Design of Medical History Collection Chatbot with Virtual Avatar Integration

Abstract: This project describes the development and implementation of an AI-based medical chatbot. The chatbot uses Large Language Models (LLM) for medical history collection conversations to provide a natural and smooth communication experience. It also introduces virtual avatars in the chatbot to provide a more immersive experience, hoping to enhance the overall user experience while encouraging users to respond. The paper outlines the development goals, design concepts, and the development and testing process of the initial prototype of the chatbot. In conclusion, this project highlights the potential of AI-based chatbots in healthcare to not only effectively bridge the communication gap between patients and healthcare services but also focuses on improving the conversational efficiency of chatbots and explores the promise of avatars in this field.

1. Introduction

1.1 Background of Research

For most medical problems, medical history is the cornerstone of the diagnostic journey. Despite the increase in diagnostic tools such as advanced imaging, molecular and laboratory assays, a comprehensive medical history is necessary to guide further steps and may sometimes even be sufficient to diagnose a disease without further testing (Hampton et al., 1975). Conversely, an inadequate medical history can jeopardize patient safety (Peterson et al., 1992). However, there are a number of problems with current medical history collection, including incomplete or unclear patient descriptions, inefficient communication, subjective bias by physicians, incomplete or inaccurate documentation, and poor patient experience. These problems may lead to the omission of critical information and affect subsequent treatment decisions (Burt et al., 2017).

Recent advances in Artificial Intelligence (AI), such as Large Language Modelling (LLM), have the potential to be used in healthcare, especially in assisting doctor-patient communication (Fadhil & Schiavo, 2019). The introduction of large language modelling (LLM) to assist in medical history taking has significant advantages. Firstly, LLM has the advantage of understanding the situation and achieving the goal (Grice et al., 2017), and natural language processing (NLP) technology enables dynamic dialogue, effective responses to patient statements, and generation of precisely asked questions (Athota et al., 2020). As for medical professionals, LLM can structure and organise patient descriptions to generate reports to assist doctors. In addition, LLM can standardise the process of history taking to ensure comprehensiveness and consistency, and avoid errors due to subjective bias (Hausberg et al., 2012). Therefore, using LLM to assist in medical history collection not only improves the efficiency and quality of consultation, but also provides solid support for the digital transformation of healthcare.

Introducing Virtual Avatar in chatbots for history collection has the opportunity to enhance user experience and interaction. Virtual avatars mimic human communication through natural speech, facial expressions, and body language, relieving users of psychological stress, especially in answering sensitive questions. Its ability to enhance immersive experiences, encourage people to respond, and provide users with a great user experience has been demonstrated in areas such as education, entertainment, and services (Kurniawan et al., 2015). It has also been shown to facilitate patient disclosure in the healthcare field (Stock et al., 2023), and its application to medical history collection may help to improve data quality and user satisfaction, and optimise the consultation process.

1.2 Research Objectives

The aim of this study is to design a history collection Chatbot that combines Virtual Avatar and Large Language Models (LLMs) to improve the efficiency of history collection while providing a good communication experience to patients. To see the potential of AI technologies such as LLM and Virtual Avatar in the medical field and the effectiveness of their application.

1.3 Research Implications

This study provides an innovative technological path for medical information collection by combining the emotional interaction capability of Virtual Avatar (Choi et al., 2020) with the natural language processing capability of LLM (Grice et al., 2017). At the research level, it promotes innovation in the field of medical artificial intelligence and human-computer interaction, and enhances the naturalness and trust of patient-system interaction through the emotional simulation of avatars and the language processing capability of LLMs (Zhen et al., 2023), and human-computer interaction provides a new research perspective in the field of medical communication. In practice, the system can efficiently collect and generate structured medical history data to reduce the workload of medical staff, while optimising the patient experience through humanised interaction design, providing an efficient, reliable and humanised solution for real medical scenarios, which is of great significance for the innovation and optimisation of medical services.

2. Literature review

2.1 Role of Chatbots in Healthcare

While the use of chatbots has increased in general, the health sector has seen particularly strong growth in recent years. Especially with the adoption of LLM technology, experts have high expectations for intelligent conversations (Fadhil & Schiavo, 2019). LLM, through its powerful natural language processing capabilities, can understand user input and dynamically generate accurate responses (Grice et al., 2017). In a study (Athota et al., 2020), chatbots, through the processing of natural language, were able to provide users with accurate medical information and medical guidance, demonstrating the ability to understand user queries and provide relevant medical guidance. A study by Medibord (Srivastava & Singh, 2020) revealed the potential for the use of chatbots in personalized healthcare solutions and their importance in improving patient outcomes and accessibility. (Kurniawan et al., 2015) research in creating a healthcare-specific e-medical bot capable of delivering timely medical assistance information through AI algorithms and real-time data demonstrates the effectiveness of chatbots in reshaping the traditional healthcare delivery model. Even if applied in different ways and performing different tasks, it is evident that chatbots have transformative potential in healthcare.

2.2 The Potential of Virtual Avatar Technology Characters and Their Design Elements

Avatar technology provides an immersive and interactive experience for users through anthropomorphic images and representations (Choi et al., 2020), which helps to build and maintain a human-chatbot relationship (Zhen et al., 2023). Avatar technology has shown potential for application in a number of fields, and in medical scenarios, avatars can enhance patient engagement and trust through natural voice, expression, and body language. In a study by (Robertson et al., 2015), the presence of avatars in service scenarios was found to have a positive effect on willingness to use. Another study (Fink et al., 2024) demonstrated the motivational effect of avatars on users. The study of (Stock et al., 2023) reveals that avatars may have a facilitating effect on user disclosure, which is a rather important element in doctor-patient communication.

In the design of avatars, visual cues are key to the effectiveness of the technology. This includes the avatar's image, movements, etc. Among them, the degree of their similarity to humans is an important factor (Kurniawan et al., 2015); in addition, the gender and social identity of the image will also affect the user's attitude towards them (Choi et al., 2020). Some studies suggest that anthropomorphism will bring about the "Valley of Terror effect," which makes humans resistant to them, but others suggest that anthropomorphism has the potential to encourage disclosure of information because they promote social bonding (Zhen et al., 2023; Robertson et al., 2015; Stock et al., 2023; Fink et al., 2024).

In conclusion, these studies shed light on the potential applications of AI-based medical chatbots in a conversational environment in healthcare. By leveraging AI techniques, NLP algorithms, and avatar technology, these chatbots offer promising solutions to a variety of healthcare challenges, including accessibility, efficiency, and personalised diagnosis. Nevertheless, further research is necessary to explore the scalability, reliability, and effectiveness of these chatbots in real-world healthcare settings.

3. Design Method

Double Diamond Design Model

In this project, I adopted the Double Diamond model (Möller, Ola, 2015), whice provides a clear framework and structured guidance for the design process. Using the "divergent-convergent" approach (Banathy, Bela H., 1996), it guided me to extensively explore doctor-patient communication issues in medical history collection during the early stages of the project and to define core needs. In the development phase, it fostered innovation and supported iterative optimization of solutions, improving the quality and efficiency of the design.

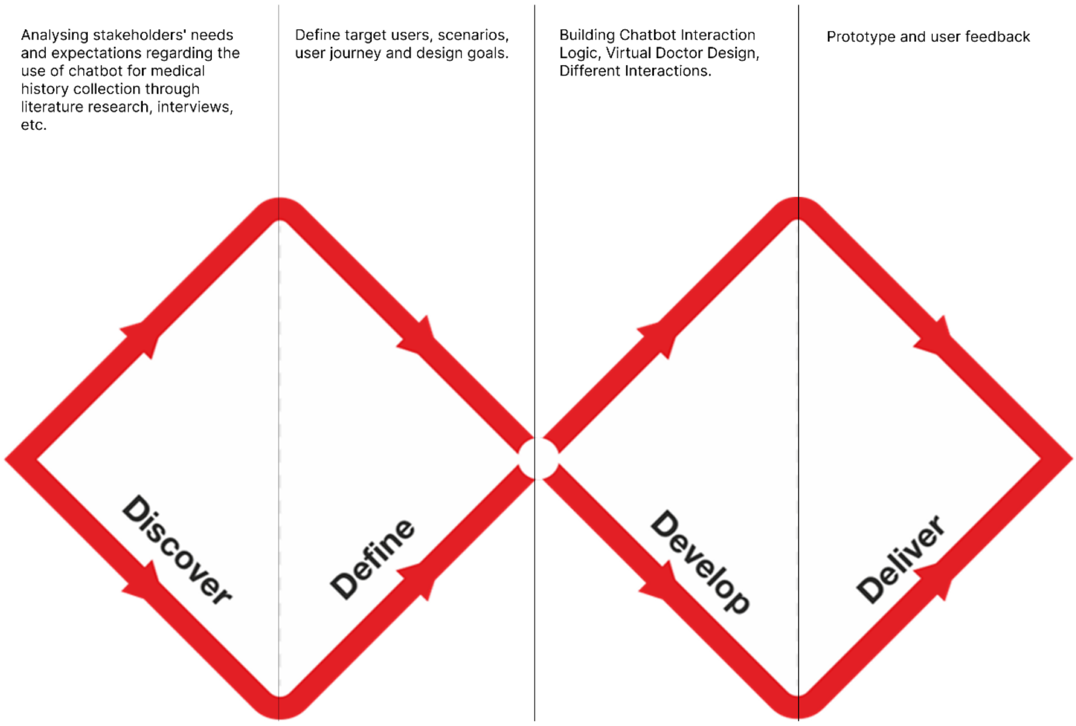

Figure 1.

In the Discover phase, I conducted literature research on doctor-patient communication and AI applications, talked to hospital experts and people about doctor-patient communication in the history taking phase, to see the gap between the ideal situation and the reality of doctor-patient communication from multiple perspectives. Thus, design goals were determined in the definition phase, and ideas for multiple interaction interfaces and interaction methods were generated. Rapid development of prototypes was carried out in the development phase to initially test the effectiveness of the prototypes and collect feedback. Then some preliminary conclusions were obtained.

4. Chatbot System Design and Demo Development

4.1 Discover and define

4.1.1 Discover

To ensure a solid foundation for the design of the chatbot, we conducted user interviews and literature review during the discovery phase.

User Interview： In the pre-project phase, we conducted a meeting with a doctor at the hospital with the aim of clarifying the doctor's requirements and expectations of the system. During the discussion with the doctor, he mentioned that he would like the system to encourage patients to describe their conditions and to allow them to express themselves actively. The data were tabulated in a way that doctors were used to and comfortable with. Therefore, we can set the patients who will interact directly with the system as the main target audience and thus define the main goals of the system design.

Literature Review： A search of the literature enabled the identification of medical expectations and standards for doctor-patient communication. The study (Berman & Chutka, 2016) clarified that in defining the effectiveness of evaluating doctor-patient communication, modern medicine is more accepting of patient-centered questioning, i.e., questioning the patient on open-ended questions, which guides the design of the conversational design in the system, as well as the features and criteria that need to be achieved in chatbot conversations.

Based on both, we constructed Persona and Scenario and identified the basic goals for the design of the system.

4.1.2 Persona

Based on the findings and design considerations, the following user roles were defined to better match the chatbot's role in medical history collection. As shown in Figure 2.

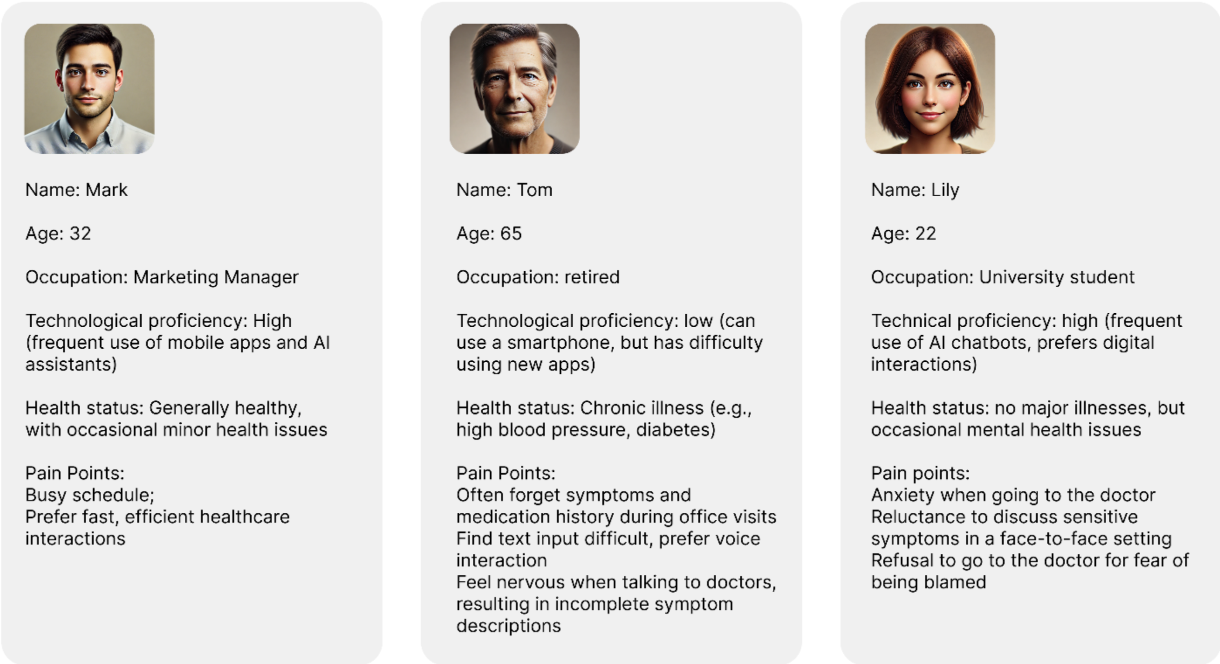

Figure 2.

4.1.3 Scenarios

The system is designed for pre-visit setups in hospitals, where the chatbot is used to complete a comprehensive history collection prior to the visit to reduce the burden of treatment. Considering that web-based digital healthcare systems are very common nowadays, the web will be set as the main platform for this system. This also enables integration with existing workflows and systems (Athota et al., 2020).

4.1.4 Design goals

a. Improve the Efficiency of History Taking

Chatbot should use patient-centered communication methods (Deveugele et al., 2005) to promote positive patient response. And A structured history report is then generated to inform the physician's diagnosis (Grice et al., 2017).

b. Providing a Good Interactive Experience

Face-to-face counselling through the use of verbal and non-verbal behaviors such as empathy and immediacy can increase trust and satisfaction between patients, thus promoting better health communication and understanding (Berman & Chutka, 2016). As well as creating a more immersive experience through the introduction of avatars, increasing user willingness to use (Stock et al. 2023).

c. Promoting Precision Medicine

Ensure the completeness and accuracy of medical history data to provide physicians with reliable decision support (Peterson et al., 1992). Encourage patients to give more comprehensive answers through multiple rounds of questioning to detect potential symptoms and hidden problems and avoid missing key information (Burt et al., 2017).

4.2 System Architecture

The chatbot system consists of three key components: user interface, dialogues management and data processing. As shown in Figure 3. Two types of user interfaces were created in this project to support text and voice interactions, respectively, and a virtual doctor avatar was used to enhance immersion. The dialogue management system integrates LLM (GPT-4) for natural language understanding, using a combination of Rule-Based Logic + LLM to ensure that the Chatbot performs its tasks accurately while responding flexibly to user input. State tracking is also used to maintain dialogue continuity in multiple rounds of interaction. Meanwhile, the data processing module includes Automatic Speech Recognition (ASR), which converts speech input into text to improve efficiency, and a structured history generation system, which organizes user input into a physician-friendly medical history report.

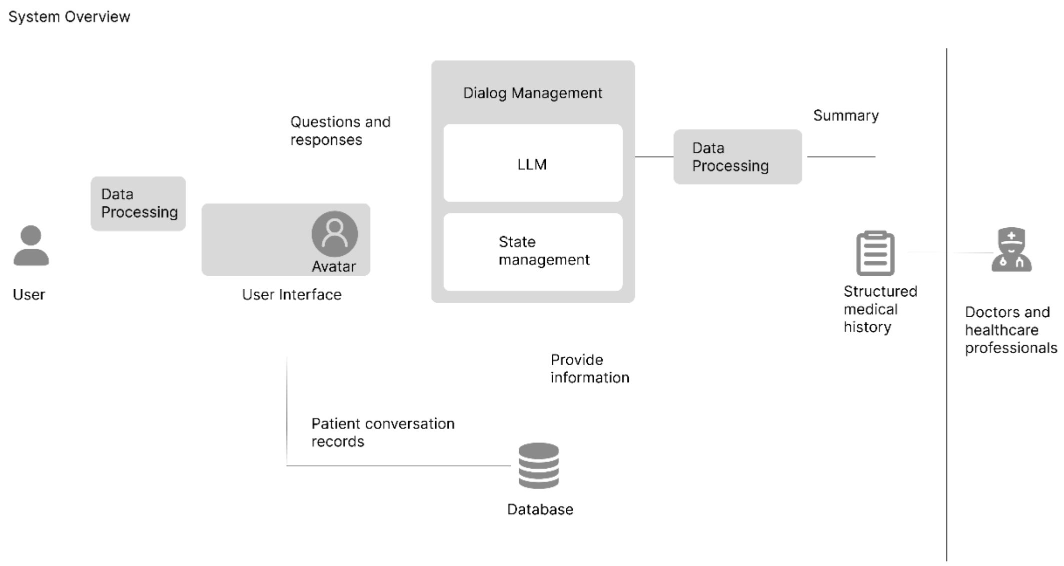

Figure 3.

4.2.1 Interface and Interaction Design

Option1: text dialog chat interface with the virtual avatar

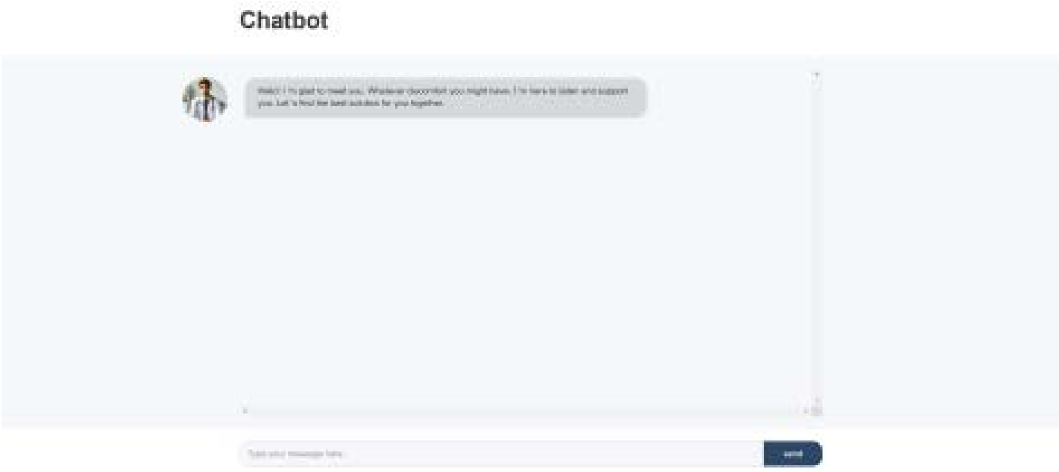

Figure 4.

The goal of the text dialogue chat window design is to provide users with a clean and efficient communication experience through text interaction. The title ‘Chatbot’ at the top of the interface indicates its function, ensuring that the user can see at a glance what the current page is for. The main area is the dialogue window, which is presented in the form of a dialogue bubble, where Chatbot's replies are presented on a left-aligned grey background and the user's inputs are presented on a right-aligned white background, which clearly distinguishes between the two sides of the conversation. The dialogue window supports scrolling through long conversations, making it easy for users to go back in time.

The input area is located at the bottom of the page and consists of a text input box with a placeholder for ‘Type your message here’ and a blue ‘Send’ button, which is highly visible and allows the user to quickly find the send function. By clicking the ‘Send’ button or pressing the Enter key, the user's input is submitted to the dialogue window and Chatbot instantly generates a response and presents it in a dialogue bubble. The overall design emphasises simplicity and intuition for users who need to express their needs in a quiet environment or with precision.

Option2: Voice Interaction Interface with the virtual avatar

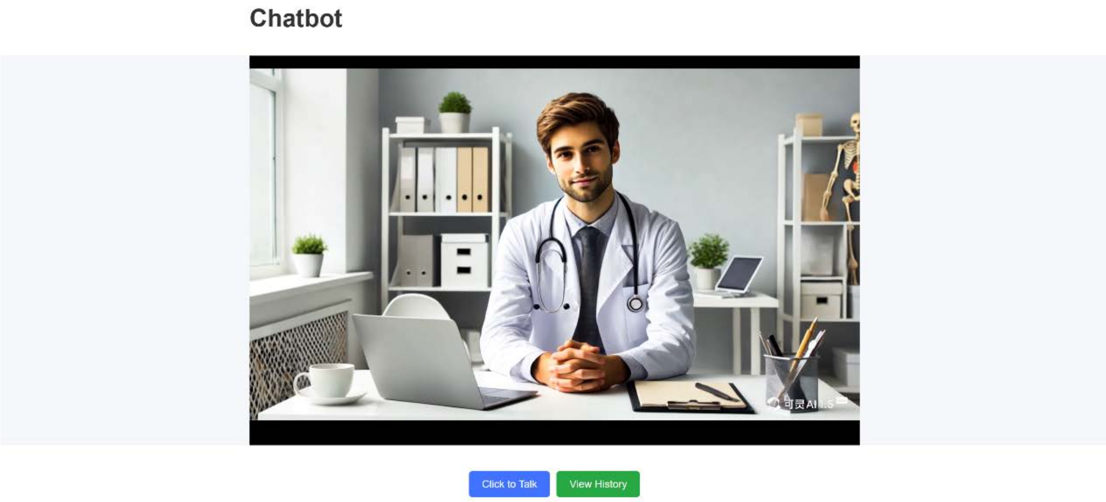

Figure 5.

The core design concept of the avatar voice interaction interface is to enhance the user's immersive experience through the dynamic representation of the avatar and voice communication. The title 'Chatbot' is also displayed at the top of the page to maintain visual consistency with the text interface. The main part is centred on the avatars, which are designed in a professional and friendly style, such as the image of a doctor, which, together with the clinic elements placed in the background, creates a believable scenario-based interactive environment. The avatar is synchronised with the user's voice input and Chatbot's voice output through dynamic facial expressions, gestures and lip-synching to enhance the realism of the dialogue.

Below the avatar area, two buttons are provided for user operation. The blue 'Click to Talk' button is used to activate the voice interaction mode. When clicked, the avatar plays a welcome message and waits for the user to input questions by voice. The green 'View History' button allows the user to view the previous dialogue history, which is suitable for scenarios that require retrospective information. The whole interface, through scenario design and voice interaction, is suitable for users who need natural communication or health consultation, providing users with a more immersive and humanised interactive experience, while ensuring a clear functional layout and easy operation.

Interaction

In option 1, the chatbot primarily communicates through text, with minimal reliance on voice interaction. The interface is text-driven, and user interactions occur mainly through typing and reading responses on the screen. The chatbot's avatar remains present but primarily serves as a visual enhancement rather than an active communication tool.

In option 2, the virtual doctor avatar takes a central role, engaging users primarily through voice interactions and facial expressions. The chatbot asks questions and responds using natural speech, while the avatar's expressions and lip-sync animations dynamically adapt to the conversation.

4.2.2 Dialog Management

The dialog flow combines rule-based logic and Large Language Modeling (LLM) for a multi-round dialog flow. The details are as follows:

1. **Session Initialization**

   At the beginning of each conversation, a System Role is set that defines the behavior and tone of the Chatbot: The Chatbot is set up as a professional health assistant with a warm and caring tone, with the goal of obtaining information about the patient's health through step-by-step questioning.

   At the same time, the system will activate state management, which is used to store the context of the conversation, including user input, emotional state and current conversation history. It is used to help the chatbot generate more accurate responses.

2. **User Input Processing**

   In this session, three or four processes are performed on the user input, including converting speech to text, providing valid information in the input, parsing the presence of emotions in the input, and adding the input to the dialogue history. If the user uses voice input, the voice is converted into text; in the text, the system identifies information related to the medical history and structures it; at the same time, the chatbot uses the OpenAI model to identify emotions in the user's input in order to form a response; and the user's input should be fully recorded in the conversational history for both the user and the doctor to view. The user's input should be recorded in the dialogue history for both the user and the doctor to view.

3. **Dynamic response**

   The chatbot's reply will consist of two parts, firstly the response to the user's reply and secondly the question based on the user's reply. Both are implemented based on State Management. Based on the current dialogue, the GPT model generates a natural language reply. To avoid lengthy replies, replies are limited to 1-2 sentences. The tone of the reply will be adjusted according to the recognized user's emotions to better match the user's emotional needs. For example, when positive emotions are recognized in the information provided by the user, the chatbot will use a positive tone such as 'I'm glad to hear that!' . When a negative emotion is recognized in the information provided by the user, chatbot will show support and concern, e.g. 'I'm sorry to hear that, I'll do my best to help you.' When no obvious emotion is recognized or a neutral emotion is recognized, chatbot will keep the response professional and clear.

The chatbot then dynamically generates questions based on the state of the conversational state, e.g., if the user mentions a 'headache', the system will further ask 'How long has it lasted?' 'Is there nausea?' etc. At the same time, the system develops a systematic history of the information needed, such as complaints, related symptoms, medical history or lifestyle habits, etc. The chatbot progressively asks questions based on these topics of information that need to be collected, in order to limit the content of the chatbot does not deviate from the scope of the history of the questioning and to ensure that the history of the information collected. When all key questions have been answered, the chatbot ends the conversation and summarizes the information.

4. **Final Summary & Report Generation**

Once all the necessary data has been collected and constructed, the system compiles it into a professionally formatted medical history report. The report is designed to be concise yet comprehensive, presenting the data in a structured format that allows the physician to quickly view the relevant information. The option to view the full patient response when necessary is also retained.

4.2.3 Data Processing

The data processing module plays a critical role in ensuring that the chatbot effectively captures, organizes, and presents user-provided medical history in a structured and usable format. It consists of two core components: Automatic Speech Recognition (ASR) and Structured Medical History Generation. Automatic Speech Recognition (ASR) is implemented to convert spoken input into text, allowing users to interact with the chatbot using natural speech rather than typing.

Once the chatbot has collected user input, it processes the data and organize it into a structured medical history report that is clinically useful to the physician. The process involves several key steps: first, symptom information is extracted from the content of the username input, key symptoms and relevant medical history are identified from the free text input; and LLM-based NLP models are used to interpret vague or ambiguous descriptions. Unstructured patient responses are then converted into a standardized medical format; in addition, medical terminology normalization ensures that lay expressions (e.g. 'I have a headache') are mapped to medically recognized terminology (e.g. 'headache'), thus improving clarity and Consistency. Finally, the final report formatting stage compiles the extracted and processed data into a structured, readable summary of the medical history, ensuring that the physician receives well-organized, actionable information.

Example:

user input	Structured medical history output
I've had a headache for three days now, mostly at bedtime. Sometimes I get nauseous.	Complaint: Headache Related symptoms: nausea Duration of symptoms: three days Pattern: recurring at bedtime

4.3 Virtual Avatar Design

Concept Design

Role Definition: Young male doctor, professional and approachable with good communication skills. The tone of voice is gentle but professional, and the behavior reflects concern for the patient.

The design of virtual doctor images plays a crucial role in shaping user perception and engagement. Studies have shown that attributes such as gender, social roles, behavioural patterns and exhibited personality affect users' trust and willingness to interact (Choi, Mehraliyev, & Kim, 2020). The goal of this project is to create a virtual doctor to enhance the realism of doctor-patient communication, facilitate natural interactions, and ultimately improve user experience and interaction with chatbots.

In terms of social identity, portraying the chatbot as a doctor can enhance the realism of the medical consultation experience. Users tend to engage more effectively when they perceive the system to be competent and responsive to their needs. (Reeves and Nass, 1996). When users perceive symbolic representations of human characteristics (e.g., virtual doctor avatars), they may respond psychologically similar to interacting with a real doctor, thus evoking trust and engagement and enhancing the credibility of the system.

The personality of the display is defined on the basis of social identity-based choices. In order to gain the user's trust, the avatar needs to convey professionalism and empathy. This is due to the fact that research (Roter, Frankel, Hall, & Sluyter, 2006) has shown that the emotional expression of doctors is crucial in healthcare, and that their warmth and caring in doctor-patient communication can alleviate patients' anxiety and fear and enhance their trust in medical decisions. Research data suggests that doctors who are able to listen to patients' needs and explain them to them are recommended and chosen by more patients. (Pękacz, Kądalska, Skoczylas, & Targowski, 2019)

The virtual assistant's male identity was chosen based on real-world healthcare demographics(World Health Organization, 2025). In most countries, men continue to make up the majority of medical doctors. Guadagno et al. (2011) found in their Second Life study that users tended to assign traditional gender roles to avatars in virtual environments. This choice was intended to align user expectations with common experiences in medical counselling. While this choice reflects existing biases and stereotypes, it does not imply endorsement of these norms, and future iterations of chatbots may explore personalization options based on user preferences.

In terms of the age of the avatar, the virtual doctor is portrayed as a young adult. Chang (2015) examined the relationship between doctor gender and facial maturity and the doctor-patient relationship, and showed that young doctors outperformed mature doctors in terms of patient expectations, satisfaction and expected loyalty. This makes them ideal candidates for digital health interactions.

Four dimensions, Visual, Dynamic effects, Background, and Sound, as shown in Figure 6, were used to construct an image of this professional, approachable medical assistant with good communication skills.

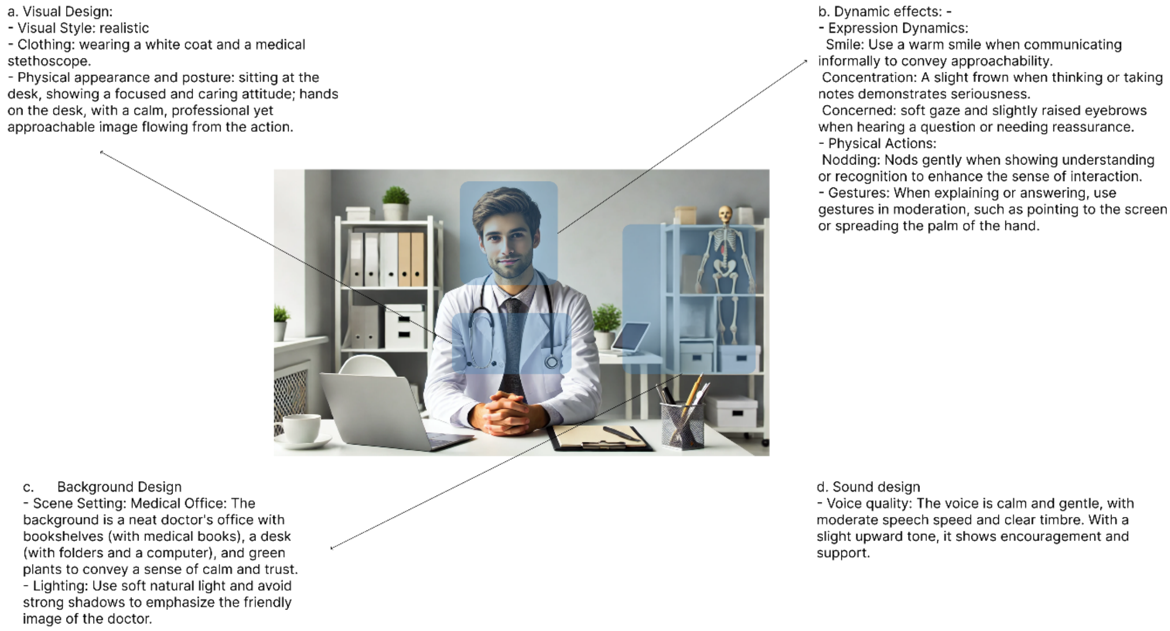

Figure 6.

4.4 Demo Development

4.4.1 LLM Selection and Configuration

a. Model Selection:

ChatGPT-4.0 is chosen as the language model in this project. The ChatGPT model has strong natural language understanding and generation capabilities and is suitable for multi-round dialog scenarios. It has a certain degree of flexibility, even if it is directly invoked, it can still generate logical answers, which is suitable for generalized medical history collection. Meanwhile, it has the advantage of rapid deployment, saving computational resources and development costs, and is suitable for rapid development.

To enhance the application of the model, ChatGPT was first provided with a clear description of the target task: “Simulate a professional doctor, collect patient history, and ask questions about relevant symptoms”. Secondly, prompts (Prompts) were entered: 'you need to be professional, gentle and caring' to make this generalized model perform more accurately in the history collection task.

b. Model Deployment:

API integration is used to invoke the model through cloud services to ensure responsiveness and scalability.

4.4.2 Generate Virtual Avatar

To generate Virtual Avatar, as shown in Figure 3, a set of short video clips covering different emotions (e.g., smiling, nodding, thinking, etc.) are first pre-generated on the basis of digital images using AI generation techniques. Then based on the chatbot responses, the voice is played in conjunction with a TTS (Text-to-Speech) speech generation system. And at the same time, the system is allowed to play an animation that matches the voice. This way to enable low-cost and rapid development, as well as the effect of a virtual avatar.

4.5. Initial testing and feedback

The initial testing phase aims to validate the usability of the chatbot system as well as to explore the effectiveness of an immersive interactive experience that incorporates avatars. Early user feedback was also collected to identify strengths and weaknesses and areas for improvement. A small-scale test was conducted with six participants, simulating a scenario from a real-life medical consultation, where participants interacted with the chatbot's two interactive interfaces

4.5.1 Participants

Six participants in this study were selected by convenience sampling involving healthy young people in the age range of 20 to 30 years. They all had experience of attending a hospital. Participants were informed about the scope of the study, safety and freedom of participation prior to the user test.

4.5.2 Setup

The experiment was conducted in a quiet room equipped with a desktop computer, a headset with microphone and a monitor displaying the virtual patient interface. The virtual patient system includes a Large Language Model (LLM) to facilitate conversational experience and an avatar to provide a more immersive experience. The microphone can be used for interaction when selecting voice input, and the keyboard is used when selecting the text input group. The home page of the system provides access to both types of interaction.

5.4.3 Procedure

1) Notify participants and obtain their consent.

2) Introduce them to the interface and provide them with a scenario: ‘You are a patient with persistent headache symptoms. Please describe your symptoms and answer the questions posed by chatbot as accurately as possible. ’

3) Participants interacted with the chatbot using each of the two interaction modes. They were free to choose the order in which they interacted with the chatbot. Participants interacted with the 'virtual doctor' for 5 minutes each.

3) During the user testing phase, the researcher made observations to collect contextual data.

4) After user testing the tool, the researcher conducted semi-structured interviews with participants.

5) The user testing was concluded with a final comment section, giving participants the opportunity to give any additional feedback.

4.5.3 Data collection and analysis

1) Ethnographic and Behavioral Observation

In order to better understand the context and behaviour of the participants and the tool, the researchers used ethnographic ( ) research techniques to record interactions and analyse user habits and difficulties. In this study, behaviours such as the order of participants' choices of the two interaction interfaces, the average input length of participants, whether there was any pause or confused behaviour during the interaction process, and the response time comparison between voice input and text input can be observed to analyse user habits and difficulties, and to improve the understanding of the design solution.

2) Semi-Structured Interviews and Thematic Analysis

In order to understand the subjective experience of the participants, the researchers designed a semi-structured interview. The interview questions were centred around a few specific themes, and the researcher could continue to ask questions if she felt there was room for more clarification or further reasoning. The themes set for the interview were:

a. User experience with chatbot

b. Dialogue Strategies

c. Users' perceptions of the virtual avatar

d. Impact of the two interaction modes and the interaction interface on willingness to express

3) Other feedback

Allows users to freely express their views on anything related to the project, allowing users to express their opinions on their own without guidance from the researcher. This approach can provide more realistic and intuitive suggestions for improvement, and uncover problems that the researcher may not have anticipated.

5.4.5 Findings and insights.

1) Observations and Insights

a. Behavioural modes:

After the introduction of the two modes, users were all first curious about the interface of audio-visual interaction and used it first.

In terms of the interactive interface of both modes of interaction interactively, users tended to have shorter and less expressive inputs during textual interactions, whereas users were more willing to be more expressive during audio interactions; the difference may be caused by the use of textual and audio inputs ( ), but it is difficult to determine whether there is a significant difference in the information revealed in the two cases.

Ambiguous or irrelevant input reveals a flaw in the chatbot's ability to maintain dialogue coherence.

b. Degree of Participation:

Participants interacted with the immersive virtual doctor interface for a longer period of time than with the text-based interface, highlighting the appeal of visual and auditory interactions, but also the potential for less effective communication as a result.

2) Outcome of Thematic Analysis

a. User Experience with the Chatbot

Participants generally found the chatbot easy to use, with a clear and structured conversation flow. The interactive elements, such as the avatar and voice options, were appreciated, but some users noted occasional difficulties in understanding lengthy responses. Additionally, some participants expressed frustration when the chatbot failed to recognize minor variations in symptom descriptions, leading to repeated questions.

b. Dialogue Strategies

The chatbot's structured questioning approach was perceived as effective for gathering medical history. However, some participants noted that open-ended questions felt more engaging, while close-ended ones made them feel restricted in expressing their symptoms. Additionally, when users provided ambiguous or unexpected responses, the chatbot sometimes failed to generate appropriate follow-up questions, affecting conversation continuity.

c. Users' Perceptions of the virtual avatar

The virtual doctor avatar was well-received for adding professionalism and engagement. However, some participants felt the avatar lacked emotional variability, making interactions feel somewhat robotic. Additionally, there was a request for customizable avatars, such as different gender options or non-human representations, to increase user comfort.

d. Impact of Interaction Modes on Willingness to Express

Audio-visual interaction encouraged longer and more expressive responses compared to text input, making users feel more engaged. However, some participants preferred text input in public settings due to privacy concerns. The voice mode was also noted to be more immersive but occasionally led to slower interactions due to processing delays.

3) other feedback

Ease of use: Most participants found the system easy to use. However, at the same time, some participants also indicated that some prompts could be added to help with the participants' expressions or to guide the participants in their responses.

User experience: Participants valued interactivity and immersion but indicated that the virtual doctor's tone and facial expressions needed more emotional variation.

Customization Preferences: Some participants mentioned wanting to personalize the image of the virtual avatar; some participants also preferred non-anthropomorphic characters.

4) Key findings

a. Pros:

The chatbot demonstrated strong language comprehension and was able to process most user input effectively, especially in structured scenarios. Generation of dynamic questions based on previous responses was logical and intuitive, facilitating comprehensive symptom exploration. The virtual doctor received positive responses and participants noted an increased sense of presence and professionalism.

b. Limitations:

The chatbot occasionally encountered vague or complex descriptions that required manual clarification in some cases. Physical interaction in voice input was confusing and required prior learning. Participants could easily notice that the voice did not match the character's mouth when the virtual doctor animation was played, making the flow of dialogue less coherent and natural.

5. Discussion

5.1 Preliminary Results.

The immersive interactive interface consisting of a speaking image built by the AI model and voice interaction did have the effect of encouraging users to respond more positively. Users tended to respond with more vivid and specific descriptions than traditional text-based interactive interfaces, although not necessarily helpful for medical history taking. In particular, when faced with open-ended questions, users tended to respond more positively in Scenario 2, which provided a more immersive experience, while responses were more terse and robotic in the text dialogue interface. These results provide clear directions for further optimisation, including enhancing the system's ability to handle ambiguous inputs, improving the virtual human's multimodal interaction techniques, and exploring the impact of different interaction modes on the efficiency of synchronous communication and information disclosure.

5.2 Significance of Research Findings

The design of a history collection Chatbot based on LLM and virtual avatar technologies opens up new possibilities for the design of interaction interfaces for future history collection tools (Fadhil & Schiavo, 2019). Through the combination of virtual doctor image and natural language processing, this design not only enhances the immersive and anthropomorphic experience of the interaction, but also guides the patient to express his/her condition in a more flexible and intelligent way, breaking the limitations of traditional questionnaire or text-based interfaces (Kurniawan et al., 2015). This also provides guidance for further design optimization (Stock et al., 2023).

5.3 Limitations

a. Single avatar image

The single avatar image and lack of personalization limit the user's trust and identification with the system. The current virtual doctor avatar fails to adapt to the user's gender, age, cultural background, or language preference, and exhibits a uniform and lack of diversity in design, which may make it difficult for some users to empathize or connect with the system (Choi et al., 2020). In addition, the virtual doctor's facial expressions and tone of voice lack emotional diversity, making it difficult to adapt to different emotional scenarios (e.g., when a user expresses anxiety or pain) (Robertson et al., 2015). This limitation not only reduces users' willingness to interact, but may also undermine the comprehensiveness and accuracy of medical history collection, affecting the overall user experience and communication effectiveness (Fink et al., 2024).

b. Insufficient emotional capture of user feedback

Despite positive feedback on the virtual doctor's visual and voice interaction design, participants noted that the virtual doctor's intonation and facial expressions lacked sufficient emotional variation to demonstrate empathy (Berman & Chutka, 2016). This deficiency may limit patients' trust and reliance on the system, especially in emotional scenarios (e.g., when patients express anxiety or distress) (Stock et al., 2023).

c. The language model is not sufficiently targeted

The technical implementation also has limitations that may have an impact on the design effectiveness in further studies. The use of ChatGPT-4.0 as the core language model in the development of the history collection Chatbot Demo has some limitations. Firstly, ChatGPT-4.0 is a general-purpose model that, although it has powerful natural language processing capabilities, is not designed specifically for the medical domain and may have a biased understanding of specialized medical terminology, structured collection of medical histories, and description of complex medical conditions (Grice et al., 2017).

d. Data Diversity and Limitations of Test Scope

The test sample included only six participants, which is a small sample size and lacks sufficient diversity to reflect usage by a wide range of people (e.g., different ages, cultural backgrounds, or medical needs) (Athota et al., 2020). In addition, the test focused primarily on descriptions of common conditions and did not adequately assess the system's performance in complex medical scenarios, such as multimorbidity or rare diseases (Peterson et al., 1992).

6.Future Work

6.1 Design Optimization

6.1.1 Providing More Personalized Solutions

To meet the needs of different patient groups, personalization features will be introduced in the future. For example, by analyzing the user's health history, language preference, and cultural background, the system is able to dynamically adjust the image of the virtual doctor (e.g., female patients may tend to confide in a female doctor (Choi et al., 2020)), the language style of the questions, and the complexity of the content. In addition, it is planned to integrate health data (e.g., biodata provided by wearable devices) to provide patients with more targeted medical history collection programs, further enhancing user trust and engagement (Stock et al., 2023).

6.1.2 Expanding Interaction Modes and Further Developing Multimodal Interaction

The current system mainly supports text and voice interaction, and will explore the combination of more modalities in the future, such as gesture recognition, expression analysis, and haptic feedback. Multimodal interaction will help the system understand the patient's intent more comprehensively and provide a more natural way of communicating with the patient. For example, by capturing the patient's facial expression through the camera, the system can adjust the virtual doctor's tone of voice and the difficulty of the questions in real time, enhancing the immersion and emotional relevance of the interaction (Zhen et al., 2023; Robertson et al., 2015).

6.2 Use of More Targeted Macrolanguage Modeling

While ChatGPT already provides robust language generation capabilities for medical history collection, the use of large language models optimized specifically for the medical domain (such as MedPaLM or other medical-specific models) will be explored in the future. These models have been specifically trained on medical corpora and perform better when dealing with specialized terminology, complex disease descriptions, and multiple rounds of medical conversations (Grice et al., 2017; Athota et al., 2020). At the same time, it will be investigated how to incorporate domain knowledge updating mechanisms into these models to ensure that the system is always up-to-date with the latest medical practices (Peterson et al., 1992).

6.3 Improved User Testing

The current user tests are limited in size, and the sample size will be expanded in the future to cover a wider range of patient groups in terms of age, language, and health status. Testing will include user response efficiency, interaction satisfaction, and the comprehensiveness and accuracy of the system's collection of medical history (Hausberg et al., 2012). Through a combination of quantitative analysis and qualitative interviews, we will gain an in-depth understanding of the system's strengths and weaknesses, providing a reliable basis for the next step of optimization (Fadhil & Schiavo, 2019).

6.4 Conduct Re-iterative

Based on the results of user testing and technology development trends, the system will enter a continuous iteration phase. Each round of iteration will focus on solving the problems found and gradually introducing new functional modules. For example, privacy protection measures will be improved, the security of data storage and transmission will be optimized, and the personalized performance of the virtual doctor will be further enhanced to better fit the needs of different patients (Athota et al., 2020). The core goal of further iterations is to make the system gradually become an efficient and credible medical history collection tool in the medical field through continuous improvement (Stock et al., 2023).

7. Conclusion

In this project, an innovative interactive system for medical history collection is designed by introducing the Large Language Model (LLM) and virtual avatar technology, which combines an immersive virtual doctor image with a traditional text-based dialogue interface, exploring a new model of patient-centered active communication (Berman & Chutka, 2016; Robertson et al., 2015). The design has significant potential to enhance the efficiency of doctor-patient communication and patient experience, both by enhancing user immersion and interactivity and by providing smarter and more flexible technical support for medical history collection (Choi et al., 2020). Future research will further expand the scale of user test samples, deeply analyze the behavioral characteristics and feedback of different user groups, and at the same time optimize the technical implementation to ensure the ethical compliance and practical application value of the system, so as to lay the foundation for constructing a more efficient and humanized medical interaction system (Fink et al., 2024; Stock et al., 2023).

Reference

Berman, A. C., & Chutka, D. S. (2016). Assessing effective physician-patient communication skills: "Are you listening to me, doc?". *Korean journal of medical education*, *28*(2), 243–249. https://doi.org/10.3946/kjme.2016.21

Banathy, B. H. (1996). Conversation in social systems design. *Educational Technology*, *36*(1), 39-41.

The Double Diamond. (2023, May 11). Design Council. https://web.archive.org/web/20230511152044/https://www.designcouncil.org.uk/our-resources/the-double-diamond/

Hampton JR, Harrison MJ, Mitchell JR, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. Br Med J. 1975;2(5969):486-489.

Dorr Goold S, Lipkin M. The doctor-patient relationship: challenges, opportunities, and strategies. J Gen Intern Med. 1999;14(Suppl 1):S26-S33

Boyd, E., & Heritage, J. (2006). Taking the history: questioning during comprehensive history-taking. In J. Heritage & D. W. Maynard (Eds.), Communication in Medical Care: Interaction between Primary Care Physicians and Patients (pp. 151–184). chapter, Cambridge: Cambridge University Press.

Burt J, Campbell J, Abel G, et al. Improving patient experience in primary care: a multimethod programme of research on the measurement and improvement of patient experience. Southampton (UK): NIHR Journals Library; 2017 Apr. (Programme Grants for Applied Research, No. 5.9.) Appendix 1, Global Consultation Rating Scale. Available from: https://www.ncbi.nlm.nih.gov/books/NBK436536/

Grice, G. R., Gattas, N. M., Prosser, T., Voorhees, M., Kebodeaux, C., Tiemeier, A., Berry, T. M., Wilson, A. G., Mann, J., & Juang, P. (2017). Design and Validation of Patient-Centered Communication Tools (PaCT) to Measure Students' Communication Skills. American journal of pharmaceutical education, 81(8), 5927. https://doi.org/10.5688/ajpe5927

Zhen, R., Song, W., He, Q., Cao, J., Shi, L., & Luo, J. (2023). Human-Computer Interaction System: A Survey of Talking-Head Generation. Electronics, 12(1), 218. https://doi.org/10.3390/electronics12010218

Kurniawan, A. A., Fachri, W. E., Elevanita, A., & Agushinta, R. D. (2015, October). Design of chatbot with 3D avatar, voice interface, and facial expression. In 2015 international conference on science in information technology (ICSITech) (pp. 326-330). IEEE.

Fink, M. C., Robinson, S. A., & Ertl, B. (2024, July). AI-based avatars are changing the way we learn and teach: benefits and challenges. In Frontiers in Education (Vol. 9, p. 1416307). Frontiers Media SA.

Stock, A., Schlögl, S., & Groth, A. (2023, July). Tell me, what are you most afraid of? Exploring the Effects of Agent Representation on Information Disclosure in Human-Chatbot Interaction. In International Conference on Human-Computer Interaction (pp. 179-191). Cham: Springer Nature Switzerland.

Zagade, A., Killedar, V., Mane, O., Nitalikar, G., & Bhosale, S. AI-Based Medical Chatbot for Disease Prediction.

Fadhil, A., & Schiavo, G. (2019). Designing for health chatbots. arXiv preprint arXiv:1902.09022.

Choi, Y., Mehraliyev, F., & Kim, S. (2020). Role of virtual avatars in digitalized hotel service. International Journal of Contemporary Hospitality Management, 32(3), 977-997.

Robertson, S., Solomon, R., Riedl, M., Gillespie, T. W., Chociemski, T., Master, V., & Mohan, A. (2015). The visual design and implementation of an embodied conversational agent in a shared decision-making context (eCoach). In Learning and Collaboration Technologies: Second International Conference, LCT 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2–7, 2015, Proceedings 1 (pp. 427-437). Springer International Publishing.

Athota, L., Shukla, V. K., Pandey, N., & Rana, A. (2020, June). Chatbot for healthcare system using artificial intelligence. In 2020 8th International conference on reliability, infocom technologies and optimization (trends and future directions)(ICRITO) (pp. 619-622). IEEE.

Srivastava, P., & Singh, N. (2020, February). Automatized medical chatbot (medibot). In 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC) (pp. 351-354). IEEE.

Hampton, J. R., Harrison, M. J., Mitchell, J. R., Prichard, J. S., & Seymour, C. (1975). Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. Br Med J, 2(5969), 486-489.

Peterson, M. C., Holbrook, J. H., Von Hales, D. E., Smith, N. L., & Staker, L. V. (1992). Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. Western Journal of Medicine, 156(2), 163.

Goold, S. D., & Lipkin Jr, M. (1999). The doctor–patient relationship: challenges, opportunities, and strategies. Journal of general internal medicine, 14(Suppl 1), S26.

Hausberg, M. C., Hergert, A., Kröger, C., Bullinger, M., Rose, M., & Andreas, S. (2012). Enhancing medical students' communication skills: development and evaluation of an undergraduate training program. BMC medical education, 12, 1-9.

Deveugele, M., Derese, A., De Maesschalck, S., Willems, S., Van Driel, M., & De Maeseneer, J. (2005). Teaching communication skills to medical students, a challenge in the curriculum?. Patient education and counseling, 58(3), 265-270.

Roter, D. L., Frankel, R. M., Hall, J. A., & Sluyter, D. (2006). The expression of emotion through nonverbal behavior in medical visits: mechanisms and outcomes. Journal of general internal medicine, 21, 28-34.

Pękacz, A., Kądalska, E., Skoczylas, A., & Targowski, T. (2019). Patient satisfaction as an element of healthcare quality–a single-center Polish survey. Reumatologia/Rheumatology, 57(3), 135-144.

Chang, C. T., & Chen, C. X. (2015). Babyface effects: A double-edged sword in healthcare service evaluations. Asian Journal of Social Psychology, 18(4), 325-336.

World Health Organization. (2025). Medical doctors by sex. WHO. https://www.who.int/data/gho/data/indicators/indicator-details/GHO/medical-doctors-by-sex