usage of Cipherbot. Over the eight-week study period, participants actively used Cipherbot as part of their coursework. The system automatically captured detailed logs of all user activity, including logins, conversation histories, and specific queries. This interaction data provided a comprehensive view of how students incorporated the chatbot into their learning workflows, offering valuable insights into usage patterns, engagement frequency, and the integration of AI support within everyday academic routines.

Post-Course Usability Assessment. To evaluate Cipherbot's usability and overall performance, all sixteen participants completed a quantitative assessment using three standardized instruments: (1) the System Usability Scale (SUS) to evaluate general usability [45], (2) the Chatbot Usability Questionnaire (CUQ) to assess the quality of dialogue-based interactions [46], and (3) the User Experience Questionnaire (UEQ) to capture perceptions across core user experience (UX) dimensions [47]. This multi-instrument approach enabled a comprehensive evaluation of Cipherbot's interface, interaction quality, and user satisfaction, addressing a key gap in standardized assessment practices within AI educational assistant research [40].

C. Results of the Usability Study

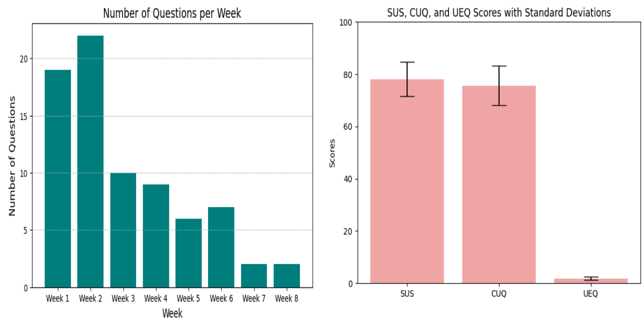
Conversation Types: Throughout the eight-week study, participants engaged with Cipherbot by posing a total of 77 questions, averaging 5 queries per student (range: 3–14). Using the underlying LLM, Cipherbot automatically classified these questions into predefined pedagogical categories: Clarifications on Content (n=34, 44%), Assignment and Homework Assistance (n=17, 22%), Practical Challenges (n=8, 11%), Deepening Understanding (n=8, 11%), and Application of Knowledge (n=5, 7%). This distribution is consistent with prior research on educational AI chatbot usage patterns [48], reinforcing the contextual validity and relevance of student interactions throughout the deployment and generalizability of findings.

Each category imposes distinct demands on information retrieval and synthesis, for example, Assignment and Homework Assistance typically require exact matches to instructional content, whereas Application of Knowledge may necessitate more abstract reasoning and cross-topic integration. Future evaluations will incorporate a category-level analysis of response success rates to better understand Cipherbot's performance across different query types. By assessing response accuracy within each category, we aim to identify areas where Cipherbot performs reliably versus where it encounters limitations. Preliminary observations indicate that most failures occurred in response to abstract or cross-cutting queries that exceeded the scope of the uploaded materials, often resulting in ungrounded or incomplete answers. A more systematic failure analysis, tracking instances of hallucination and ambiguity, will be essential to refine prompt engineering, improve content structuring, and enhance the system's overall reliability. However, from our preliminary findings, the professor setting the expectations of course content is also a key factor for generating in-scope conversations.

Student Engagement: As illustrated in Fig. 6a, student interaction with Cipherbot declined progressively over the eight-week study—a trend consistent with prior research on educational chatbot engagement [49], [50], [51]. The most notable drop occurred during the final two weeks, coinciding with a shift in course focus toward project completion and final report submissions. This pattern aligns with expected academic workflows and mirrors documented engagement trajectories in earlier studies [52]. Despite the decline, all participants remained active users of Cipherbot throughout the study, and their interaction patterns closely followed established benchmarks [48], providing a robust basis for post-study evaluation and longitudinal analysis [53].

Post-Course Usability Assessment: At the end of the eight-week study, participants completed three standardized evaluations to assess Cipherbot's usability and user experience. As shown in Fig. 6b, the SUS yielded an average score of 78.1 (max = 87.5, min = 65.5, SD = 6.61), indicating strong overall usability [45]. The CUQ averaged 75.6 (max = 85.9, min = 60.9, SD = 7.57), reflecting high-quality dialogue-based interaction [46]. The UEQ produced a mean score of

1.91 (max = 3.0, min = 0.50, SD = 0.72), placing Cipherbot within the good to excellent UX range [47]. Collectively, these results validate the system's effectiveness in delivering a usable and engaging learning interface. However, the observed decline in weekly usage suggests a need to incorporate ongoing, interactive learning activities to sustain long-term engagement with educational assistants [54].



(a) Number of Conversations by Week.

(b) Comparison of SUS, CUQ, and UEQ scores with standard deviations.

Fig. 6. Conversations by week from the students and the SUS, CUQ, and UEQ results.

Although a gradual decline in student engagement was observed over the eight-week deployment, this study does not systematically investigate the underlying causes of that trend. Potential explanations include interface fatigue, diminished perceived value of chatbot-mediated tasks, or the natural shift in course focus toward final projects. However, in the absence of qualitative data, these interpretations remain speculative, requiring future work incorporating interviews and open-ended survey instruments to elicit student perspectives on engagement dynamics. These insights will inform targeted design refinements, such as adaptive prompts, personalized interface features, or gamification, to support sustained, meaningful interaction over extended learning periods.

While Cipherbot leverages a RAG framework to enhance factual accuracy by constraining responses to educator-uploaded materials, the binary classification of responses as “answered” or “not answered” serves as a preliminary grounding indicator that the RAG approach worked well. The complication is that determining what is ‘out of scope’ and what is just ‘basic knowledge’ has no clear boundary for many courses. From a review of the conversations, Cipherbot accurately addressed out-of-scope questions with a “that is not covered in the course material” response.

V. Strengths, Limitations, and Future Work

Cipherbot offers a promising direction for scalable, personalized education through AI-driven, instructor-aligned conversational interfaces. Its core strength lies in integrating educator-curated content with a RAG framework, enabling contextually grounded and reliable responses while reducing the risk of misinformation common in general-purpose chatbots. The platform's content management system supports flexible organization and annotation of materials, while its automatic categorization of student queries provides instructors with actionable insights into learning needs. Broad format compatibility and cross-device access enhance usability, and 24/7 availability supports self-paced, flexible engagement across diverse learning contexts.

Concerning limitations, the evaluation was conducted in a single academic context, involving sixteen undergraduate students enrolled in a HCI course. This relatively homogeneous and tech-savvy cohort limits the generalizability of findings to broader educational settings, particularly for learners from non-technical disciplines, such as the arts or humanities, who may exhibit different levels of digital literacy and usability expectations. Additionally, Cipherbot's performance may vary depending on the complexity and structure of the subject matter, underscoring the need for cross-disciplinary testing. While usage data was collected, the study did not investigate long-term learning outcomes, content retention, or academic performance metrics. The decline in student engagement toward the end of the study further highlights the ongoing challenge of sustaining interaction with educational chatbots, even when usability and content alignment are strong.

Future work will extend the evaluation of Cipherbot across a broader set of educational domains to assess its adaptability and pedagogical effectiveness in diverse instructional contexts. In parallel, we plan to apply advanced chat log analytics to uncover patterns in unanswered or frequently asked questions, enabling the identification of content gaps and informing the development of targeted supplemental materials. These insights will support educators in refining course content, addressing learner confusion, and optimizing instructional strategies. While the current study emphasizes usability, engagement, and functional performance, future research will also evaluate Cipherbot's influence on learning outcomes and student experiences compared to traditional educational support methods.

Future research should include controlled comparative studies to assess differences in academic performance, content retention, and learner satisfaction between Cipherbot students and those who do not [55]. Such evaluations will offer deeper insight into Cipherbot's effectiveness in delivering scalable, personalized educational support, extending beyond question-answering accuracy to its broader pedagogical impact, as outlined in the introduction. In addition, future work will conduct benchmarking against both general-purpose systems (e.g., ChatGPT, Phind) and domain-specific agents (e.g., Khanmigo), employing standardized hallucination detection metrics to evaluate factual grounding and reliability. These comparative analyses will provide stronger empirical validation of Cipherbot's core contribution—enhancing trust and accuracy in AI-mediated educational interactions.

While the quantitative evaluation using SUS, CUQ, and UEQ provided a solid foundation for assessing Cipherbot's usability and overall user experience, these instruments offer limited visibility into users' contextual expectations and specific functional preferences. A limitation of the present study is the absence of qualitative data, such as interviews or open-ended responses, that could deepen our understanding of how students interpret and experience Cipherbot in real-world settings. Future evaluations will integrate qualitative methods to enable a more nuanced exploration of user needs, reveal pain points and strengths not captured by quantitative scales, and inform the iterative development of features grounded in learner and educator perspectives.

Cipherbot will explore the integration of knowledge graphs, mobile elements such as podcasts, adaptive learning pathways, and immersive features aligned with metaverse technologies to promote deeper, ongoing interaction by students. These enhancements will hopefully enrich the learning experience, increase motivation, and support individualized educational journeys. The goal is to position Cipherbot and other AI educational assistants as flexible, trustworthy, and engaging platforms for personalized learning and digital education [56].

VI. Conclusion

This study introduces Cipherbot, an AI-driven educational assistant designed to support scalable, personalized learning by integrating educator-curated content and LLM capabilities. By combining a RAG framework with structured prompt engineering, Cipherbot addresses two challenges in AI-enhanced education, which are delivering personalized support at scale and ensuring factual accuracy in system responses. Results from an eight-week deployment with undergraduate students demonstrated high usability, high engagement at the course start, and pedagogically grounded interactions, validating Cipherbot's effectiveness in facilitating course-specific question answering.

As educational institutions continue integrating AI and immersive technologies into instructional practice, Cipherbot presents a compelling design model for embedding educational assistants within pedagogical workflows. Cipherbot's design emphasizes intelligent, adaptive, and trustworthy interaction, aligning with broader goals of value creation in education. While challenges such as maintaining long-term engagement persist, Cipherbot establishes a strong foundation for advancing research in personalized, AI-mediated learning systems. This work contributes to ongoing discourse around the role of AI technologies in shaping more inclusive, responsive, and effective educational experiences.

References

[1] R. Meyer Von Wolff, J. Nörtemann, S. Hobert, and M. Schumann, "Chatbots for the Information Acquisition at Universities – A Student's View on the Application Area," in *Chatbot Research and Design*, A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, and P. B. Brandtzaeg, Eds., in Lecture Notes in Computer Science, vol. 11970. Cham: Springer International Publishing, 2020, pp. 231–244. doi: 10.1007/978-3-030-39540-7_16.

[2] R. Winkler and M. Soellner, "Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis," *Proceedings*, vol. 2018, no. 1, p. 15903, Aug. 2018, doi: 10.5465/AMBPP.2018.15903abstract.

[3] C. G. Brinton, R. Rill, S. Ha, M. Chiang, R. Smith, and W. Ju, "Individualization for education at scale: MIIC design and preliminary evaluation," *IEEE Transactions on Learning Technologies*, vol. 8, no. 1, pp. 136–148, 2014.

[4] K. S. Hone and G. R. El Said, "Exploring the factors affecting MOOC retention: A survey study," *Computers & Education*, vol. 98, pp. 157–168, 2016.

[5] L. Vermette, J. McGrenere, C. Birge, A. Kelly, and P. K. Chilana, "Freedom to Personalize My Digital Classroom: Understanding Teachers' Practices and Motivations," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland Uk: ACM, May 2019, pp. 1–14. doi: 10.1145/3290605.3300548.

[6] L. B. Portnoy and M. Rabinowitz, "What's in a domain: understanding how students approach questioning in history and science," *Educational Research and Evaluation*, vol. 20, no. 2, pp. 122–145, Feb. 2014, doi: 10.1080/13803611.2014.885844.

[7] J. T. Dillon, "A norm against student questions," *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, vol. 55, no. 3, pp. 136–139, 1981.

[8] M. Micari and S. Calkins, "Is it OK to ask? The impact of instructor openness to questions on student help-seeking and academic outcomes," *Active Learning in Higher Education*, vol. 22, no. 2, pp. 143–157, July 2021, doi: 10.1177/1469787419846620.

[9] T. Wambsganß, L. Haas, and M. Söllner, *Towards the Design of a Student-Centered Question-Answering System in Educational Settings*. 2021.

[10] C. Zheng, Y. Zhang, Z. Huang, C. Shi, M. Xu, and X. Ma, "DiscipLink: Unfolding Interdisciplinary Information Seeking Process via Human-AI Co-Exploration," in *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, Pittsburgh PA USA: ACM, Oct. 2024, pp. 1–20. doi: 10.1145/3654777.3676366.

[11] H. Zhang *et al.*, "ProtoDreamer: A Mixed-prototype Tool Combining Physical Model and Generative AI to Support Conceptual Design," in *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, Pittsburgh PA USA: ACM, Oct. 2024, pp. 1–18. doi: 10.1145/3654777.3676399.

[12] L. Curtis, A. Ueno, R. Wood, and C. Yu, "The Metaverse in Higher Education: The Metaverse Conference 2023," May 2023. Accessed: Mar. 21, 2025. [Online]. Available: https://airsi.unizar.es/wp-content/uploads/docs/2023/airsi2023-proceedings.pdf

[13] E. Mohammadi, M. Thelwall, Y. Cai, T. Collier, I. Tahamtan, and A. Eftekhar, "Is generative AI reshaping academic practices worldwide? A survey of adoption, benefits, and concerns," *Information Processing & Management*, vol. 63, no. 1, p. 104350, 2026.

[14] D. Feng, E. Shaw, J. Kim, and E. Hovy, "An intelligent discussion-bot for answering student queries in threaded discussions," in *Proceedings of the 11th international conference on Intelligent user interfaces*, 2006, pp. 171–177.

[15] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.

[16] R. Zhao, J. Tang, W. Zeng, Y. Guo, and X. Zhao, "Towards human-like questioning: Knowledge base question generation with bias-corrected reinforcement learning from human feedback," *Information Processing & Management*, vol. 62, no. 3, p. 104044, 2025, doi: https://doi.org/10.1016/j.ipm.2024.104044.

[17] G. Liu, M. Ren, L. Guo, J. Li, and M. Ma, "Comprehensive exercise recommendation with practicality, generalizability, and versatility in AI-driven education," *Information Processing & Management*, vol. 62, no. 3, p. 104051, 2025, doi: https://doi.org/10.1016/j.ipm.2024.104051.

[18] W. Maroengsit, T. Piyakulpinyo, K. Phonyiam, S. Pongnumkul, P. Chaovalit, and T. Theeramunkong, "A Survey on Evaluation Methods for Chatbots," in *Proceedings of the 2019 7th International Conference on Information and Education Technology*, Aizu-Wakamatsu Japan: ACM, Mar. 2019, pp. 111–119. doi: 10.1145/3323771.3323824.

[19] D. Kaczorowska-Spychalska, "How chatbots influence marketing," *Management*, vol. 23, no. 1, pp. 251–270, 2019.

[20] Grand View Research, "AI In Education Market Size & Share | Industry Report, 2030." Accessed: July 30, 2025. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report

[21] N. Editorials, "Tools such as ChatGPT threaten transparent science; here are our ground rules for their use," *Nature*, vol. 613, no. 612, pp. 10–1038, 2023.

[22] Y. Zhu, Y. Lu, H. Xie, J. Ye, and M. Chen, "A quasi-experimental analysis of capabilities and limitations of generative AI in academic content evaluation in social sciences," *Information Processing & Management*, vol. 63, no. 1, p. 104365, 2026.

[23] M. Yoo, H. Jin, and J. Kim, "How Do Teachers Create Pedagogical Chatbots?: Current Practices and Challenges," Apr. 04, 2025, *arXiv*: arXiv:2503.00967. doi: 10.48550/arXiv.2503.00967.

[24] J. Yin, T.-T. Goh, and Y. Hu, "Interactions with educational chatbots: the impact of induced emotions and students' learning motivation," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 1, p. 47, 2024.

[25] C. Zhao *et al.*, "CyberBOT: Towards Reliable Cybersecurity Education via Ontology-Grounded Retrieval Augmented Generation," Apr. 01, 2025, *arXiv*: arXiv:2504.00389. doi: 10.48550/arXiv.2504.00389.

[26] Z. Li, Z. Wang, W. Wang, K. Hung, H. Xie, and F. L. Wang, "Retrieval-augmented generation for educational application: A systematic survey," *Computers and Education: Artificial Intelligence*, p. 100417, 2025.

[27] R. Winkler, S. Hobert, A. Salovaara, M. Söllner, and J. M. Leimeister, "Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, Apr. 2020, pp. 1–14. doi: 10.1145/3313831.3376781.

[28] H. Yusuf, A. Money, and D. Daylamani-Zad, "Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art," *Education Tech Research Dev*, vol. 73, no. 2, pp. 815–874, Apr. 2025, doi: 10.1007/s11423-025-10447-4.

[29] G. Liu, M. Ren, L. Guo, J. Li, and M. Ma, "Comprehensive exercise recommendation with practicality, generalizability, and versatility in AI-driven education," *Information Processing & Management*, vol. 62, no. 3, p. 104051, 2025.

[30] V. Ashrafimoghari, N. Gürkan, and J. W. Suchow, "Evaluating Large Language Models on the GMAT: Implications for the Future of Business Education," Jan. 01, 2024, *arXiv*: arXiv:2401.02985. doi: 10.48550/arXiv.2401.02985.

[31] Q. Li, Y. Li, S. Zhang, X. Zhou, and Z. Pan, "A theoretical framework for human-centered intelligent information services: A systematic review," *Information Processing & Management*, vol. 62, no. 1, p. 103891, 2025, doi: https://doi.org/10.1016/j.ipm.2024.103891.

[32] P. P. Ray, "How true is the role of large language models in nursing?," *European Journal of Cardiovascular Nursing*, p. zvad123, 2024.

[33] I. J. Akpan, Y. M. Kobara, J. Owolabi, A. A. Akpan, and O. F. Offodile, "Conversational and generative artificial intelligence and human–chatbot interaction in education and research," *Int Trans Operational Res*, vol. 32, no. 3, pp. 1251–1281, May 2025, doi: 10.1111/itor.13522.

[34] M. Park, S. Kim, S. Lee, S. Kwon, and K. Kim, "Empowering personalized learning through a conversation-based tutoring system with student modeling," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–10.

[35] R. Sajja, Y. Sermet, and I. Demir, "An Open-Source Dual-Loss Embedding Model for Semantic Retrieval in Higher Education," May 08, 2025, *arXiv*: arXiv:2505.04916. doi: 10.48550/arXiv.2505.04916.

[36] M. Younas, D. A. S. El-Dakhs, and Y. Jiang, "A Comprehensive Systematic Review of AI-Driven Approaches to Self-Directed Learning," *IEEE Access*, 2025, Accessed: Mar. 21, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10906568/

[37] H. Abu-Rasheed *et al.*, "Llm-assisted knowledge graph completion for curriculum and domain modelling in personalized higher education recommendations," in *2025 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, 2025, pp. 1–5. Accessed: July 30, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11016377/?casa_token=nP1twH0IP7wAAAAA:3yJ3fpQvo7CXv9WjoOvmbokd1pg4tkYqX6GeGKJxShLAZfXIXNSZ1GTh-uQBL_DIsmmsqJOJNQ

[38] M. Fidan and N. Gencel, "Supporting the Instructional Videos With Chatbot and Peer Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation," *Journal of Educational Computing Research*, vol. 60, no. 7, pp. 1716–1741, Dec. 2022, doi: 10.1177/07356331221077901.

[39] X. Ma *et al.*, "DRESSing Up LLM: Efficient Stylized Question-Answering via Style Subspace Editing," Jan. 24, 2025, *arXiv*: arXiv:2501.14371. doi: 10.48550/arXiv.2501.14371.

[40] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Educ Inf Technol*, vol. 28, no. 1, pp. 973–1018, Jan. 2023, doi: 10.1007/s10639-022-11177-3.

[41] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[42] Z. Jiang *et al.*, "Active Retrieval Augmented Generation." arXiv, Oct. 2023. Accessed: Sept. 14, 2024. [Online]. Available: http://arxiv.org/abs/2305.06983

[43] P. Zhao *et al.*, "Retrieval-Augmented Generation for AI-Generated Content: A Survey." arXiv, June 2024. Accessed: Sept. 14, 2024. [Online]. Available: http://arxiv.org/abs/2402.19473

[44] D. T. K. Ng, C. W. Tan, and J. K. L. Leung, "Empowering student self-regulated learning and science education through CHATGPT : A pioneering pilot study," *Brit J Educational Tech*, vol. 55, no. 4, pp. 1328–1353, July 2024, doi: 10.1111/bjet.13454.

[45] J. Brooke, "SUS-A quick and dirty usability scale," in *Usability evaluation in industry*, vol. 189, London, England, 1996, pp. 4–7.

[46] S. Holmes, A. Moorhead, R. Bond, H. Zheng, V. Coates, and M. Mctear, "Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?," in *Proceedings of the 31st European Conference on Cognitive Ergonomics*, BELFAST United Kingdom: ACM, Sept. 2019, pp. 207–214. doi: 10.1145/3335082.3335094.

[47] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (UEQ-S)," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 6, pp. 103–108, 2017.

[48] L. Labadze, M. Grigolia, and L. Machaidze, "Role of AI chatbots in education: systematic literature review," *Int J Educ Technol High Educ*, vol. 20, no. 1, p. 56, Oct. 2023, doi: 10.1186/s41239-023-00426-1.

[49] J. Salminen *et al.*, "Communication Design for an Educational AI Chatbot: Analyzing Cipherbot's Communication Style and Challenges," in *Proceedings of the 27th International Academic Mindtrek Conference*, Tampere Finland: ACM, Oct. 2024, pp. 176–187. doi: 10.1145/3681716.3681727.

[50] J. Salminen *et al.*, "Using Cipherbot: An Exploratory Analysis of Student Interaction with an LLM-Based Educational Chatbot," in *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, Atlanta GA USA: ACM, July 2024, pp. 279–283. doi: 10.1145/3657604.3664690.

[51] J. Van De Pol, M. Volman, and J. Beishuizen, "Scaffolding in Teacher–Student Interaction: A Decade of Research," *Educ Psychol Rev*, vol. 22, no. 3, pp. 271–296, Sept. 2010, doi: 10.1007/s10648-010-9127-6.

[52] L. K. Fryer, M. Ainley, A. Thompson, A. Gibson, and Z. Sherlock, "Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners," *Computers in Human Behavior*, vol. 75, pp. 461–468, Oct. 2017, doi: 10.1016/j.chb.2017.05.045.

[53] Y. Chen, S. Jensen, L. J. Albert, S. Gupta, and T. Lee, "Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success," *Inf Syst Front*, vol. 25, no. 1, pp. 161–182, Feb. 2023, doi: 10.1007/s10796-022-10291-4.

[54] Y. Wang, S. Guo, L. Ling, and C. W. Tan, "Nemobot: Crafting Strategic Gaming LLM Agents for K-12 AI Education," in *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, Atlanta GA USA: ACM, July 2024, pp. 393–397. doi: 10.1145/3657604.3664671.

[55] A. Farooq *et al.*, "Representing groups of students as personas: A systematic review of persona creation, application, and trends in the educational domain," *Computers and Education Open*, p. 100242, 2025.

[56] Y. Sharma, A. Suri, R. Sijariya, and L. Jindal, "Role of education 4.0 in innovative curriculum practices and digital literacy– A bibliometric approach," *E-Learning and Digital Media*, vol. 22, no. 1, pp. 1–32, Jan. 2025, doi: 10.1177/20427530231221073.