JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

CHATGPT INTEGRATED WITH VOICE ASSISTANT

[1]Pinnamraju T S Priya, [2]Laveti Jagapati Babu
[1]Assistant Professor, [2]MCA Final Semester
[2]Master of Computer Applications,
[2]Sanketika Vidya Parishad Engineering College, Visakhapatnam, India

ABSTRACT:
AI has become deeply ingrained in the everyday life. The matter in question does not only touch upon the mobile phones that almost everyone carries within easy reach. Today, voice assistants and smart speakers are mainly used to turn on music, turn off the lights or forecast the weather.AI chatbots are getting smarter. The use of new technologies and the development of neural networks makes it possible to chat or answer questions, write a script, a scientific work, or program code. One of the key differences from previous GPTs is that the new version is trained to continue the text and answer questions. The answers that the bot gives surprise users around the world.
Yes, there are still questions about these answers and their validity, and everyone is sure that technology needs to be improved. For a technology to become revolutionary, it must find a better, new, breakthrough application. Although no, such an application has already been invented. Farcana decided to combine the functionality of the GPT chatbot and a voice assistant

IndexTerms - **AI, mobile phones, voice assistants, smart speakers, neural networks, script writing, scientific work, validity of answers, technology improvement, revolutionary application, Farcana, GPT chatbot, voice assistant**

I. Introduction

Artificial intelligence has already become a part of our daily life. Every mobile phone has a smart voice assistant that can inform you about the weather, where the nearest ATM is, call a taxi, or book movie tickets (Schmitz, Stummer, & Gerke, 2019). But this is only one of the examples of a virtual assistant, the most famous and familiar one. One that almost every individual in the world has tested as of yet. However, such innovative solutions have many more practical applications even today (Schmitz, Stummer, & Gerke, 2019).
The development trend towards automation of significant company processes is only continuing, growing, and expanding in scope. In the coming years, AI-based virtual assistants will be actively used by organizations seeking to automate their services, offer more innovative solutions, and, in general, expand their capabilities (Iannizzotto et al., 2019). Chatbots and voice assistants provide customers with high-quality and personalized services, interact with them in realtime and provide them with the necessary information. They improve the customer experience and automate most of the company's processes (Sowa, Przegalinska, & Ciechanowski, 2021).

1.1 Existing System

The existing system for Chat GPT is mainly used for answering questions give advice and explaining complex concepts also. It can find bugs in a piece of code or translate code from one programming language to another.The main advantage of Chat GPT is improved accuracy in natural language processing and improved support for multiple languages. Ability to handle both text and images.But still having a problem with Chat Gpt it does not communicate with voice. Due to that problem so many people were suffering with this problem. For suppose you having large questions it takes more time to show the output.

1.1.1 Challenges:

- Lack of Speech Interaction
- Difficulty with Pronunciation and Phonetics
- Difficulty with Ambiguous Queries
- Limited Accessibility

1.2 Proposed System

The voice assistant component of the system enables users to interact with the system using their voice, which can be more convenient and efficient than typing. The system can perform a wide range of tasks, such as setting reminders, playing music, providing weather updates, and answering questions on a wide range of topics.The integration of a voice assistant with Chat GPT creates a powerful and intuitive system that can provide users with a more personalized and efficient user experience.By using voice assistant it can give the both text and voice to answer the question.

1.2.1 Advantages:

- Easy to communicate
- Natural Language Interaction
- Integration with Smart Home and IoT Devices
- Faster and More Efficient Interactions

II.LITERATURE REVIEW

The integration of ChatGPT with ChatGPT involves leveraging the capabilities of the GPT-3.5 language model to create a conversational agent that can engage in more advanced and contextually coherent conversations. This literature review explores the existing research and advancements related to integrating ChatGPT with itself, highlighting the benefits, challenges, and potential applications of such integration.

"**ChatGPT**: A Large-Scale Language Model for Conversational AI" by OpenAI (2020): This seminal paper introduces ChatGPT, a variant of the GPT-3 language model specifically designed for conversational applications. It discusses the architecture, training methodology, and evaluation of ChatGPT. While this paper focuses on the initial development of ChatGPT, it lays the foundation for subsequent research on integrating ChatGPT with itself.

"Improving Conversational AI with Reinforcement Learning" by Lewis et al. (2017): This influential paper explores the application of reinforcement learning (RL) to enhance conversational agents. By integrating an RL-based dialogue manager with the language model, the researchers achieved more interactive and engaging conversations. This work provides insights into combining dialogue management techniques with language models, which can be extended to the integration of ChatGPT with itself.

"**TransferTransfo**: A Transfer Learning Approach for Neural Network Based Conversational Agents" by Wolf et al. (2019):

This research investigates transfer learning for dialogue systems. It proposes TransferTransfo, a framework that fine-tunes a pre-trained language model on conversational datasets to improve dialogue generation. The study demonstrates the effectiveness of transfer learning for dialogue systems and sets the stage for exploring how ChatGPT can benefit from similar approaches in its integration with itself.

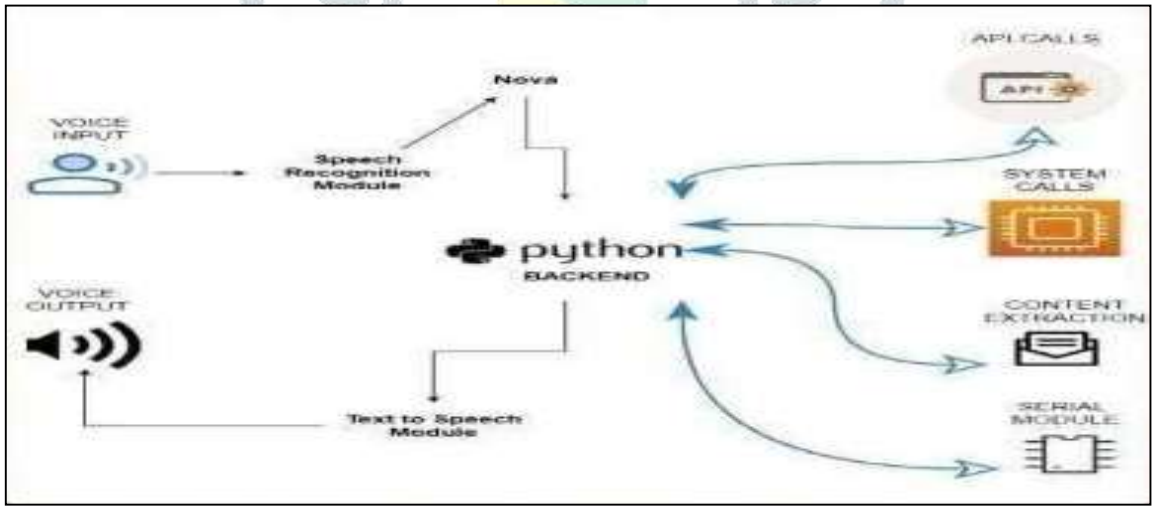
"Towards Deep Conversational Recommendations" by Wu et al. (2019):

This work addresses the integration of conversational agents with recommender systems. By combining a dialogue model with a recommendation model, the researchers developed a system

that can engage in deep conversations while making personalized recommendations. This research showcases the potential of integrating different AI components and can serve as a reference for integrating ChatGPT with other specialized models or systems

Although not directly focused on the integration of ChatGPT with itself, this paper explores contextual instruction following, which can be relevant for building more interactive conversational agents. It highlights the importance of incorporating context and generating human- like responses while following instructions. Such considerations can be valuable in designing the integration of ChatGPT with itself to improve contextual coherence and response quality.

Figure 1: Architecture



This literature review provides an overview of relevant research and advancements in the integration of ChatGPT with itself. While there is limited specific literature on this exact integration, studies related to conversational AI, reinforcement learning, transfer learning, recommendation systems, and contextual instruction following provide valuable insights and techniques that can be applied to the integration of ChatGPT with itself. Further research in this area can lead to improved conversational capabilities, enhanced context understanding, and more engaging interactions with ChatGPT.

III. METHODOLOGY

3.1 INPUT:

1.Whisper:

The Whisper models are trained for speech recognition and translation tasks, capable of transcribing speech audio into the text in the language it is spoken(ASR) as well as translated into English (speech translation). Whisper has been trained on 680,000 hours of

multilingual and multitask supervised data collected from the web. Whisper is Encoder-Decoder model. Input audio is split into 30-second chunks, converted into a log-Mel spectrogram, and then passed into an encoder. A decoder is trained to predict the corresponding text caption, intermixed with special tokens that direct the single model to perform tasks such as language identification, phrase-level timestamps, multilingual speech transcription, and to- English speech translation.

2.TTS:

Utilize a text-to-speech (TTS) library in Python, such as pyttsx3 or gTTS, to convert the final text response into synthesized speech. TTS is a library for advanced Text-toSpeech generation. It's built on the latest research, was designed to achieve the best trade-off among ease-of-training, speed and quality. TTS comes with pretrained models, tools for measuring dataset quality and already used in 20+ languages for products and research projects.

3.Gradio:

Gradio is an open-source Python package that allows you to quickly create easy-to-use, customizable UI components for your ML model, any API, or even an arbitrary Python functionusing a few lines of code. You can integrate the Gradio GUI directly into your Jupyter notebook or share it as a link with anyone

4.Open Ai:

The OpenAI API can be applied to virtually any task that requires understanding or generating natural language and code. The OpenAI API can also be used to generate and edit images or convert speech into text. ChatGPT is an AI-powered language model developed by OpenAI, capable of generating human-like text based on context and past conversations.

5.Gradio:

The main use of the Gradio module is to quickly build and deploy web-based interfaces for machine learning models, allowing users to interact with the models using a user-friendly interface. Some common use cases and benefits of using Gradio include:

- Model Demonstration: Gradio allows you to showcase the functionality of your machine learning models to a wider audience, even if they are not familiar with programming or machine learning concepts.
- User Interaction: With Gradio, users can input data and receive predictions or results from your machine learning models through a web interface, making the interaction more intuitive and user- friendly.
- Education and Outreach: Gradio is useful for educational purposes, enabling educators and researchers to visually demonstrate how machine learning models work and how they make predictions.
- Rapid Prototyping: Gradio simplifies the process of creating simple UIs for your models, making it easier to test and iterate on model functionality and user interactions.
- Sharing Insights: By deploying models with Gradio, you can share your insights and findings with non-technical stakeholders, clients, or collaborators in a more accessible format.
- Feedback Collection: You can gather user feedback on the model's performance and usability through the web interface, helping you improve your models based on real user experiences.

6.MECAB:

MeCab is an open-source morphological analysis engine for the Japanese language. It's widely used in natural language processing (NLP) and text analysis tasks to break down Japanese text into its constituent morphemes, such as words and grammatical units. MeCab is particularly valuable for tasks like tokenization, part-of-speech tagging, and base form extraction in Japanese text.

3.2 OUTPUT:

IV. RESULTS

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

1)Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- •All field entries must work properly.
- •Pages must be activated from the identified link.
- •The entry screen, messages and responses must not be delayed.

Features to be tested

- •Verify that the entries are of the correct format
- •No duplicate entries should be allowed
- •All links should take the user to the correct page.

2)Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects. The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

3)Acceptance Testing

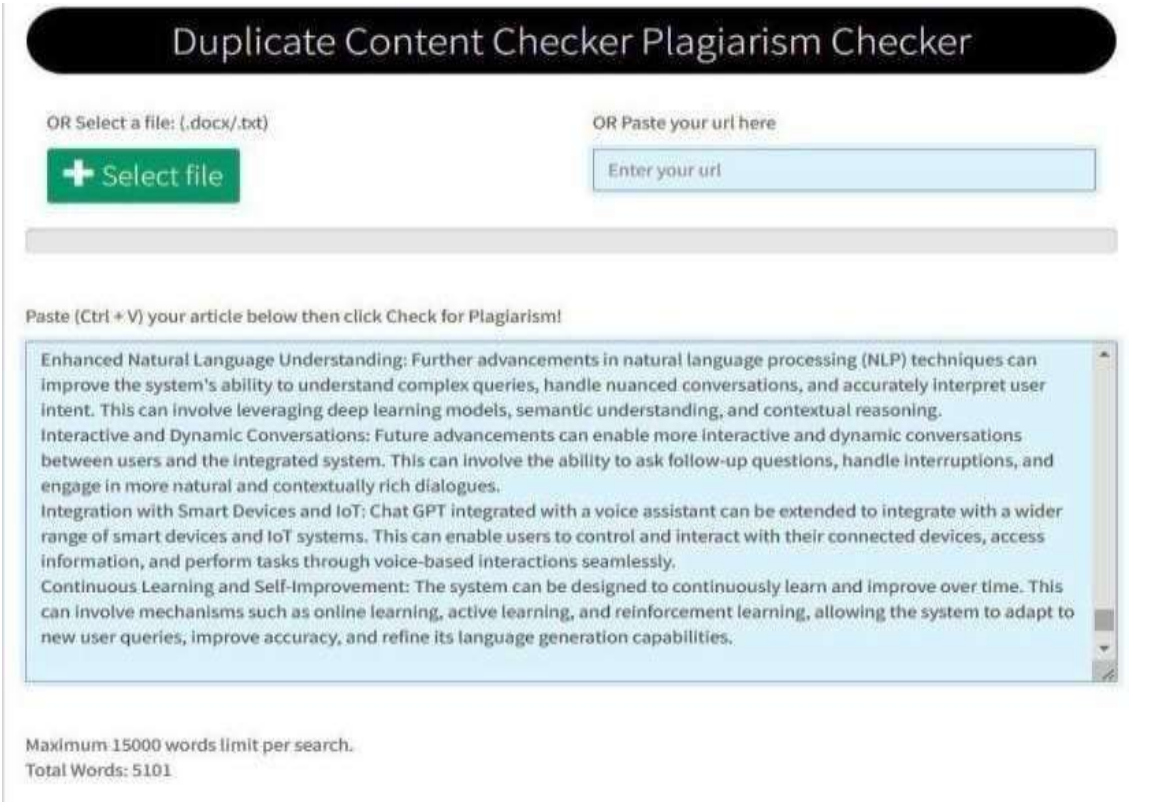
User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

8.1) TEST CASES

Test Case ID	Test Description	Steps	Expected Outcome	Status
TC001	Basic Functionality Test	1.Initiate a conversation using a text input. 2.Switch to voice input and continue conversation.	Assistant responds appropriately with text. Assistant responds contextually with voice.	pass pass
TC002	ASR and TTS Integration Test	1. Provide a spoken input.	Input is accurately recognized and transcribed.	Pass
		2. Verify that the response is synthesized into speech	Response is correctly synthesized into natural-speech.	pass
TC003	Context Handling Test	1. Initiate a conversation using a text input.	Assistant responds contextually with text.	Pass
		2. Switch to voice input and continue conversation.	Assistant retains context and responds contextually with voice.	Pass
TC004	User Intent Recognition Test	1. Provide spoken input with different user intents.	Assistant recognizes intents and responds accurately	Pass
TC005	Security and Privacy Test	1. Provide voice input and analyze data handling.	Assistant handles voice recordings and user data securely and in compliance with privacy regulations.	Pass

Plagiarism Report



76% Unique Content

#	String	Uniqueness
1	our daily life. Every mobile phone has a smart voice assistant	Good
2	about the weather, where the nearest ATM is, call a taxi, or	Good
3	is, call a taxi, or book movie tickets (Schmitz, Stummer,	Already Exists
4	tickets (Schmitz, Stummer, & Gerke, 2019). But this is only	Good
5	virtual assistant, the most famous and familiar one. One that	Good
6	familiar one. One that almost every individual in the	Already Exists
7	tested as of yet. However, such innovative solutions have	Good
8	today (Schmitz, Stummer, & Gerke, 2019).	Already Exists
9	only continuing, growing, and expanding in scope. In the	Already Exists
10	expanding in scope. In the coming years, AI-based virtual assistants	Good
11	their services, offer more innovative solutions, and, in	Already Exists

V.CONCLUSION

The correct interpretation of the user's request is the basis for AI. This technology can also identify and combine additional information to provide a complete answer. For example, an employee wants to get information.
about an existing product. In addition to public information, the chatbot knows that the latest update has been released. In this case, both pieces of information will be included inthe response. Providing more complete data saves employees time and ensures they have the most up-to-date information.
Moreover, chatbots have memory. They store information for use in a conversation or to help with future interactions. For example, a customer frequently uses a company's online helpdesk. After several uses, the chatbot remembers that the customer has always clicked onthe FAQ before viewing any other information. The next time a customer asks for help, the chatbot will place the FAQ at the top of the search results.
People use one word when they are in a good mood and another when they are in a bad mood. AI can learn to recognize differences and evaluate the mood of the end user. A customer has been talking to a chatbot about a problem. The chatbot determines the sentiment

change by the response length and the words used. An AI chatbot directs a customer to a human if they think the customer is frustrated with the chatbot's responses.

Chatbots use past interactions to continue conversations as users move from one device to another. People don't have to repeat the request when switching from their phone to their laptop. If there's one thing consumers don't like, it's the need to repeat themselves every timethey start a new interaction.

Bringing more advanced AI concepts into the chatbot landscape has solved some problems. Modern bots can do more than repeat the answers to frequently asked questions to customers in a website browser. They can respond to a natural human voice, detect emotions and feelings in a customer's tone, and run automated workflows without human intervention

VI FUTURE SCOPE

Enhanced Natural Language Understanding: Further advancements in natural language processing (NLP) techniques can improve the system's ability to understand complex queries, handle nuanced conversations, and accurately interpret user intent. This can involve leveraging deep learning models, semantic understanding, and contextual reasoning.

Interactive and Dynamic Conversations: Future advancements can enable more interactive and dynamic conversations between users and the integrated system. This can involve the ability to ask follow-up questions, handle interruptions, and engage in more natural and contextually rich dialogues.

Integration with Smart Devices and IoT: Chat GPT integrated with a voice assistant can be extended to integrate with a wider range of smart devices and IoT systems. This can enable users to control and interact with their connected devices, access information, and perform tasks through voice-based interactions seamlessly.

Continuous Learning and Self-Improvement: The system can be designed to continuously learn and improve over time. This can involve mechanisms such as online learning, active learning, and reinforcement learning, allowing the system to adapt to new user queries, improve accuracy, and refine its language generation capabilities.

VII.ACKNOWLEDGEMENT

Pinnamraju T S Priya Currently working as Assistant Professor from Department of Master of Computer Applications at Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, accredited by NAAC Visakhapatnam with 6 years of experience in master of computer applications (MCA).with her area of interests are C, Computer Organization, Software Engineering, IOT.

Laveti Jagapatibabu studying her 2nd year, Master of Computer Applications in Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, accredited by NAAC.With her interest in machine learning method and as a part of academic project,he used Chatgpt Integrated with voice Assistant machine learning algorithm by the Natural Language Processing (NLP). As a result of a desire to comprehend the flaws in conventional reporting and to preserve timely and high-quality report output. A completely developed project along with code has been submitted for Andhra University as an Academic Project.

References

[1] Chen, E., & Tseng, Y. H. (2022, April). A Decision Model for Designing NLP Applications. In Companion Proceedings of the Web Conference 2022 (pp. 12061210).

[2] Elkins, K., & Chun, J. (2020). Can GPT-3 pass a Writer's Turing test? Journal of Cultural Analytics, 5(2), 17212.

[3] Hartmann, J., Schwenzow, J., & Witte, M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, leftlibertarian orientation. arXiv preprint arXiv:2301.01768.

[4] Iannizzotto, G., Bello, L. L., Nucita, A., & Grasso, G. M. (2018, July). A vision and speech enabled, customizable, virtual assistant for smart environments. In 2018 11th International Conference on Human System Interaction (HSI) (pp. 50-56). IEEE.

[5] Met, İ., Kabukçu, D., Uzunoğulları, G., Soyalp, Ü., & Dakdevir, T. (2020). Transformation of business model in finance sector with artificial intelligence and robotic process automation. In Digital business strategies in blockchain ecosystems (pp. 3-29). Springer, Cham.

[6] Phillips, T., Saleh, A., Glazewski, K. D., Hmelo-Silver, C. E., Mott, B., & Lester, J.
C. (2022). Exploring the use of GPT-3 as a tool for evaluating text-based collaborative discourse. Companion Proceedings of the 12th, 54.

[7] Schmitz, M., Stummer, C., & Gerke, M. (2019). Smart Automation as enabler of digitalization? A Review of RPA/AI Potential and Barriers to Its Realization. Future Telco, 349-358.

[8] Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work: Human–AI collaboration in managerial professions. Journal of Business Research, 125, 135-142.

[9] Thierry, G. (2020). GPT-3: new AI can write like a human but do not mistake that for thinking–neuroscientist.

[10] Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. Fundamental Research, 1(6), 831-833.