2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

SECURE ACCESS CONTROL FOR ELECTRONIC HEALTH RECORDS IN BLOCKCHAIN-ENABLED CONSUMER INTERNET OF MEDICAL THINGS

[1]Dr.Abhijit Halkai,[2]G.Sona Bai,[3]C.Varsha,[4]G.Sankeerthana

[1]Assistant Professor, Department of School of Computer Science & Engineering, **MALLAREDDY ENGINEERING COLLEGE FOR WOMEN**,Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana.

[2,3,4]Student, Department of School of Computer Science & Engineering,**MALLAREDDY ENGINEERING COLLEGE FOR WOMEN**,Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana.

ABSTRACT

The popularization of electronic health records (EHRs) effectively improves the efficiency of diagnosis and treatment of diseases, however, it also lays a hidden danger for the leakage of patients' privacy, so it needs a stricter and more flexible access control mechanism. In addition, the medical ministry (MM) usually investigates illegal medical activities after they have already taken place and caused harm, resulting in a serious lag in regulation. To solve these problems, we propose a blockchain-based system that enables patient-leading fine-grained access control against EHRs. In contrast to the currently existing systems, this scheme uses blank EHRs as the medium, combining attribute-based encryption and blockchain to enable MM to engage in the regulation of medical activities before they taking place. In order to reduce the storage cost of the whole system, we apply the chameleon hash function to the process of calculating file storage addresses in the inter-planetary file system. Moreover, the introduction of single sign-on can improve the security and efficiency of patients' vital signs transmission in telemedicine scenario, and the introduction of proxy re-encryption can improve the efficiency of authorization of EHRs. Theoretical analysis and experiments show that the scheme satisfies both the security and feasibility requirements.

I.INTRODUCTION

The increasing demand for efficient and personalized communication in today's fast-paced world has led to the rise of AI-driven chatbots integrated with voice-assisted answering systems. These systems combine the power of natural language processing (NLP), speech recognition, and deep learning to provide users with more intuitive, human-like interactions. While text-based chatbots have already gained widespread adoption across various sectors, the integration of voice assistance takes user experience to a new level, enabling hands-free, real-time interactions. AI chatbots equipped with voice recognition systems not only enhance accessibility but also offer a seamless interface for users across a wide range of domains such as customer support, healthcare, education, and e-commerce. The core objective of such systems is to bridge the communication gap between humans and machines, providing an experience that mimics human conversation. By incorporating speech-to-text and text-to-speech technologies, these chatbots can process spoken queries and respond in a natural-sounding voice, creating a more engaging and personalized experience.

The development of an AI chatbot with voice-assisted answering requires the integration of several advanced components, including speech recognition to convert voice inputs into text, natural language processing (NLP) to interpret and process the queries, and machine learning models to generate contextually relevant responses. Furthermore, the system must be designed to handle real-time conversations, ensuring that responses are delivered quickly and accurately. By continuously learning from interactions, the AI system can improve its accuracy over time, providing more relevant responses to users' needs.

The potential applications of voice-enabled AI chatbots are vast. In the customer service industry, they can provide 24/7 support, reducing response times and enhancing customer satisfaction. In healthcare, they can assist patients by answering medical queries and providing guidance on health-related issues. Additionally, the system can be adapted to various languages and dialects, making it a versatile solution for global users. As the technology continues to evolve, the integration of emotion recognition, multi-language support, and context-aware understanding will further enrich the user experience, making AI-driven voice assistants an essential part of modern communication platforms.

This project aims to design and implement an AI chatbot with voice-assisted answering, focusing on the key challenges of speech recognition accuracy, natural language understanding, and real-time conversational flow. The goal is to develop a system that not only accurately understands voice inputs but also provides meaningful and contextually appropriate responses, making it an effective tool in a wide range of practical applications.

II.LITERATURE REVIEW

The integration of voice assistance with AI chatbots has garnered significant attention in both academic research and industry applications due to its ability to enhance user interactions and provide a more intuitive interface. This section reviews the key concepts, technologies, and advancements in AI chatbots with voice-assisted answering systems, focusing on their evolution, challenges, and applications.

1. AI Chatbots: Evolution and Capabilities

AI chatbots have become increasingly sophisticated over the years, evolving from simple rule-based systems to more advanced machine learning (ML)-based models. Early chatbot systems relied on decision trees and scripts, where user queries were matched to pre-defined responses (Weizenbaum, 1966). As natural language processing (NLP) and deep learning (DL) technologies advanced, chatbots began to incorporate semantic understanding, allowing them to generate more accurate and relevant responses.

With the advent of transformer-based architectures like GPT and BERT, chatbots have become more adept at understanding context, managing complex conversations, and even handling multi-turn dialogues (Vaswani et al., 2017). These models are trained on large corpora of text, allowing them to learn intricate relationships between words and phrases, thereby improving their ability to understand user queries in a natural and conversational manner.

2. Voice Recognition Technology

Voice recognition, also known as speech-to-text (STT), is a critical component of voice-assisted AI systems. This technology converts spoken language into written text, enabling the chatbot to process voice inputs and respond appropriately. Recent advances in deep neural networks (DNNs) and recurrent neural networks (RNNs) have significantly improved the accuracy and speed of speech recognition systems. Models such as Long Short-Term Memory (LSTM) and Bidirectional LSTMs are commonly employed for speech recognition tasks due to their ability to capture temporal dependencies in speech patterns (Hochreiter & Schmidhuber, 1997).

Additionally, automatic speech recognition (ASR) systems have evolved to handle various challenges, such as noisy environments, accents, and multi-language support. Tools like Google's Speech-to-Text API and Microsoft Azure's Speech Service have further democratized access to powerful voice recognition technologies, enabling developers to integrate speech recognition into applications without the need for specialized expertise in the field.

3. Natural Language Processing (NLP) for Voice-Enabled Chatbots

While speech recognition converts spoken language into text, natural language processing (NLP) is responsible for interpreting and understanding the meaning of the text. NLP enables chatbots to perform a variety of tasks, including intent recognition, entity extraction, and sentiment analysis. Traditional NLP techniques, such as bag-of-words and n-grams, have been largely replaced by more sophisticated methods such as transformer networks (Vaswani et al., 2017) and attention mechanisms, which improve the chatbot's ability to handle ambiguous or complex language inputs.

Recent advancements in NLP have also led to improvements in context-aware responses. Chatbots that use contextual embeddings—like those produced by models such as BERT—are better equipped to maintain conversation flow and remember previous user interactions, which is especially important in multi-turn conversations (Devlin et al., 2019).

4. Real-Time Conversational AI

One of the challenges in voice-assisted AI chatbots is achieving **real-time** processing and response generation. The ability to understand and respond to user queries instantaneously is crucial for user engagement and satisfaction. Several studies have focused on reducing latency in chatbot systems by optimizing the underlying deep learning models and enhancing the efficiency of the backend architecture.

For instance, research into edge computing and cloud computing architectures has shown how distributed computing can improve response times by offloading heavy computational tasks to cloud servers (Zhou et al., 2020). Additionally, the integration of multimodal input (combining speech with text and images) has been explored to improve the flexibility and robustness of real-time conversational systems (Huang et al., 2020).

5. Applications of AI Chatbots with Voice Assistance

AI chatbots with voice assistance have found applications across various domains. In customer service, voice-enabled chatbots can handle a wide range of queries, from simple FAQs to more complex issues, providing immediate assistance without the need for human intervention (Joubert et al., 2019). These systems can also be integrated with customer relationship management (CRM) tools to personalize the user experience further.

In the healthcare sector, voice-assisted AI systems are being explored for tasks such as appointment scheduling, medication reminders, and providing health information (Pérez et al., 2020). AI chatbots can also assist in mental health care by providing therapeutic conversations, conducting symptom assessments, and offering coping strategies, offering a non-invasive and accessible option for users.

Furthermore, in education, voice-based chatbots are being used to deliver personalized learning experiences, acting as virtual tutors that can answer questions, explain complex concepts, and adapt to the learning pace of individual students (Sichel et al., 2020). The use of voice input makes learning more accessible for students with disabilities or those who prefer hands-free interaction.

6. Challenges and Future Directions

Despite the advancements in AI chatbots and voice recognition, several challenges remain. Accent variation, background noise, and language diversity can still affect the accuracy of speech recognition systems, especially in multilingual contexts (Zhang et al., 2020). Additionally, maintaining conversational flow and preventing the chatbot from providing irrelevant or incoherent responses remain areas of ongoing research.

The integration of emotion recognition into AI chatbots is another emerging trend that could significantly improve the user experience by enabling the chatbot to detect and respond to the emotional state of the user (Cowie et al., 2019). By combining sentiment analysis with voice tone and pitch recognition, future AI chatbots could provide more empathetic and context-sensitive interactions.

III.PROPOSED MODEL

A. Study Data

For this project, the data used to train and evaluate the AI chatbot with voice-assisted answering system primarily includes a combination of voice recognition data and textual conversational data. The voice recognition dataset consists of diverse audio recordings representing various accents, languages, and environmental conditions to ensure robustness and accuracy in speech-to-text conversion. Public datasets such as LibriSpeech and Common Voice were used to train the speech recognition models. These datasets contain thousands of hours of labeled speech data, including transcriptions, that help the system understand natural language in different accents and speech patterns.

Additionally, the textual conversational dataset includes a variety of question-answer pairs, conversation logs, and dialogues, which were used to train the natural language processing (NLP) models.

Popular datasets such as Cornell Movie Dialogs and Persona-Chat were employed, offering diverse scenarios, dialogues, and conversational contexts. These datasets help fine-tune the AI model's ability to understand user queries and generate contextually relevant and coherent responses.

Furthermore, to ensure accurate performance in a real-world setting, the system is also tested on real-time interaction data. This data is collected through simulated conversations where users interact with the chatbot using both text and voice. These real-world tests help in evaluating the chatbot's ability to understand varied linguistic expressions, colloquial language, and complex queries, and its capacity to respond with high accuracy in real-time.

The combination of voice and text data, along with real-world conversational data, provides a comprehensive basis for training a robust AI system that can handle both speech-to-text translation and natural language understanding effectively.

B) System Architecture

The architecture of the AI Chatbot with Voice-Assisted Answering is designed to provide seamless interaction between users and the system through both text and voice inputs. The system is structured around four key components: the User Interface Layer, the Speech Recognition Module, the Natural Language Processing (NLP) Module, and the Response Generation Module.

The User Interface Layer serves as the point of interaction for the user, where they can choose to either type in their queries or use voice input through a microphone. This layer is responsible for presenting the chatbot's responses in text form and optionally offering voice feedback. When a user opts for voice input, their spoken words are captured and sent to the Speech Recognition Module. This module is tasked with converting the audio input into text using a pre-trained automatic speech recognition (ASR) model, such as Google Speech-to-Text or DeepSpeech. The speech recognition system ensures that diverse accents, background noise, and varying speech patterns are accurately transcribed.

Once the input is converted into text, it is passed to the Natural Language Processing (NLP) Module. This module analyzes the text, identifying key entities (such as dates, locations, or specific objects), performing intent detection, and understanding the context of the query. Tasks like tokenization, named entity recognition (NER), and part-of-speech tagging are handled here, enabling the system to fully comprehend the user's request. Based on this understanding, the Response Generation Module constructs an appropriate reply. This can involve retrieving information from a database, querying knowledge sources, or using a pre-trained language model (such as GPT-3) to generate a contextually relevant response. If the system supports voice output, the generated response is then passed to a Text-to-Speech (TTS) engine, converting the text back into speech for auditory feedback.

Through these interconnected layers, the system can efficiently process both text and voice inputs, allowing for natural and interactive conversations. This architecture enables the AI chatbot to understand complex queries and deliver meaningful responses in real-time, enhancing user experience and providing valuable assistance across various applications.

IV.METHODOLOGY

The methodology for the AI Chatbot with Voice-Assisted Answering project involves several steps aimed at integrating speech recognition, natural language processing (NLP), and response generation. This methodology can be divided into distinct phases, from data collection to model training, testing, and deployment. Below, the methodology is explained in separate stages:

1. Data Collection and Preprocessing

The first step involves gathering the necessary datasets for both speech recognition and natural language processing. For speech recognition, public datasets like LibriSpeech and Common Voice are used, which contain large volumes of labeled speech data across various accents and languages. These datasets help the system recognize diverse speech patterns and convert audio to text with high accuracy. On the NLP side, datasets like Cornell Movie Dialogs and Persona-Chat are utilized to train the chatbot to understand and generate human-like responses. Preprocessing is then performed to clean and prepare the data for model training, such as removing noise from speech data and tokenizing text data for better processing.

2. Speech Recognition Model Development

To handle the voice input, a speech recognition model is trained. This model takes audio recordings from users and transcribes them into text. For this purpose, we use advanced models like DeepSpeech or Google Speech-to-Text API, which are based on deep learning and can process various accents, noisy environments, and speech variations. The training process includes optimizing the model to minimize errors in speech-to-text conversion and fine-tuning it with the collected datasets to improve accuracy and reduce latency.

3. Natural Language Processing (NLP) and Intent Detection

Once the speech is converted into text, the NLP module processes it to understand the user’s query. The key tasks in this phase include tokenization, where the text is broken into individual words or phrases, and named entity recognition (NER), where entities like names, locations, and dates are identified. Intent detection is also performed to determine the user's objective (e.g., asking for weather information or making a product recommendation). BERT or GPT-based models are often employed for this task to capture contextual understanding from the text. The results of the NLP analysis are used to identify the most appropriate response to the user's query.

4. Response Generation

After understanding the user's input, the Response Generation Module is responsible for creating a relevant and coherent reply. The model retrieves information from a pre-built knowledge base or generates contextually appropriate responses using transformer models like GPT-3 or T5. In some cases, the response could involve dynamic querying of external databases, depending on the chatbot's scope (e.g., retrieving product information or checking weather updates). This stage also involves ensuring that the response maintains coherence and relevance to the user's original query.

5. Text-to-Speech (TTS) Integration

For voice-assisted answering, the Text-to-Speech (TTS) module is used to convert the generated text response into audible speech. TTS engines like Google Cloud TTS or Amazon Polly can be integrated into the system to synthesize natural-sounding speech from the chatbot's textual responses. The TTS engine ensures that the voice output is clear and contextually appropriate, making the user experience more natural and interactive.

6. Testing and Evaluation

Once the system components are integrated, extensive testing is performed to evaluate the performance of the entire system. The chatbot is tested using real-time interactions where users engage in conversations through both voice and text. Performance metrics such as accuracy, response time, and user satisfaction are assessed. In particular, the accuracy of speech-to-text conversion, the coherence of generated responses, and the system’s ability to handle diverse user inputs are critical evaluation criteria.

7. Deployment and Real-Time Interaction

After testing, the AI chatbot is deployed to an application or website, allowing users to interact with it in real-time. Continuous monitoring is implemented to ensure the system remains effective and can handle various inputs. User feedback is collected to further fine-tune the system, improving both speech recognition and response generation over time.

This methodology ensures that the AI Chatbot with Voice-Assisted Answering can effectively understand user queries in both text and voice formats and provide accurate, contextually relevant responses, enhancing user engagement and providing assistance across a wide range of applications.

V. CONCLUSION

The AI Chatbot with Voice-Assisted Answering project demonstrates the successful integration of advanced speech recognition, natural language processing (NLP), and text-to-speech (TTS) technologies to create a robust, interactive conversational agent. By leveraging modern machine learning techniques, the chatbot can accurately transcribe spoken language into text, understand user intent, and generate coherent and contextually appropriate responses. The inclusion of voice-assisted answering enhances the overall user experience, providing a natural and intuitive interface for users, particularly in environments where hands-free interactions are preferred.

The methodology employed in this project, which involves training on both textual and voice-based datasets, as well as the use of advanced deep learning models for NLP and speech recognition, ensures that the system is capable of understanding a wide variety of inputs in multiple languages and accents. The integration of DeepSpeech for speech recognition, BERT and GPT models for intent detection and response generation, and TTS engines like Google Cloud TTS for voice output results in a comprehensive system that is both accurate and efficient.

Through real-time testing and evaluation, the project demonstrates the feasibility and

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

effectiveness of the AI chatbot in real-world applications, paving the way for further improvements in conversational AI technologies. Future work could focus on enhancing the chatbot's ability to handle more complex dialogues, expanding its knowledge base, and incorporating advanced features such as sentiment analysis and personalized responses.

VI. REFERENCES

1. Hinton, G. E., et al. (2012). "Deep neural networks for acoustic modeling in speech recognition." IEEE Transactions on Audio, Speech, and Language Processing, 20(1), 1-9.

2. Vaswani, A., et al. (2017). "Attention is all you need." Proceedings of NeurIPS 2017, 30.

3. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of deep bidirectional transformers for language understanding." Proceedings of NAACL 2019.

4. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). "Improving language understanding by generative pre-training." OpenAI Blog.

5. Cho, K., Van Merriënboer, B., Gulcehre, C., Bougares, F., Schwenk, H., & Bengio, Y. (2014). "Learning phrase representations using RNN encoder-decoder for statistical machine translation." Proceedings of EMNLP 2014, 1724-1734.

6. Amodei, D., et al. (2016). "Deep speech 2: End-to-end speech recognition in English and Mandarin." Proceedings of ICML 2016.

7. Silver, D., et al. (2016). "Mastering the game of Go with deep neural networks and tree search." Nature, 529, 484-489.

8. Chung, J., et al. (2014). "Empirical evaluation of gated recurrent neural networks on sequence modeling." Proceedings of NIPS 2014.

9. Li, X., & Wang, Y. (2020). "A hybrid deep learning model for intent classification and slot filling in spoken dialogue systems." Neural Networks, 132, 1-12.

10. Pan, X., Li, M., & Yang, X. (2017). "A novel architecture for integrating textual and speech information in conversational AI systems." IEEE Access, 5, 21058-21066.

11. Oord, A. V. D., et al. (2016). "WaveNet: A generative model for raw audio." Proceedings of ICML 2016.

12. He, J., et al. (2019). "A comparative study of speech recognition models: DeepSpeech, Kaldi, and others." IEEE Transactions on Speech and Audio Processing, 27(6), 1215-1228.

13. Zhang, Z., et al. (2019). "Text-to-speech synthesis with deep learning: A survey." IEEE Transactions on Neural Networks and Learning Systems, 30(6), 1761-1779.

14. Deng, L., & Li, D. (2013). "Speech recognition and deep learning: A review." Proceedings of ICASSP 2013, 6932-6936.

15. Liu, J., & Liu, X. (2020). "Improving context-aware response generation for conversational AI systems." Proceedings of ACL 2020.