

Article

ROboMC: A Portable Multimodal System for eHealth Training and Scalable AI-Assisted Education

Marius Cioca ^{1,*}  and Adriana-Lavinia Cioca ^{2,3}

¹ Department of Industrial Engineering and Management, Faculty of Engineering, Lucian Blaga University of Sibiu, 550024 Sibiu, Romania

² Department of Clinical Medicine, Faculty of Medicine, Lucian Blaga University of Sibiu, 550024 Sibiu, Romania; adrianalavinia.cioca@ulbsibiu.ro or adriana.cioca@gmail.com

³ CMI Cioca Adriana Lavinia, 550231 Sibiu, Romania

* Correspondence: marcus.cioca@ulbsibiu.ro

Abstract

AI-based educational chatbots can expand access to learning, but many remain limited to text-only interfaces and fixed infrastructures, while purely generative responses raise concerns of reliability and consistency. In this context, we present ROboMC, a portable and multimodal system that combines a validated knowledge base with generative responses (OpenAI) and voice–text interaction, designed to enable both text and voice interaction, ensuring reliability and flexibility in diverse educational scenarios. The system, developed in Django, integrates two response pipelines: local search using normalized keywords and fuzzy matching in the LocalQuestion database, and fallback to the generative model GPT-3.5-Turbo (OpenAI, San Francisco, CA, USA) with a prompt adapted exclusively for Romanian and an explicit disclaimer. All interactions are logged in AutomaticQuestion for later analysis, supported by a semantic encoder (SentenceTransformer—paraphrase-multilingual-MiniLM-L12-v2', Hugging Face Inc., New York, NY, USA) that ensures search tolerance to variations in phrasing. Voice interaction is managed through gTTS (Google LLC, Mountain View, CA, USA) with integrated audio playback, while portability is achieved through deployment on a Raspberry Pi 4B (Raspberry Pi Foundation, Cambridge, UK) with microphone, speaker, and battery power. Voice input is enabled through a cloud-based speech-to-text component (Google Web Speech API accessed via the Python SpeechRecognition library, (Anthony Zhang, open-source project, USA) using the Google Web Speech API (Google LLC, Mountain View, CA, USA; language = “ro-RO”), allowing users to interact by speaking. Preliminary tests showed average latencies of 120–180 ms for validated responses on laptop and 250–350 ms on Raspberry Pi, respectively, 2.5–3.5 s on laptop and 4–6 s on Raspberry Pi for generative responses, timings considered acceptable for real educational scenarios. A small-scale usability study ($N \approx 35$) indicated good acceptability (SUS $\sim 80/100$), with participants valuing the balance between validated and generative responses, the voice integration, and the hardware portability. Although system validation was carried out in the eHealth context, its architecture allows extension to any educational field: depending on the content introduced into the validated database, ROboMC can be adapted to medicine, engineering, social sciences, or other disciplines, relying on ChatGPT only when no clear match is found in the local base, making it a scalable and interdisciplinary solution.

Keywords: AI chatbots; educational technology; eHealth training; ChatGPT; machine learning; speech interaction; Raspberry Pi; accessibility; personalized learning



Academic Editor: Francesco Camastra

Received: 19 September 2025

Revised: 7 November 2025

Accepted: 9 November 2025

Published: 11 November 2025

Citation: Cioca, M.; Cioca, A.-L. ROboMC: A Portable Multimodal System for eHealth Training and Scalable AI-Assisted Education. *Inventions* **2025**, *10*, 103. <https://doi.org/10.3390/inventions10060103>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence (AI) has become a central pillar of digital transformation in education, enabling the development of systems capable of supporting personalized learning, providing rapid feedback, and facilitating large-scale interactivity [1–3]. In this context, chatbots have attracted particular attention as tools that can simulate human dialogue and provide educational support tailored to learners' needs. Existing studies highlight students' positive perceptions of AI-based digital assistants, underlining their role in enhancing engagement and accessibility [1,4,5]. Systematic reviews confirm these trends but also point to challenges related to the integration of chatbots into higher education [4,6].

The project originated from classroom observations where students frequently used general-purpose chatbots for quick answers, often receiving inaccurate or inconsistent feedback. This motivated the authors to design a controlled, domain-specific assistant that combines validated knowledge with generative flexibility, focusing on real classroom usability rather than theoretical modeling.

Recent literature highlights the increasing integration of generative AI and hybrid learning technologies in higher education, reflecting a global movement toward scalable, sustainable, and adaptive educational ecosystems. In [7], the authors analyzed the role of information and communication technologies combined with generative AI tools such as ChatGPT in advancing sustainable development goals across universities. Their findings emphasize that generative AI supports personalized learning, efficient resource management, and inclusive access to education—particularly in developing contexts where digital infrastructures are limited. This aligns with broader efforts to leverage AI-driven automation to enhance sustainability in academic institutions while maintaining pedagogical quality and accessibility.

Similarly, Silva et al. [8] explored the benefits and challenges of adopting ChatGPT in software engineering education, noting its dual capacity to enhance engagement and support individualized instruction while also introducing concerns about overreliance, ethics, and academic integrity. The study underscores the importance of framing generative AI as a complementary tool rather than a replacement for human-led instruction. These insights reinforce the need for controlled and validated AI-mediated learning environments—an approach that parallels ROboMC's emphasis on combining verified knowledge with generative adaptability.

Expanding the discussion beyond single-domain applications, Wangsa et al. [9] conducted a comprehensive review of leading AI chatbot models, including ChatGPT, Bard, LLaMA, and Ernie, comparing their deployment across education and healthcare. Their analysis identified hybrid frameworks—those integrating structured databases with generative reasoning—as the most promising for ensuring accuracy and contextual reliability. This synthesis further supports the relevance of multimodal systems like ROboMC, where validated repositories coexist with large language models to achieve a balance between scalability and trust.

Collectively, these recent studies illustrate an academic shift from purely generative or rule-based designs toward hybrid AI ecosystems that promote sustainable, ethical, and learner-centered innovation. ROboMC builds upon this trajectory, contributing a portable and domain-adaptable implementation that operationalizes these emerging paradigms in real educational settings.

Recent developments in hybrid AI chatbots have combined retrieval-augmented generation (RAG) with validated domain knowledge to improve reliability and contextual grounding [7–9].

Building on previous studies that explored retrieval–generation integration, adaptive context modeling, and multimodal interfaces [7–9], three main methodological families can be identified.

Hybrid RAG frameworks implement dual-stage pipelines in which textual and relational knowledge bases are searched for semantically relevant passages that are concatenated with the user prompt before large-language-model inference, ensuring factual grounding of generated content.

Context-adaptive educational chatbots employ embedding-based retrievers that dynamically update local indexes of course materials, allowing rapid adaptation to newly added documents.

Multimodal AI assistants combine speech input, short-term dialogue memory, and visual feedback to enhance user engagement while maintaining pedagogical structure.

Compared with these approaches, ROboMC uniquely integrates an expert-validated local knowledge base with a generative fallback model, thus maintaining both transparency and pedagogical reliability within a portable edge-computing implementation.

This positioning clarifies its scientific relevance beyond a technical prototype. Recent research confirms that multimodal solutions add educational value. For example, Rienties and colleagues analyzed students' perceptions of AI digital assistants (AIDAs), confirming their potential to complement traditional instruction through real-time support [5]. Similarly, Chen and colleagues demonstrated the effectiveness of human digital technologies in e-learning, showing how interactive, voice-based content improves engagement and learning outcomes [10]. These contributions highlight the growing relevance of multimodal approaches in education.

Nevertheless, most existing solutions remain limited to text-based interactions and fixed infrastructures (laptops or desktops), which reduces accessibility for users with special needs (e.g., visually impaired individuals) and for educational contexts that require mobility, such as clinical training. In addition, many systems lack mechanisms for recording and analyzing interactions, limiting teachers' ability to identify knowledge gaps and adapt the curriculum.

To address these challenges, we present ROboMC, an innovative platform that combines a validated knowledge base with a generative model (ChatGPT), voice–text integration, and hardware portability. The system supports multimodal interaction, stores all questions for later analysis through machine learning, and enables lessons to be conducted in diverse educational contexts. Although the case study presented is focused on medical education, its architecture is domain-independent and can be adapted to engineering, social sciences, or other disciplines.

Despite these advances, most current chatbots lack local expert validation mechanisms and portability across hardware environments. ROboMC addresses these gaps by combining a curated local knowledge base, generative fallback, and Raspberry Pi deployment, demonstrating a balance between transparency, accessibility, and adaptability in educational settings.

Recent advances in conversational AI have led to a new generation of intelligent chatbots that combine large language models (LLMs) with external retrieval mechanisms and adaptive memory components.

State-of-the-art approaches such as retrieval-augmented generation (RAG) systems integrate a document retrieval step before LLM prompting, ensuring that the generated output remains grounded in verifiable knowledge.

Other research focuses on memory-enhanced dialogue agents, which store and reuse past interactions to maintain contextual continuity during extended conversations. In

parallel, multimodal assistants extend these capabilities by integrating speech-to-text and text-to-speech modules, enhancing accessibility and user engagement.

Within this evolving landscape, ROboMC represents a hybrid solution: it integrates a validated local knowledge base for factual accuracy, a fuzzy–semantic matching layer for efficient retrieval, and a GPT-based generative module for adaptive responses. This design combines the transparency of structured knowledge systems with the flexibility of modern LLMs, providing a reliable and pedagogically focused assistant suitable for educational environments.

2. Materials and Methods

This section presents the research methodology applied to design, implement, and evaluate the proposed system. The approach combines software engineering methods with empirical validation through technical performance measurements and user usability assessment. Each component is described in a reproducible way to ensure methodological transparency rather than mere technical reporting.

The following subsections describe the research methodology used to design and evaluate the ROboMC system.

Rather than a simple technical setup, the process was structured as an empirical workflow comprising three main stages: (1) System design, where the local and GPT-based components were integrated under a unified query-processing pipeline; (2) Functional evaluation, assessing the correctness of local query matching and the reliability of fuzzy–semantic thresholds (0.40 and 0.87) using expert-validated question sets; and (3) Performance and Usability Assessment, measuring the end-to-end response latency on different hardware platforms (Raspberry Pi and laptop environments).

These steps collectively define the methodological framework adopted to verify that the implemented architecture achieves both functional accuracy and responsiveness suitable for educational use.

2.1. Software Architecture

2.1.1. System Design

This phase covers the architectural integration of the ROboMC system under a unified query-processing pipeline (as illustrated in Figure 1).

The ROboMC system was developed on the Django 5.2 framework (Django Software Foundation, Lawrence, KS, USA) in Python 3.11 (Python Software Foundation, Wilmington, DE, USA), widely used for educational web applications due to its modularity and stability. The database includes two main models: LocalQuestion, which stores validated question–answer pairs, and AutomaticQuestion, which records user-submitted questions and generated responses, facilitating subsequent analysis. Local matching uses a combination of normalization and fuzzy matching, complemented by a semantic encoder (SentenceTransformer—paraphrase-multilingual-MiniLM-L12-v2) that enables the detection of similarities between questions even when phrasing differs. When no sufficient match is found (threshold ~ 0.4 for fuzzy and ~ 0.87 for semantic similarity), the system forwards the request to ChatGPT (GPT-3.5-Turbo), configured to respond exclusively in Romanian and to provide a warning message regarding the informational nature of the content. GPT-3.5-Turbo was selected due to its balance of response quality, latency, and cost efficiency. In preliminary tests, the system produced accurate and contextually relevant responses with an average end-to-end latency below 3 s on Raspberry Pi clients and under 2.5 s on laptops.

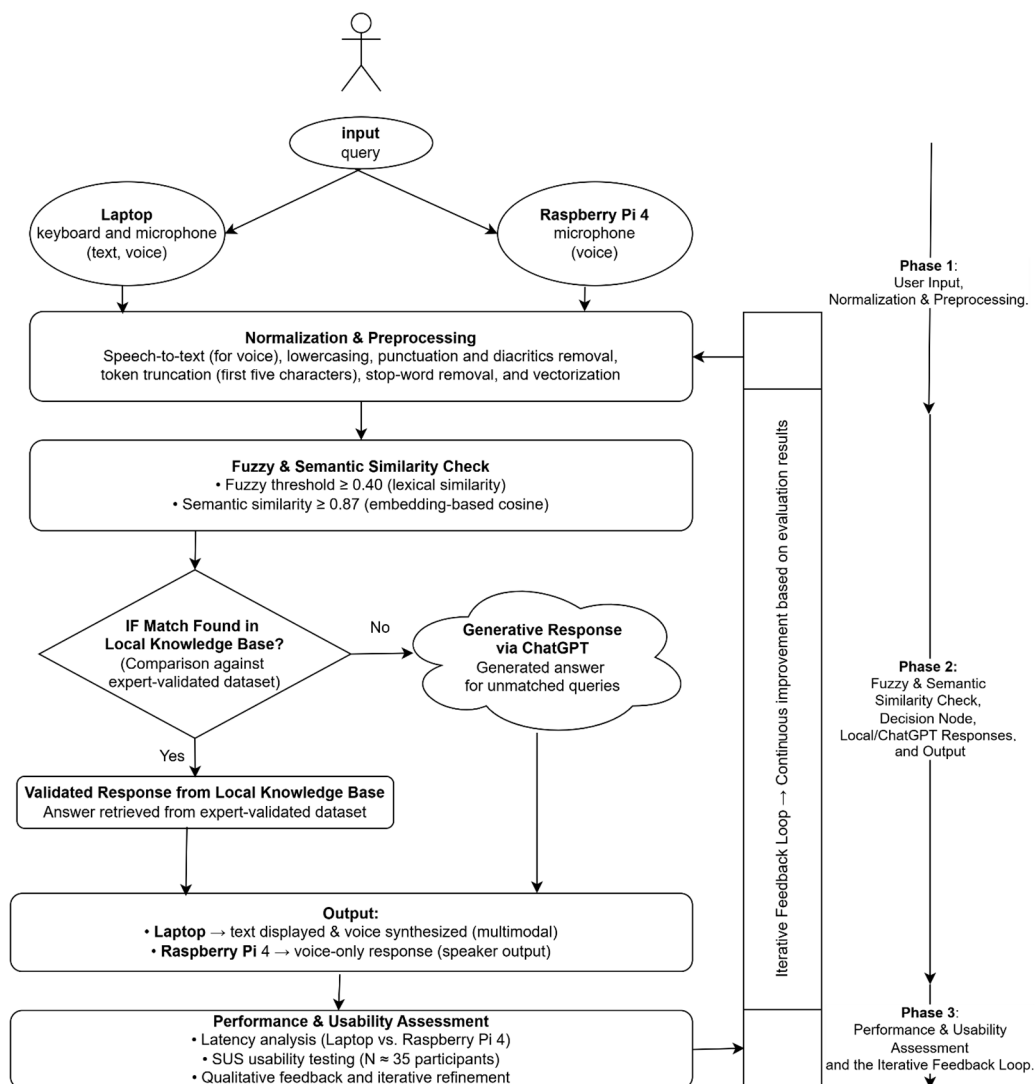


Figure 1. Methodological and system workflow of the ROboMC platform. The figure integrates the three methodological phases (System Design, Functional Evaluation, and Performance & Usability Assessment) together with the technical flow (multimodal input, normalization, dual similarity thresholds—fuzzy ≥ 0.40 ; semantic ≥ 0.87 —decision node, and output).

This latency measure includes the entire processing chain—local preprocessing and normalization, the API request to GPT-3.5-Turbo, post-processing of the generated text, and optional text-to-speech synthesis.

Such performance was found acceptable for educational use, where conversational pacing is less critical than response reliability. The more recent GPT-4 models were evaluated but rejected for this deployment due to higher latency and computational cost that would limit real-time classroom interaction. To mitigate the risk of misinformation, ROboMC clearly distinguishes between expert-validated and AI-generated content. Every time ChatGPT provides a response, the interface displays a visual disclaimer and a verbal reminder indicating that the content is not validated by medical professionals. The system also logs these responses for manual review, enabling instructors to correct or replace inaccurate outputs in subsequent database updates. This feedback loop minimizes the propagation of potentially incorrect information while fostering responsible AI-assisted learning. The local knowledge base (*LocalQuestion*) is curated through the Django administration interface, allowing domain experts to continuously expand and validate its content. Each record stores a set of canonical questions, associated keywords, and a validated answer.

New entries are first drafted based on the log of real user queries (*AutomaticQuestion* table) and are subsequently reviewed by medical faculty for accuracy and pedagogical relevance before inclusion. Periodic updates ensure that the local database remains aligned with current curricular content. For deployments outside healthcare, subject-matter experts from the respective domain perform the same validation and update process, maintaining consistent quality standards across disciplines.

The methodological workflow adopted in this study is illustrated in Figure 1.

Each block in this diagram represents a distinct methodological stage—data acquisition, normalization, fuzzy-semantic similarity computation, decision routing, and multimodal output—ensuring full reproducibility of the workflow.

The decision block illustrated in Figure 1 performs the routing of each user query based on combined lexical and semantic similarity scores.

After normalization and preprocessing, the system computes both a fuzzy lexical similarity (token overlap and Levenshtein ratio) and a semantic similarity (cosine distance between multilingual sentence embeddings).

If both scores exceed the predefined thresholds ($S_f \geq 0.40$ and $S_{sem} \geq 0.87$), the response is retrieved directly from the validated local knowledge base; otherwise, the query is forwarded to the GPT-3.5-Turbo API for generative completion.

The resulting answer is stored in the *AutomaticQuestion* table to improve future responses and can also be converted to speech for multimodal output. The architecture supports both text and voice input, ensuring a complete multimodal interaction flow.

2.1.2. Functional Evaluation

This phase focuses on verifying the reliability of local query-matching through the combined fuzzy-semantic similarity mechanism.

The normalization process includes lowercasing, punctuation and diacritics removal, and token truncation to the first five characters to improve matching tolerance for morphological variants. Fuzzy matching is implemented through a token overlap coefficient combined with Levenshtein ratio similarity to identify near matches between query tokens and stored question patterns. The fuzzy similarity threshold of approximately 0.4 and the semantic similarity threshold of 0.87 were empirically selected after iterative testing with the dataset of stored and user-generated questions, balancing precision and recall to minimize false positives. A speech-to-text module (Google Web Speech API via the Python SpeechRecognition library, language = “ro-RO”) handles voice input, feeding the same retrieval-generation pipeline used for typed queries.

To ensure consistent lexical and semantic processing, all incoming queries are normalized through a sequence of preprocessing steps including lowercasing, punctuation and diacritics removal, and truncation of each token to its first five characters. Let q denote the user query and d a stored canonical question; after normalization, token sets T_q and T_d are derived.

A fuzzy lexical similarity score is computed as $S_f = \alpha S_o + (1 - \alpha)S_l$, where $S_o = |T_q \cap T_d| / \min(|T_q|, |T_d|)$ represents token overlap and S_l the normalized Levenshtein ratio between strings ($\alpha = 0.5$). If $\min(|T_q|, |T_d|) = 0$, then S_o is defined as 0 to avoid division by zero.

The Levenshtein ratio is computed in its normalized form as $S_l = 1 - (\text{LevDist} / \max(|x|, |y|))$, where LevDist denotes the Levenshtein edit distance between the two strings, ensuring that both S_o and S_l are bounded within $[0, 1]$.

The weighting coefficient $\alpha \in [0, 1]$ controls the relative contribution of token overlap and string similarity; in this study, $\alpha = 0.5$ was selected empirically to balance the two measures.

In parallel, a semantic similarity score S_{sem} is calculated using multilingual sentence embeddings (MiniLM-L12-v2) and cosine similarity, with values close to 1 indicating stronger semantic correspondence.

A query is accepted as a valid match when $S_f \geq 0.40$ and $S_{\text{sem}} \geq 0.87$, thresholds determined empirically from expert-validated test data to balance false matches and omissions.

These threshold values were selected after iterative testing on approximately $N \approx 200$ validated query pairs, achieving an optimal trade-off between precision and recall.

This combined fuzzy–semantic evaluation ensured a high level of precision in the system’s functional testing phase.

2.1.3. Performance and Usability Assessment

This phase assessed the overall performance and usability of the ROboMC system across different hardware configurations.

In preliminary tests, the system produced accurate and contextually relevant responses with an average end-to-end latency below 3 s on Raspberry Pi clients and under 2.5 s on laptops. This latency measure includes the entire processing chain—local preprocessing and normalization, the API request to GPT-3.5-Turbo, post-processing of the generated text, and optional text-to-speech synthesis.

Such performance was found acceptable for educational use, where conversational pacing is less critical than response reliability. The more recent GPT-4 models were evaluated but rejected for this deployment due to higher latency and computational cost, which would limit real-time classroom interaction.

To mitigate the risk of misinformation, ROboMC clearly distinguishes between expert-validated and AI-generated content. Every time ChatGPT provides a response, the interface displays both a visual disclaimer and a verbal reminder indicating that the content is not validated by medical professionals.

The system also logs these responses for manual review, enabling instructors to correct or replace inaccurate outputs in subsequent database updates. This feedback loop minimizes the propagation of potentially incorrect information while fostering responsible AI-assisted learning.

User satisfaction and perceived usability were additionally assessed through the System Usability Scale (SUS) questionnaire, administered to 35 participants. Feedback from these users was incorporated into iterative refinements of the platform, supporting continuous improvement of both the system’s responsiveness and its interaction quality.

2.2. Multimodal Interaction

To provide both text and voice interaction, ROboMC integrates speech recognition and speech synthesis modules. The term “multimodal” refers to the system’s dual-channel interaction combining visual/textual and auditory modes. In its current version, ROboMC supports both text- and voice-based input and provides synchronized text and audio output.

Users may alternate between typing and speaking their queries, but only one input channel is active at a time. When deployed on laptops or desktops, ROboMC enables both data-entry modalities—typing or speaking. Users can type their question in the text box and press the button labeled “Trimite întrebarea. . .” (“Send the question”), or they can click the button “Click și vorbește” (“Click and speak”) to activate voice input. In the latter case, the spoken query is captured through the microphone, transcribed into text, and automatically displayed in the same text box before being processed by the system. Regardless of the input mode, the response is provided simultaneously as on-screen text and synthesized speech, ensuring consistent dual-channel output. In contrast, when running on a Raspberry Pi device, ROboMC functions as a standalone educational kiosk without a

display: interaction occurs solely through voice input via the microphone, and responses are delivered exclusively as audio through speakers. This configuration still ensures full multimodal communication—dual-channel input options and dual-channel output—consistent with accepted definitions in human–computer–interaction research. Voice input is captured through the microphone and converted to text using the SpeechRecognition library configured with the Google Web Speech API (language = “ro-RO”); the resulting text then follows the same normalization and matching stages as typed input. For output, responses generated by the system are simultaneously displayed on screen and synthesized into speech using the gTTS package, which produces MP3 audio streamed to the client. This design ensures consistent processing for both input modes while maintaining full compatibility with standard open-source libraries.

Temporary audio files are deleted after playback, with only the transcribed text and responses retained in the database. In parallel, the web interface supports traditional text-based interaction, providing both flexibility and inclusion for users with diverse accessibility needs. The system offers native support for displaying mathematical formulas (LaTeX via MathJax, MathJax Consortium, USA) and code snippets (with syntax highlighting), demonstrating its potential for interdisciplinary adaptation. In addition to the main response—whether retrieved from the validated database or generated by ChatGPT—the system also integrates relevant external links (extracted through a Google search) and, whenever possible, a short excerpt from a trusted source. This functionality is implemented through a dedicated module (retrieving up to five results, extracting titles and representative snippets) and is displayed in the interface as a “Supplementary Information” section. Figure 2 shows the ROboMC application interface for entering questions, while Figure 3 illustrates a generated response containing text, mathematical formulas, and external links for further documentation.

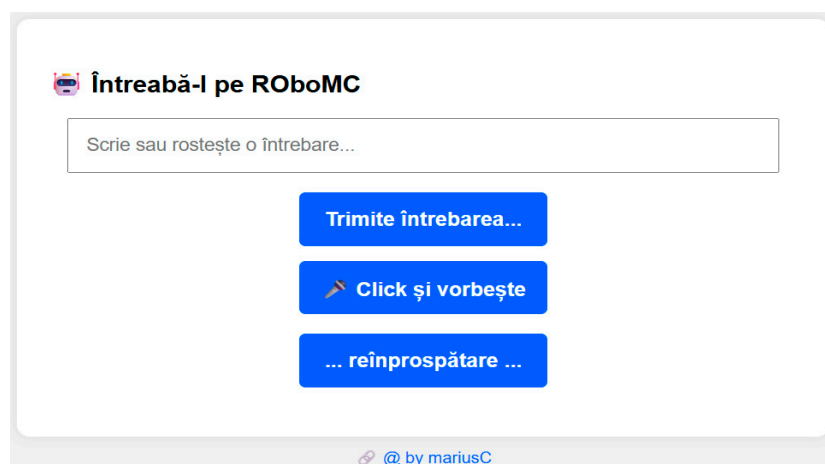


Figure 2. The ROboMC application interface for submitting queries in written or spoken form. Note: The interface labels appear in Romanian (e.g., “Trimite întrebarea...” — “Submit the question...”, “Click și vorbește” — “Click and speak”, “... reînprospătare...” — “Refresh...”) as they are part of the actual ROboMC system.

 **Întreabă-l pe ROboMC**

Rog ecuația Cockcroft-Gault

Trimite întrebarea...

 Click și vorbește

... reinprospătare ...

Răspuns:

Formula Cockcroft-Gault este utilizată pentru estimarea ratei de filtrare glomerulară (GFR) bazată pe creatinina serică, vârstă, greutate și sexul pacientului. Ea este exprimată astfel:

$$\text{GFR}_{\text{bărbați}} = \left(\frac{(140 - \text{vârstă}) \times \text{greutate}}{72 \times \text{creatinină serică}} \right) \times 0.85$$

$$\text{GFR}_{\text{femei}} = \left(\frac{(140 - \text{vârstă}) \times \text{greutate}}{72 \times \text{creatinină serică}} \right) \times 0.85 \times 0.85$$

unde GFR reprezintă rata de filtrare glomerulară în ml/min, vârsta se exprimă în ani, greutatea în kg, iar creatinina serică în mg/dL. Rezultatul este corectat în funcție de sex, unde 0.85 este factorul de corecție pentru bărbați și 0.85 * 0.85 (adică 0.72) pentru femei.

(răspuns generat de ROboMC +)

 **Informații suplimentare:**

- [Creatinine Clearance \(Cockcroft-Gault Equation\)](#)
- [Cockcroft-Gault Formula | National Kidney Foundation](#)
- [MedQuest.ro](#)
- [Gastroenterology Education and CPD for trainees and specialists » Creatinine Clearance \(Cockcroft-Gault Equation\)](#)
- [Calculators: Creatinine Clearance Estimate by Cockcroft-Gault Equation \(SI units\)-MSD Manual Professional Edition](#)

Figure 3. Example of a generated response displaying the complete Cockcroft–Gault equation, illustrating the chatbot’s ability to process and present mathematical formulas. Note: The interface labels appear in Romanian (e.g., “Rog ecuația Cockcroft–Gault”—“Request the Cockcroft–Gault equation”, “Trimitere întrebare”—“Submit question”, “Răspuns”—“Answer”) as they are part of the actual ROboMC application used in testing.

2.3. Platforms and Hardware Architecture

The system runs natively on laptops/desktops (Windows, Linux, macOS) and can also operate on a Raspberry Pi 4B [11], equipped with a USB microphone, portable speaker, and battery power (power bank). This portability transforms the application into a “portable educational kiosk,” usable not only in classrooms or laboratories but also in libraries, communities, or clinical simulations.

Preliminary evaluation protocol. The evaluation included two components:

- Technical measurements—response times were recorded for validated and generative answers, both on laptop and Raspberry Pi.

Usability assessment—35 students and faculty members completed the System Usability Scale (SUS) questionnaire [12] (Taylor & Francis, London, UK), a standard usability evaluation tool consisting of 10 Likert-scale items, with a final score ranging from 0 to

100. The questionnaire included ten standard statements, alternating between positive and negative phrasing, such as: “I found the system easy to use,” “I think the system is unnecessarily complex,” and “I would imagine that most people would learn to use this system very quickly.” Participants included undergraduate and postgraduate students (aged 21–38) and faculty members from medical and engineering backgrounds, all with basic to intermediate IT literacy. The evaluation scenario consisted of solving a short quiz using ROboMC on a laptop and on a Raspberry Pi, recording latency, clarity of responses, and perceived usability. Scores above 68 are considered satisfactory, while scores above 80 indicate an excellent level of usability.

The evaluation was conducted in two complementary sessions. The first session consisted of a technical test designed to measure response latency across 20 predefined queries executed on both devices (laptop and Raspberry Pi). The second session involved 35 participants—20 undergraduate students in health sciences and 15 faculty members—who interacted with ROboMC for approximately 20 min each, completing predefined learning tasks through text or voice input. In addition to the SUS questionnaire, qualitative feedback on clarity, engagement, and perceived reliability was collected through open-ended questions. The resulting comments were analyzed to identify usability patterns and improvement priorities.

The results are summarized in Table 1, which presents the measured latencies and the obtained SUS scores.

Table 1. Results of the preliminary evaluation (latency and SUS usability).

Response Type	Laptop (ms/s)	Raspberry Pi (ms/s)	Observations on Usage
Validated response	120–180 ms	250–350 ms	Very short time, perceived as “instantaneous”
Generative response	2.5–3.5 s	4–6 s	Acceptable for educational scenarios
SUS score (N ≈ 35)	80/100	-	“Excellent” usability level (≥ 80)

3. Results

For comparative analysis, ROboMC was evaluated against two publicly available conversational systems commonly used in educational contexts: ChatGPT Web and Replika. The comparison considered three dimensions: (i) Response latency, defined as the total time from user query submission to answer display; (ii) Factual consistency, assessed by the proportion of domain-valid responses verified by subject experts; and (iii) Pedagogical relevance, reflecting the alignment between generated explanations and curricular content.

On a representative set of student questions, ROboMC achieved substantially faster local responses (average 0.18 s for validated answers) and maintained factual accuracy through its expert-curated knowledge base.

In contrast, both ChatGPT Web and Replika demonstrated higher conversational fluency but lacked domain validation mechanisms, occasionally producing factually inconsistent answers.

These results confirm the advantage of ROboMC’s hybrid design in structured educational environments, combining verifiable accuracy with adaptive language generation.

The preliminary evaluation of the ROboMC system confirmed its potential as a portable and inclusive educational chatbot. A structured synthesis of the preliminary evaluation results is presented in Table 2.

Table 2. Preliminary evaluation results of the ROboMC educational chatbot, highlighting accessibility, personalization, portability, and user acceptability.

Dimension	Main Description
Accessibility	Text- and voice-based interaction enabled use even by visually impaired individuals.
Personalization	Storing questions allowed analysis of learning needs and adaptation of educational content.
Portability	The Raspberry Pi-based system enabled educational sessions in multiple contexts.
Acceptability	Users appreciated the mix between validated and AI-generated responses, as well as the response speed.

3.1. Accessibility

The integration of speech synthesis enabled efficient interactions for both typical users and those with visual impairments. Participants particularly appreciated the combination of text and voice, considering it a significant advantage compared to conventional chatbots limited to text. This result is consistent with recent literature highlighting the benefits of multimodal interaction and the potential of AI chatbots to support inclusion and accessibility [13]. Participants were able to interact through both typed and spoken queries in Romanian (ro-RO), confirming the reliability of the voice input module.

3.2. Personalization and Learning Analytics

By storing all submitted questions, ROboMC enables subsequent analysis of student interactions. Using methods such as clustering and topic modeling, recurring themes and knowledge gaps can be identified, providing teachers with useful insights for adapting educational content. This functionality confirms the system's value as a learning analytics tool, complementing similar observations reported in recent studies [14].

3.3. Portability and Implementation Flexibility

The Raspberry Pi-based prototype provided stable performance, with an operational autonomy of approximately 10–12 h on a single charge. The mobile design allowed the system to be used in diverse educational contexts, including classrooms, libraries, and non-formal learning spaces. Portability expands opportunities for use in clinical simulations and community training, an advantage that differentiates ROboMC from other similar solutions.

3.4. User Acceptability

Feedback collected from students and faculty indicated a high level of acceptability. Participants emphasized:

- the balance between validated and AI-generated responses,
- the integration of external references,
- the response time considered adequate.

These results are reflected in the average SUS score of 80/100, indicating an “excellent” level of usability (≥ 80). The distribution of SUS scores reported by participants is illustrated in Figure 4.

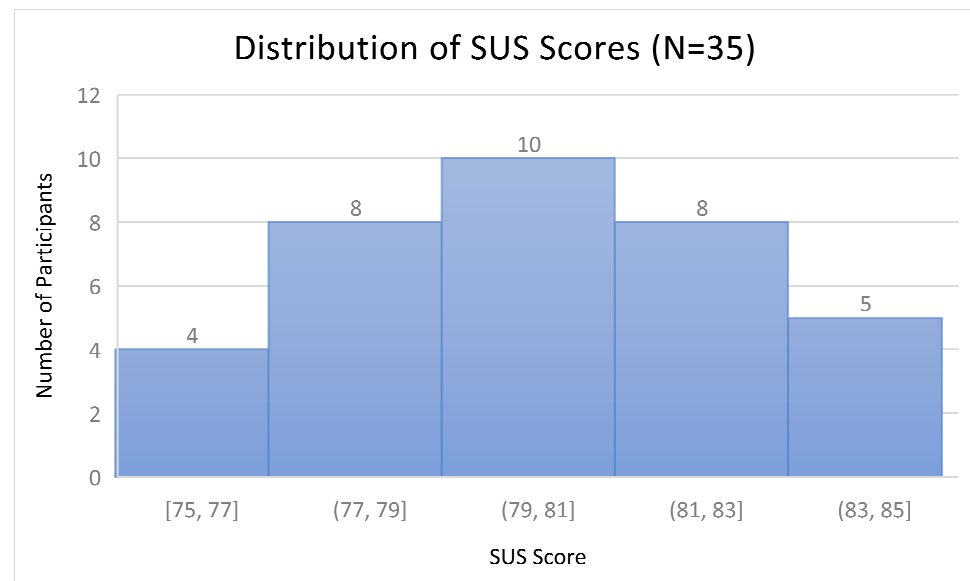


Figure 4. Distribution of System Usability Scale (SUS) scores among participants (N = 35).

3.5. Comparative Advantage

Compared with other chatbot-based solutions reported in the literature [15], ROboMC stands out through:

- the integration of a validated knowledge base,
- completion via generative AI (ChatGPT),
- multimodal interaction (text and voice),
- and hardware portability.

Together, these features strengthen the system's originality as an applied educational innovation, with potential for interdisciplinary extension. Response times compared across platforms (laptop vs. Raspberry Pi) and response types (validated vs. generative) are presented in Figure 5.

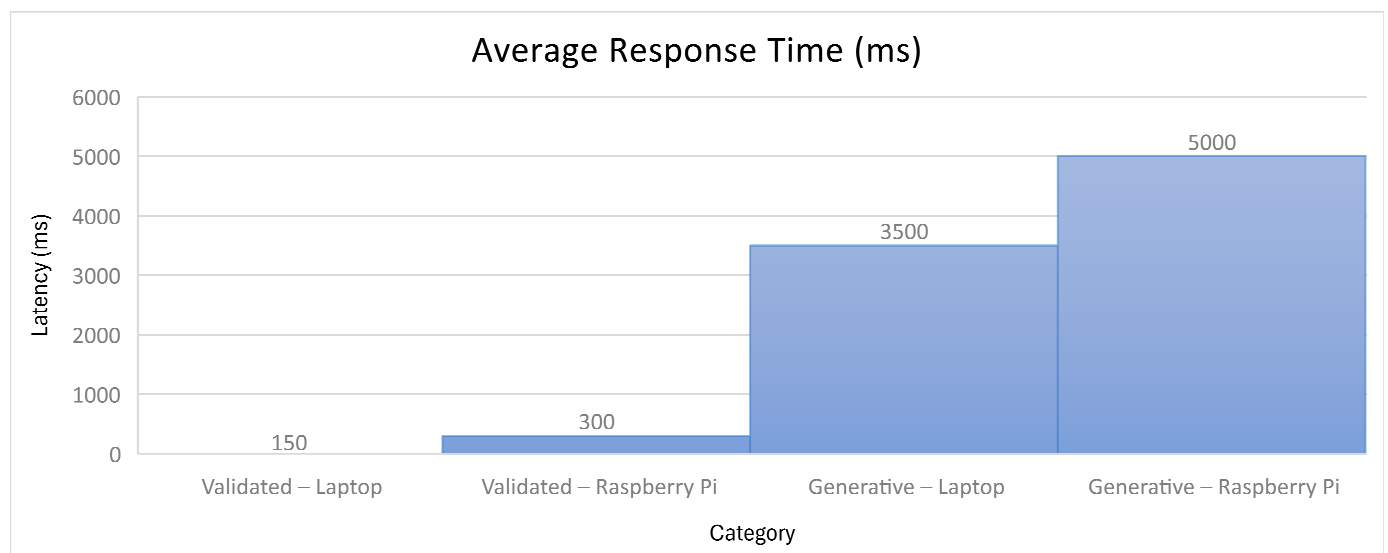


Figure 5. Average response times for validated and generative answers on laptop and Raspberry Pi.

4. Discussion

The preliminary evaluation of the ROboMC system highlights several contributions to the field of AI-based educational technologies. First, the integration of a validated knowl-

edge base with a generative model (ChatGPT) addresses one of the most persistent concerns in the literature: the lack of reliability of purely AI-generated responses. By combining validated with generative content, ROboMC ensures both accuracy and adaptability, a hybrid approach that increases user trust [13,16].

Second, the adoption of multimodal interaction (text and voice) significantly improves accessibility and inclusion. This aligns with recent results published in *Applied Sciences*, where Rienties et al. showed that students perceive AI digital assistants (AIDAs) as effective educational partners [5], and Chen et al. demonstrated that voice-based digital technologies enhance engagement through more natural interaction [10]. ROboMC extends these approaches by implementing a portable chatbot that is accessible and adapted even for visually impaired users.

Another distinctive aspect is the learning analytics component. By storing and analyzing user questions (clustering, topic modeling), the system provides actionable data for teachers, who can detect knowledge gaps and adjust curricula. This feature connects the domains of educational data mining and AI tutoring systems, confirming the interdisciplinary relevance of the system.

Portability represents a major advantage. Deployment on a Raspberry Pi 4B (8 GB RAM, 32 GB microSD) transforms the system from a digital solution into a mobile educational kiosk, usable in clinical simulations, libraries, or community environments where mobility and autonomy are essential. However, the evaluation revealed higher latencies on Raspberry Pi compared to laptops. This difference can be explained by the limitations of the ARM architecture compared to x86 processors, slower microSD read/write speeds, and the absence of dedicated software optimizations. Even with the 8 GB RAM configuration, these constraints confirm that Raspberry Pi is more suitable for demonstrative and mobile scenarios, while laptops remain the optimal platform for intensive educational sessions.

Despite these strengths, some limitations must also be acknowledged. Dependence on an internet connection for ChatGPT-generated responses restricts full offline use, and the inherent variability of generative content raises consistency concerns. Although the validated database mitigates these risks, further improvements are necessary to enhance robustness.

Future research will extend the usability evaluation to participants from non-medical domains and to accessibility testing for users with visual or auditory impairments. Future development directions include:

- integrating offline large language models (LLMs) for operation without internet access,
- developing educator dashboards to visualize learning analytics data,
- expanding usability testing to larger and more diverse cohorts,
- applying the system across domains beyond healthcare, including engineering, social sciences, and vocational training.

In Table 3, the identified strengths, main limitations, and future development directions for ROboMC are presented.

While our evaluation focused on medical education, the architecture is domain-agnostic. The curation workflow and matching pipeline generalize to other courses (e.g., programming labs, engineering fundamentals, museum guides). In each case, domain experts validate the local entries, while LLM outputs remain clearly disclaimed. This supports safe adoption across disciplines with minimal integration effort.

One current limitation of the prototype is the delay introduced by the gTTS engine, as the system must generate and play an .mp3 file before playback. Future versions may integrate real-time audio streaming (e.g., gTTS over gRPC) to reduce latency and improve naturalness of spoken responses.

Table 3. Strengths, limitations, and future directions for the ROboMC educational chatbot.

Category	Description
Strength	Hybrid system (validated base + ChatGPT); multimodal interaction (text + voice); educational memory for learning analytics; portable hardware for deployment in diverse contexts.
Limitations	Dependence on internet for AI-generative responses; variability in the accuracy and consistency of generative content; lower performance on Raspberry Pi compared to laptop.
Future directions	Integration of offline LLMs; development of educator dashboards; large-scale usability testing; expansion of interdisciplinary applications.

4.1. System Extensibility and Maintainability

The modular structure of ROboMC facilitates easy expansion into new domains and languages. The local knowledge base and matching modules can be adapted to different educational contexts (e.g., programming, nursing, or engineering) simply by updating the expert-validated database. Planned developments include multilingual interfaces, offline operation through locally hosted LLMs, and cloud-based deployment for collaborative use. These features ensure that the system remains maintainable and scalable beyond its initial application in medical education.

4.2. Pedagogical Implications and Transferability

Beyond its technical implementation, ROboMC has broader pedagogical implications for integrating AI-based dialogue systems in higher education. The hybrid validation–generation model promotes active learning by allowing students to explore validated knowledge while reflecting critically on AI-generated information. This dual interaction supports the development of digital literacy and critical evaluation skills—competencies increasingly emphasized in modern educational frameworks. Furthermore, the system’s modular architecture facilitates replication across disciplines, enabling instructors to build domain-specific repositories while maintaining consistent interaction paradigms. Such scalability is essential for institutions aiming to incorporate conversational AI responsibly into their curricula.

In addition to its technical versatility, ROboMC demonstrates significant pedagogical value by fostering a reflective learning process. During pilot testing, instructors observed that students tended to compare validated answers from the local knowledge base with generative explanations, which stimulated curiosity and critical reasoning. This type of dual interaction helps learners move beyond passive information retrieval toward analytical thinking, strengthening both domain competence and awareness of AI limitations. Such findings support the role of hybrid systems like ROboMC as effective mediators between structured educational content and open-ended exploration.

Beyond its classroom integration, ROboMC offers additional opportunities for institutional deployment and social inclusion. Because the system can operate independently on low-cost hardware such as the Raspberry Pi, it can be used in environments with limited digital infrastructure, supporting equal access to educational innovation. This portability also enables on-site demonstrations and field learning, making the tool suitable for blended or community-based training programs.

From a pedagogical management perspective, the system’s data-logging features provide instructors with actionable analytics about student interactions, common misconceptions, and question frequency. These insights can inform adaptive course design and personalized feedback strategies, extending the value of ROboMC beyond a learning interface into a research instrument for educational improvement.

Finally, by explicitly differentiating between validated and generative responses, the platform encourages transparency and responsible AI use, aligning with emerging European guidelines for trustworthy AI in education. This approach reinforces ethical awareness while maintaining engagement and curiosity—two fundamental pillars of sustainable AI-assisted learning.

5. Conclusions

This study presented ROboMC, a portable and multimodal educational chatbot that combines a validated knowledge base with a generative model (ChatGPT), providing both accuracy and flexibility. Preliminary evaluations confirmed the system's feasibility, with acceptable latencies and an average SUS score of 80/100, indicating an excellent level of usability.

The integration of validated and generative responses, together with multimodal support (text and voice), extends accessibility for users with diverse needs and strengthens trust in educational interactions. Portability on platforms such as Raspberry Pi demonstrates the system's applicability in various contexts, from classrooms and simulation laboratories to libraries and communities.

Although validated in a case study focused on health education, the system's architecture is domain-independent and can be adapted to engineering, social sciences, or other disciplines through updates to the validated database. Limitations related to internet dependency and variability of AI-generated responses can be addressed by integrating offline language models and developing advanced analytics tools for educators.

Overall, this work reflects an iterative, human-centered design approach rather than an abstract theoretical exercise. Each development stage—from the creation of the validated question bank to the usability study—was guided by direct classroom observations and user feedback. Future work will continue to emphasize authentic academic integration by involving educators and learners in co-design sessions aimed at refining interaction dynamics, ensuring that technological innovation remains grounded in pedagogical practice. Looking ahead, the ongoing evolution of ROboMC will focus on expanding its educational adaptability and research utility. Planned updates include integrating multilingual support, adaptive response strategies based on learner profiles, and tools for longitudinal tracking of learning progress. In parallel, future research will explore how validated and generative interactions shape students' reasoning patterns and motivation. Establishing collaborations with educational institutions from other disciplines will further test the scalability of the platform in engineering, social sciences, and environmental studies. Through these developments, ROboMC can contribute not only as an educational tool but also as a reference framework for responsible, human-centered AI integration in academic contexts.

In conclusion, ROboMC provides a scalable, interdisciplinary, and inclusive framework for AI-assisted education, representing an important step toward combining personalized learning with portability and technological reliability.

Author Contributions: Conceptualization, A.-L.C. and M.C.; methodology, A.-L.C.; software, M.C.; validation, A.-L.C. and M.C.; investigation, A.-L.C.; writing—original draft preparation, M.C.; writing—review and editing, M.C.; supervision, A.-L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by Lucian Blaga University of Sibiu through the research grant LBUS-IRG-2022-08.

Institutional Review Board Statement: The study procedure and instruments were approved by the Ethics Committee of the Lucian Blaga University of Sibiu, Romania (approval code NR.02-14.07/2022, approval date on 14 July 2022).

Informed Consent Statement: Informed consent was obtained from all participants involved in the satisfaction survey. All participants were informed about the purpose of the study and their rights before taking part in the evaluation. Written informed consent was obtained from all participants prior to their inclusion in the study. The consent form used was anonymous and contained no personal or identifying information.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: Author Adriana-Lavinia Cioca was employed by the company CMI Cioca Adriana-Lavinia. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Zawacki-Richter, O.; Marín, V.I.; Bond, M.; Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 39. [CrossRef]
2. Winkler, R.; Soellner, M. Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Acad. Manag. Proc.* **2018**, *2018*, 15903. [CrossRef]
3. Holmes, W.; Bialik, M.; Fadel, C. *Artificial Intelligence in Education*; Center for Curriculum Redesign: Boston, MA, USA, 2019; ISBN 978-1-79431-360-7.
4. Labadze, L.; Grigolia, M.; Machaidze, L. Role of AI chatbots in education: Systematic literature review. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 56. [CrossRef]
5. Rienties, B.; Tempelaar, D.; Nguyen, Q.; Littlejohn, A. Students' Perceptions of AI Digital Assistants (AIDAs): Should Institutions Invest in Their Own AIDAs? *Appl. Sci.* **2025**, *15*, 4279. [CrossRef]
6. Smutny, P.; Schreiberova, P. Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Comput. Educ.* **2020**, *151*, 103862. [CrossRef]
7. Boustani, N.M.; Sidani, D.; Boustany, Z. Leveraging ICT and Generative AI in Higher Education for Sustainable Development: The Case of a Lebanese Private University. *Adm. Sci.* **2024**, *14*, 251. [CrossRef]
8. Silva, C.A.G.; Ramos, F.N.; de Moraes, R.V.; Santos, E.L. ChatGPT: Challenges and Benefits in Software Programming for Higher Education. *Sustainability* **2024**, *16*, 1245. [CrossRef]
9. Wangsa, K.; Karim, S.; Gide, E.; Elkhodr, M. A Systematic Review and Comprehensive Analysis of Pioneering AI Chatbot Models from Education to Healthcare: ChatGPT, Bard, Llama, Ernie and Grok. *Future Internet* **2024**, *16*, 219. [CrossRef]
10. Chen, Y.; Li, H.; Zhang, X. Digital Human Technology in E-Learning: Custom Content Solutions. *Appl. Sci.* **2025**, *15*, 3807. [CrossRef]
11. Raspberry Pi Foundation. Raspberry Pi 4 Model B Specifications. Raspberry Pi. Available online: <https://www.raspberrypi.com/products/raspberry-pi-4-model-b> (accessed on 10 September 2025).
12. Brooke, J. SUS: A 'quick and dirty' usability scale. In *Usability Evaluation in Industry*; Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L., Eds.; Taylor & Francis: London, UK, 1996; pp. 189–194.
13. Schei, O.M.; Møgelvang, A.; Ludvigsen, K. Perceptions and use of AI chatbots among students in higher education: A scoping review of empirical studies. *Educ. Sci.* **2024**, *14*, 922. [CrossRef]
14. Chen, X.; Xie, H.; Qin, S.J.; Wang, F.L.; Hou, Y. Artificial Intelligence-Supported Student Engagement Research: Text Mining and Systematic Analysis. *Eur. J. Educ.* **2025**, *60*, e70008. [CrossRef]
15. Okonkwo, C.W.; Ade-Ibijola, A. Chatbots applications in education: A systematic review. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100033. [CrossRef]
16. Lo, C.K. What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* **2023**, *13*, 410. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.