*International Journal of Engineering Applied Sciences and Technology, 2024*
**Vol. 8, Issue 11, ISSN No. 2455-2143, Pages 202-206**
**Published Online March 2024 in IJEAST (http://www.ijeast.com)**

# INTEGRATION OF OPENAI API FOR MULTILINGUAL VOICE CHATBOT SYSTEMS

Mr. H. P. Khandagale
Assistant Professor
Computer Science and Technology Department,
Department of Technology,
Shivaji University,
Kolhapur, Maharashtra, India.

Mr. Anish Balaji Murshetwar, Mr. Yash Sunil Amane, Mr. Aryan Shirish Karanjkar, Mr. Shubham Ananda Bhilare
Student,
Computer Science and Technology Department,
Department of Technology,
Shivaji University, Kolhapur, Maharashtra, India.

*Abstract*— **The modern era of technology has a significant impact on society. With the creation of digital personal assistants, chatbots have become a popular entity in conversational services. Chatbots, also known as dialogue systems, interactive conversational agents, or virtual agents are an example of such systems used in a wide variety of applications ranging from customer support in the business domain to companionship in the healthcare sector.This paper investigates and compares popular existing chatbots API offerings and then propose and develop a voice interactive and multilingual chatbot that can effectively respond to users' mood, tone, and language using IBM Watson Assistant, Tone Analyzer, and Language Translator. Chatbots guarantee that all the data will be made accessible to any individual day in and day out.**

*Keywords*— **OPEN AI, Chatbot, NLP, API, Speech Recognition, Artificial Intelligence**

## I. INTRODUCTION

In a globalized world where communication is cross-border, multilingual chatbot creation has become an important field for study and innovation. In order to improve user experiences and bridge linguistic gaps, this article explores the field of multilingual chatbots and examines their importance, difficulties, and developments in various linguistic communities [1]. Chatbots are becoming a necessity in several industries [2]. Conversational agents like Apple's Siri, Microsoft's Cortana and XiaoIce, Google Assistant, Facebook Messenger bots, IBM's Watson Assistant, and Amazon's Alexa have all been developed with a significant financial commitment [3]. Still a work in progress, though, is the creation of chatbots that resemble emotional intelligence. To improve the user experience of a particular technology, it can be essential to comprehend complex user behaviour under a variety of scenarios and conditions [4]. Chatbots should be able
to do more than just comprehend concepts offered in text-based queries; they should also be able to analyze facial expressions and interpret text/voice to better replicate human emotional skills [5].

The goal of this paper is to give the advance features which are not present in any other chatbots or to use them subscription is needed. This will help common people to use this chatbot without any hurdles and any language barrier, and applications in the field of multilingual chatbot creation. It explores the complexities of language processing, translation, and adaptation in chatbot frameworks, looking at the difficulties brought on by cultural settings, linguistic subtleties, and user expectations [6].

Chatbots can operate in two domains: open, where users can carry on a conversation from anywhere, like social media platforms, or closed, where input and output options are limited, like customer service or shopping assistants [7]. By investigating this further, we hope to add to the current conversation about multilingual chatbots and highlight how they may promote inclusiveness, cross-cultural dialogue, and international communication in the digital era. The objective of our research is to stimulate additional investigation, creativity, and execution endeavours with the goal of utilising multilingual chatbots to enhance human communication across language boundaries in a globalised society [8]. By employing chatbots to assist students in carrying on discussions in English, Arabic, or a variety of other languages, language and translation education also gains from the potential of AI technologies. Enhance their grammar,

*International Journal of Engineering Applied Sciences and Technology, 2024*
**Vol. 8, Issue 11, ISSN No. 2455-2143, Pages 202-206**
**Published Online March 2024 in IJEAST (http://www.ijeast.com)**

vocabulary, and four language abilities (writing, speaking, reading, and listening).

## II. OBJECTIVE

**1**.Evaluate the effectiveness of multilingual voice chatbot in facilitating seamless communication across different language speakers.
**2.**Examine the scalability of multilingual voice chatbots in accommodating diverse language requirements and user populations.
**3.** Explore the potential challenges and opportunities in designing conversational user interfaces for multilingual voice chatbots.
**4.** Investigate the role of machine learning techniques in improving the language understanding and response generation capabilities of multilingual voice chatbots.
**5**.Explore strategies for enhancing the adaptability and personalization features of multilingual voice chatbots to better cater to individual user preferences and linguistic nuances.

## III. LITERATURE REVIEW:

A multilingual voice chatbot's background research would normally require reading through the body of literature and researching new advancements in number of the important areas:

**1. Aditya Jain1*, Gandhar Kulkarni, Vraj Shah [10] NLP, or natural language processing:**
Recognise the fundamental ideas and methods of natural language processing (NLP), such as entity recognition, sentiment analysis, generation, and language interpretation. Examine the difficulties and applications of NLP in multilingual settings.

**2. Elliott, W. S. [11] Chatbot Development**:
Examine how chatbots have changed over time, moving from rule-based to conversational agents driven by AI. Analyse the many platforms, frameworks, and architectures used to create voice-activated chatbots.

**3. Tomalin, B. [12] Multilingual Computing:**
Examine the fundamentals and techniques of multilingual computing, such as localization techniques, Unicode standards, character encoding, and language detection.

**4. Om Prakash Prabhakar, Navneet Kumar Sahu [15] Voice Recognition and Synthesis:**
With an emphasis on multilingual applications, examine the developments in voice recognition and synthesis technologies, such as text-to-speech (TTS) synthesis, prosody modelling, and automatic speech recognition (ASR).

**5. Mr. H. P. Khandagale, Ms. Shraddha Vaibhav Mane [13] Use Cases and Applications:**
Examine actual use cases and applications of multilingual speech chatbots in a variety of industries, including customer service, healthcare, education, and entertainment, to comprehend the advantages and practical ramifications of these systems.

**6. [14] H. P. Khandagale, R. Zambare, P. Pawar, P. Jadhav, P. Patil, S. Mule [14] Emerging Trends and Future Directions:**
Examine new developments, obstacles, and prospects in the field of multilingual voice chatbots. Some of these include cross-platform compatibility, multimodal interaction, integration of AI assistants, and the possible influence of cutting-edge technologies like edge computing and blockchain.

## IV. PROPOSED SYSTEM:

Speech-to-text and text-to-speech conversion operations are the main means of enabling conversational interactions between users and an AI chatbot in the proposed system architecture. It starts by taking user voice input, which can be in any language. After that, this input is processed by a "Speech-to-Text" component, which uses Automatic Speech Recognition (ASR) technology to convert spoken language into text.
The user's question is forwarded to an API for processing by the system once the speech has been translated into text. Here, the system makes use of the "OpenAI API," most likely alluding to the OpenAI Completion API, which can produce imaginative text replies in response to cues.

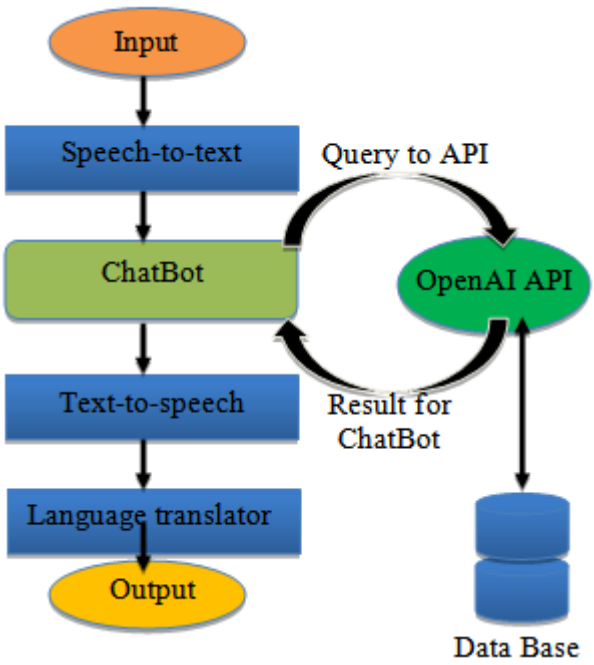



Figure: Proposed System for AI voice ChatBot.

*International Journal of Engineering Applied Sciences and Technology, 2024*
**Vol. 8, Issue 11, ISSN No. 2455-2143, Pages 202-206**
**Published Online March 2024 in IJEAST (http://www.ijeast.com)**

The "ChatBot" component, which uses the OpenAI API to operate as a large language model (LLM), is the brains of the system. These models are skilled at comprehending and producing human-like responses to a broad range of cues and inquiries because they have been trained on enormous volumes of text data. The chatbot's interaction with the user is captured in the answer that the OpenAI API generates, which offers insightful and contextually relevant data.

After the written response is generated, the "Text-to-Speech" component of the system then turns it back into speech. This phase is essential for providing the chatbot's response in an audible manner so that users can interact with the system via speech-based natural language interfaces.

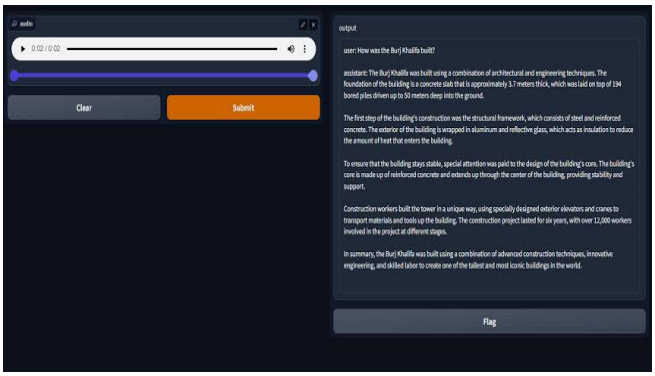
The user receives the system's output, which consists of the synthesized speech response, closing the conversational loop. Users may easily communicate with the system and get instant answers to their questions.

The graphic also includes a "Database" element, implying that a database might be integrated to provide pertinent data to the chatbot in addition to its responses. For example, when the system receives a weather query, it may first use the OpenAI API to formulate its response after retrieving the most recent weather data from the database.

In order to enable smooth and natural interactions between users and the AI chatbot, the system architecture as a whole includes a complex conversational AI framework that smoothly blends speech recognition, massive language model processing, and text-to-speech synthesis.

**Tools/Technologies: OpenAI API, tts, whisper, gradio, LLM, SpeechRecognition.**

## V. IMPLENTATION:




**1. Recording Voice of the User:**

First, the user's voice input is captured to start the adventure. The SpeechRecognition library is used in this situation. It serves as a link, translating spoken words into text that is comprehensible to your chatbot. This library is multilingual, but depending on the languages you wish to cover, you may require additional language packs to provide a truly multilingual experience.

**2. Determining Language:**

Determining the language that the speech recognition produced is essential after you have the text. Whisper excels in this situation. As a language investigator, Whisper examines the text to determine the language that is being used correctly. In order to later route the user's query to the relevant LLM, this step is crucial.

**3. Making the Most of LLMs' Power:**

This is the Large Language Model (LLM), which serves as the conversation's central component. To meet the needs of your users, you'll need an LLM that supports different languages. Popular choices are Jurassic-1 Jumbo and OpenAI's GPT-3. Here, you'll transmit the user's inquiry (if needed translated to the LLM's supported language) over the OpenAI API, and you'll get the LLM's perceptive response.

**4. Converting Text to Speech Again:**

The LLM responds with text, but we need to turn that back into audio so the user can hear it for a voice-based experience. Text-to-speech (TTS) libraries come into play here. Several well-liked options are Microsoft Azure Text-to-Speech, Amazon Polly, and Google Text-to-Speech. Choose a library that offers the languages you wish to be supported by your chatbot.

**5. Optional User Interface:**

Gradio is a useful tool for chatbot design and testing, however it's not necessarily required for functionality. With a text box to show the chatbot's response in an understandable and interactive manner and a microphone for voice input, you may develop an intuitive user experience.

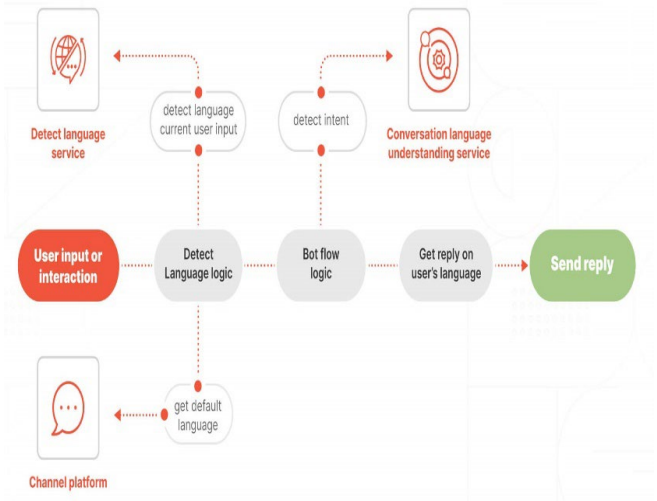
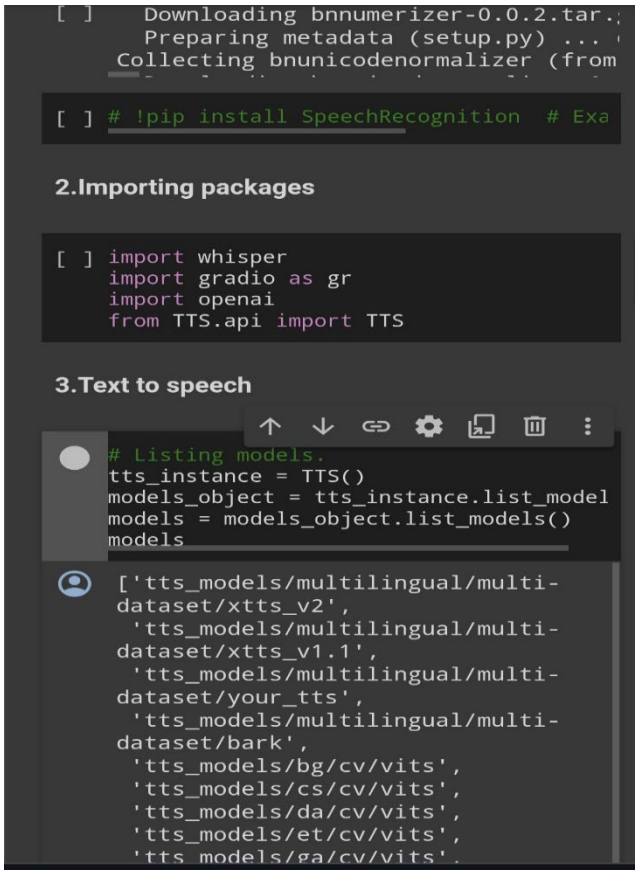



Figure: Implementation of Proposed System

*International Journal of Engineering Applied Sciences and Technology, 2024*
**Vol. 8, Issue 11, ISSN No. 2455-2143, Pages 202-206**
**Published Online March 2024 in IJEAST (http://www.ijeast.com)**




## VI. CONCLUSION:

To sum up, the suggested conversational AI system running on Gradio provides a reliable and effective way to communicate with users using natural language input. Because of its incorporation of automatic voice recognition (ASR), spoken inquiries may be converted into text with ease, improving accessibility and user experience. The system can deliver precise and contextually appropriate answers on a variety of topics by utilizing the OpenAI API as a large language model (LLM). Incorporating text-to-speech synthesis also guarantees users hear output that is clear and natural-sounding, which improves the conversational experience in general. Furthermore, by adding to the data obtained from outside sources, the possible use of a database component improves the system's capacity to deliver well-informed responses. All things considered, this system promises accuracy and user happiness in interactions, showcasing a smart and effective approach to conversational AI.

## VII. REFERENCES

[1]. Aleedy Moneerh 1,2(B) , Atwell1 Eric, 2017 ,'Chatbots for ,University Education'University of Leeds,Leeds LS29JT, UKscmmale @leeds.ac.uk, (pp. 67-81).

[2]. L. Vaira, M. A. Bochicchio, M. Conte, F. M. Casaluci, and A. Melpignano, 2018, "MamaBot: a System based on ML and NLP for supporting Women and Families during Pregnancy," in Proceedings of the 22nd International Database Engineering & Applications Symposium,: ACM, (pp. 273-277).

[3]. Shum H.-Y., X.-d. He, and D. Li, 2018 ,"From Eliza to XiaoIce: challenges and opportunities with social chatbots," Frontiers of Information Technology & Electronic Engineering, vol. 19, no. 1, (pp. 10-26).

[4]. M. Mostafa, T. Crick, A. C. Calderon, and G. Oatley, 2016 ,"Incorporating emotion and personality-based analysis in user-centered modelling," in International Conference on Innovative Techniques and Applications of Artificial Intelligence,: Springer, (pp. 383-389).

[5]. Justo A. V, dos Reis J. C, I. Calado, R. Bonacin, and F. R. Jensen, 2018 ,"Exploring Ontologies to Improve the Empathy of Interactive Bots," in IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), (pp. 261-266).

[6]. Ralston Kennedy ' 2019 , A Voice Interactive Multilingual Student Support System using IBM Watson',School of Computing Queen's University Kingston,ON,Canadaralson @cs.queensu.ca, (pp. 383-389).

[7]. B. Denny, vol. 2019, ed, 2016 ,"Deep Learning for Chatbots," in WILDML: Artificial Intelligence, Deep Learning, and NLP (pp 6-18).

[8]. Yuhao Chen 2019 ,'A Voice Interactive Multilingual Student Support System using IBM Watson' , School of Computing Queen's University Kingston, ON, Canada14yc37@queensu.ca, (pp. 273-277).

[9]. Na-young, K., Cha, Y.J., Kim, H.-S.: 2019 "Chatbots and artificial intelligence. Multimedia-Assist". Lang. Learn.22, (pp.32–53).

[10]. Jain Aditya, Kulkarni Gandhar, Shah Viraj DOI:10.26438/ijcse/v6i1.161167 ,Natural Language Processing International Journal of Computer Sciences and Engineering 6(1):161-16

[11]. Elliott, W. S. "Computer-Aided Mechanical Engineering: 1958 to 1988." Computer-Aided Design, vol. 21, no. 5, 1989, (pp. 275–88).

[12]. Tomalin, B. Doi: 10.29366/2018tlc.2.4.10 ,Multilingual computer assisted language learning (a review). Training, Language and Culture, 2(4), (pp.106-109).

[13]. Mr. Khandagale H. P., Ms. Mane Shraddha Vaibhav May 2023 in IJEAST "INTERACTIVE CHATBOTS: A WAY FOR THE DEVELOPMENT OF THE EDUCATIONAL SECTOR" International Journal of Engineering Applied Sciences and

Technology, 2023 Vol. 8, Issue 01, ISSN No. 2455-2143, (Pages 220-227).

[14]. Khandagale H. P., Zambare R., P. Pawar, P. Jadhav, P. Patil, S. Mule February 2020 "STREET LIGHT CONTROLLER WITH GSM TECHNOLOGY" International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 4, Issue 10, ISSN No. 2455-2143, (Pages 268-271).

[15]. Smadi Kalid Al , Rababah's Mohammad A.A. , 2018 ,"Analytical Survey : Speech Recognition Methods Used In Voice Recognition Techniques" (pp.1-8).