



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 3, March 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

AI Desktop Voice Assistant

Mr.Pawar.V.D¹, Mr.Banegaon.R.P², Savare.A.V³, Patil.S.N⁴, Pujari.V.B⁵, Dhadake.H.P⁶

Vice Principal & HOD, Department of Computer Engineering, Shree Vatavruksha Swami Maharaj Devasthan's

Kai.Kalyanarao(Balasaheb) Ingale Polytechnic College, Akkalkot, Maharashtra, India¹

Lecturer, Department of Computer Engineering, Shree Vatavruksha Swami Maharaj Devasthan's

Kai.Kalyanarao(Balasaheb) Ingale Polytechnic College, Akkalkot, Maharashtra, India²

Student, Department of Computer Engineering, Shree Vatavruksha Swami Maharaj Devasthan's

Kai.Kalyanarao(Balasaheb) Ingale Polytechnic College, Akkalkot, Maharashtra, India³⁻⁶

ABSTRACT: A desktop computer software application called an AI desktop voice assistant is made to help users and carry out tasks via voice commands. It makes use of artificial intelligence technology to comprehend and converse with users in response to their questions and orders. In addition to managing calendars and answering queries, the assistant may also make reminders, provide information, play music, and operate other computer programs. It provides a hands-free, simple method for users to interface with computers in an effort to increase user productivity and convenience.

KEYWORDS: Artificial Intelligence, Conversational Interface, Natural Language Processing (NLP), Speech Recognition, Desktop, Voice Assistant, SpeechCommands.

I. INTRODUCTION

An artificial intelligence (AI) desktop voice assistant is a piece of software that uses voice commands to carry out activities on a desktop computer and offer assistance. It provides customers with an easy-to-use, hands-free method of interacting with their computers.

These voice assistants are meant to converse with users by comprehending and reacting to their questions and orders. Users can speak with the assistant in a more casual and human-like manner by using natural language processing (NLP) to interpret and comprehend spoken language.

Artificial intelligence (AI) desktop voice assistants can do a lot of things, like playing music, controlling other computer programs, organizing calendars, asking inquiries, giving information, and more. They provide a smooth and effective means of computer interaction with the goal of increasing user productivity and ease.

Google Assistant, Microsoft Cortana, and Apple's Siri are a few well-known AI desktop voice assistants. With the development of AI technology, these assistants are always changing and getting better, which enables them to respond to user inquiries in a more precise and tailored manner.

All things considered, AI desktop voice assistants provide a practical and effective means of interacting with your computer, simplifying tasks and increasing accessibility through voice commands.

J.A.R.V.I.S. stands for Just A Rather Very Intelligent System. Tony Stark's AI helper, J.A.R.V.I.S., manages and maintains all of the internal systems of Stark's structures and the Iron Man suits. In the Iron Man novel written by Peter David, J.A.R.V.I.S. Facebook uses Jarvis as a testbed for both its present and future AI systems. Jarvis, whose name derives from J.A.R.V.I.S., or "Just a Rather Very Intelligent System," an artificial intelligence program featured in the 2008 superhero movie Iron Man, is voiced by actor Morgan Freeman. Refer to AI and virtual assistants. Jarvis, which is written in Python, PHP, and Objective C, employs a number of artificial intelligence techniques, including as natural language processing, speech recognition, face recognition, and reinforcement learning. The voice-based AI assistant JARVIS was created in the programming language Python. It adds new, distinctive features by utilizing various technologies. With just a single voice command, tasks can be automated. This AI assistant is desktop-based. Some virtual assistants, such as Microsoft's Cortana, Apple's Siri, Amazon's Alexa, and Google's Google Assistant, are well-known to us. This approach is specifically made to function The Python programming language. It uses a variety of technologies to create new, unique features. It is possible to automate tasks with a single voice command. This AI

helper runs on a desktop. We are familiar with a few virtual assistants, like Google Assistant, Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant. This strategy is designed specifically to work.

II. LITERATURE REVIEW

[1] Well, I had the similar thought before I started making my very own “Digital” Personal Assistant. Though it is not as capable and high as like Amazon’s Alexa or Google Assistant, Home or Apple’s Siri or JARVIS from Iron Man. Nowadays, People are troubled by typing commands into the computer. Be it procrastination or a busy schedule. Typing is a big obsolete process. The solution to this is that we switch over to an assistant which understands us and do the initial work for us. An assistant is the best replacement for typing commands.

[2] With time, computers have become increasingly significant tools that are also getting cheaper. The goal of the personal virtual assistant is to provide a trustworthy, affordable, and simple to use helper. The term "virtual assistant" (VA) refers to computer simulated environments that can approximate physical presence in both real-world and fictional settings. A real-time and interactive technology is a virtual assistant. It implies that the computer can instantly alter the virtual reality in response to user input. The user's perception of being a part of the action in their environment is enhanced through interaction and its gripping power. A high-level encounter can be had by utilising all human sensory pathways. The majority of virtual assistant environments today are primarily visual, shown on a computer screen, but some simulations also contain extra sensory data, such sound through speakers or headphones. The development of virtual assistants has shown promise in a number of fields, including training simulators, medicine and health care, rehabilitation, education, engineering, scientific visualisation, and the entertainment sector. The software functions similarly to Siri and Google Assistant. Yet, the primary focus of the application is the computer. A voice assistant is a digital assistant that helps people through gadgets and voice recognition software by using speech synthesis, natural language processing, and voice recognition. The foundation of this research is speech recognition, one of the fundamental ideas in artificial intelligence.

[3] In our project, we mainly use voice as communication means, so the basically the Speech recognition application. Same kind of application is also developed by the Google that is “Google Voice Search” which is used for in Android Phones. But this Application mostly works with Internet Connections. The concept of speech technology really encompasses two technologies: Synthesizer and recognizer. A speech synthesizer takes as input and produces an audio stream as output. A speech recognizer, on the other hand, does the direct opposite. It takes an audio stream as input and thus turns it into text transcription.

[4] What is voice assistant and how it works. Many of us might have already known about this voice assistant and we use this in our day-to-day life. A voice assistant is a digital assistant that uses voice recognition, language processing algorithms, and voice synthesis to listen to specific voice commands and return relevant information or perform specific functions as requested by the user. A brief description is given about them in this chapter. Speech is an effective and natural way for people to interact with applications, complementing or even replacing the use of mice, keyboards, controllers, and gestures. A handsfree, yet accurate way to communicate with applications, speech lets people be productive and stay informed in a variety of situations where other interfaces will not. Speech recognition is a topic that is very useful in many applications and environments in our daily life. Generally speech recognizer is a machine which understands humans and their spoken word in some way and can act thereafter. A different aspect of speech recognition is to facilitate for people with functional disability or other kinds of handicap. To make their daily chores easier, voice control could be helpful. With their voice they could operate the light switch turn off/on or operate some other domestic appliances. This leads to the discussion about intelligent homes where these operations can be made available for the common man as well as for handicapped.

[5] Everything in the twenty-first century is trending toward automation, whether it's your home or your transportation. Over the past couple of years, there has been an incredible development, or rather advancement, in new tech. You can, presume it or not, but you can engage with your gadget in current period. What does it mean to engage with a machine? Providing it some input is obvious, but if the input data is not in the traditional method of typing, but rather your own voice? What if you communicate to the computer, give it commands, and just want the system to engage with you as if it were your private assistant? What about if the system does more than simply display user the best outcome? What about if it also advises them on a good option? The revolutionary method of human system interchange is to easily accessible machine via voice commands. To accomplish this, we must be using a API which converts voice messages into text messages to understand the input. Many companies, including Google, Amazon, and Apple, are attempting to attain this in a more universal manner. Isn't it great how you can make reminders simply by

stating "reminds me to " or "set an alarm" with wake me up at ? Recognizing the significance of this, we decided to create a platform that can be installed anywhere in the neighborhood and can be asked to assist anybody with anything simply by chatting with it. Furthermore, you can link two similar devices via Wi-Fi and have them interact with one another in the future. These devices can be highly useful for day-to-day use and can assist you perform better by providing you with frequent alerts and updates. Why else would we require it? As our own voice is becoming a better input device than just a standard enter key. All Operating Systems offer a plethora of apps and services to users. The most well-known iPhone application is "SIRI," which enables people to communicate with their phones via voice commands and responds to voice instructions. Google has also created a similar programme, "Google Assistant", which is utilized on Android smartphones. However, that application relies heavily on Internet connections. However, the proposed system may operate with or without using Internet connectivity, taking input from users in the form of speech or text and processing it before returning the outcome in various formats such as action to be taken. Voice-controlled based home automation technologies could provide consumers with a much more comfortable living and make routine tasks easier. Voice control in energy conservation building is especially advantageous for those with disabilities, allowing them to live a previously unattainable lifestyle. Implementing voice activated systems might provide significant benefits, including aid with tasks at work.

[6] Rather Very Intelligent System (J.A.R.V.I.S.) J.A.R.V.I.S. is an AI that functions as Tony Stark's assistant, running and taking care of all the internal systems of Stark's buildings and the Iron Man suits. In Peter David's novelization of Iron Man, J.A.R.V.I.S. Jarvis is a testbed of current and future AI systems used by Facebook. Actor Morgan Freeman is the voice of Jarvis, while the name comes from J.A.R.V.I.S., "Just a Rather Very Intelligent System," the artificial intelligence program in the 2008 superhero film Iron Man. See virtual assistant and AI. Jarvis uses several artificial intelligence techniques, including natural language processing, speech recognition, face recognition, and reinforcement learning, written in Python, PHP, and Objective C. JARVIS is a Voice- Based AI Assistant which is developed in Python Programming Language. It uses Different Technologies to Add New Unique Features. It can Automate Tasks with just One Voice Command. It is a Desktop Based AI Assistant. We know some of the virtual assistants, like Google's Google Assistant, Apple's Siri, Amazon's Alexa, and Microsoft's Cortana. this method is specifically designed to work effectively on desktops. Personal Assistant application code ameliorate.

VII. METHODOLOGY

Speech Recognition: The voice assistant transcribes spoken words into text by using speech recognition technology. To do this, audio input must be analyzed in order to identify specific words and phrases. Natural Language Processing (NLP): After the spoken words are translated into text, NLP methods are used to interpret the user's intentions and commands. To do this, the text must be parsed and analyzed in order to retrieve pertinent information. Intent Recognition: Based on the input it receives, the voice assistant uses machine learning techniques to determine the user's intent. This aids in identifying the precise activity or action that the user desires to be carried out by the assistant. Knowledge Base: To deliver precise and pertinent answers to customer inquiries, the voice assistant consults a sizable knowledge base or database of facts. To ensure current information, this knowledge base is expanded and updated on a regular basis. Task Execution: The voice assistant carries out the relevant task or activity after determining the user's purpose. This could entail using the computer to access other programs or services, getting data from the internet, or carrying out system-level operations. Setting. Understanding: In order to deliver more precise and tailored responses, the voice assistant strives to keep the conversation contextually intact. To provide a more customized experience, it considers user preferences and past interactions. Continuous Learning: To get better over time, AI desktop voice assistants frequently use

Fig.1 System Architecture

i. Python: Python is a high-level, interpreted programming language with an OOPs (Object Oriented Programming) foundation. It is a powerful, extremely practical language with a rapid application development (RAD) focus. Python facilitates simple code creation and execution. Compared to other OOP languages, Python may implement the same logic in as little as one-fifth of the code. Python offers a plethora of advantages to everyone. Python is used in so many different contexts that it cannot be restricted to a single task. Due to its increasing popularity, it is now able to participate in some of the most well-liked and intricate procedures, such as data science, machine learning, artificial intelligence, and natural language processing. For all the libraries this project would require, Python has a ton of them. Libraries used by JARVIS include selenium for web automation, Pyttsx3 for text-to- speech, and speech recognition for voice recognition. Python operates fairly efficiently. Small instances typically don't have any issues with efficiency. If you find that a portion of your Python code is using too much time, you can usually make it more efficient by implementing that portion in a lower-level language. As a result, you will have to write far less programming and produce code that is more efficient.

ii. Que.py: is a Python framework designed to convert questions in natural language into database query language queries. It is easily adaptable to various database queries and natural language inquiry types. Thus, you can create a system enabling natural language access to your database with a little bit of coding.

iii. Que.py: Que.py is a Python framework designed to convert questions in natural language into database query language queries. It is easily adaptable to various database queries and natural language inquiry types. Thus, you can create a system enabling natural language access to your database with a little bit of coding.

iv. Python Text to Speech, or Pyttsx3, is an acronym. It is a Python wrapper for text-to-speech synthesis that works across platforms. It is a Python package that works with Linux, Windows, and Mac OS X text-to-speech engines. It is compatible with Python versions 3.x and 2.x. The fact that it is offline is its primary benefit.

v. NLP and Voice Recognition: The desktop voice assistant processes and comprehends speech commands by using natural language processing (NLP) techniques. To accurately interpret the user's requests, this may require tasks like context extraction, intent recognition, language understanding, and speech recognition.

vi. SQLite: The competent library SQLite offers an in-process relational database for effective small-to-medium-sized data set storing. With very few exceptions, it supports the majority of SQL (Structured Query Language)'s standard capabilities. The best part is that most Python users can start using SQLite right away because the sqlite3 module is already included in the standard library of most distributions. You can simply enhance SQLite with your own Python code because it runs in memory alongside your application. A substantial portion of the hooks that SQLite offers are handled by the default library database driver.

Voice Assistant various applications

The idea behind a voice assistant is to enable users to operate a system by speaking commands. Work on the virtual assistant began with the analysis of audio commands using a microphone during its early stages of development [2]. There are numerous voice assistants on the market, including Cortana, Alexa, Google Assistant, and many more. The voice assistant created by Apple that is the oldest on the market is called Siri. Individuals with disabilities who want to utilize computers or laptops have been introduced to the idea of a desktop virtual assistant. This voice recognizer may operate both online and offline, carrying out different tasks in response to user requests and commands. [3]

System Architecture:

Voice assistance system architecture revolves around different phases. The design consists of:

- Taking voice input from the user through microphone.
- Voice Input data recognition and converting it into text.
- Manipulation of data by executing python script
- Synthesizing speech from processed text output [6].

The user provides speech input in the first phase, which is followed by the use of natural language processing (NLP) to turn the speech data into text. It turns speech input from users into computer executables. Python scripts are used to manipulate data in the third phase. The procedure involved in this would be text to voice. In the final step, the result is again given the form of speech [6].



Speech is input into the speech recognition module, which converts it into signals. The central processor processes the transformed text and outputs the required result [2]. A few other installed modules and libraries, including Speech Recognition, Wikipedia, GTTS, date time, OS, pyaudio, requests, and a web browser, are also included in the system architecture[4].

Characteristics:

- It should always be on, even when not in use, and respond quickly when a task is asked for [6].
- It must comprehend the user's speech, taking into account their tone, pronunciation, and fluency. When evaluating the web search, it ought to provide the intended result both audibly and on screen simultaneously.

Difficulties: Since voice assistants can answer to our orders without needing our hands, they may have become indispensable in our daily lives. Despite all of its development and advancement, there are still certain difficulties. Other elements still play a role in recognition even after hundreds of inputs and trainings. The primary factor is the background noise. The technology fails to distinguish verbal input from the user from ambient noise. It must be educated with various sounds and have the ability to filter out outside noise and inputs. In addition to adjusting for background noise, users can improve the effectiveness of the system's interpretation of voice input by altering its pitch [2].

VIII. RESULTS

A desktop AI voice assistant offers users multiple advantages. It provides a practical and hands-free method of interacting with a computer by leveraging voice recognition and artificial intelligence technologies. Simply give your commands or questions to the assistant, and it will comprehend and react appropriately. In addition to managing calendars and answering queries, the assistant may also make reminders, provide information, play music, and operate other computer programs. By eliminating the time and effort that would otherwise be required to complete these operations manually, this helps to increase productivity. A more intuitive and natural engagement is made possible by the voice assistant's conversational interface. Conversations between users and the assistant enhance user-friendliness and personalization of the experience. A smooth communication experience is also made possible by the voice assistant's capacity to comprehend and interpret natural language using technologies like speech recognition and natural language processing (NLP). Users can communicate in a natural way, and the assistant will correctly understand what they say.

IX. CONCLUSION

With the use of voice recognition, natural language processing, and artificial intelligence, the AI desktop voice assistant is a potent technology that offers hands-free, user-friendly support on desktop computers. Voice commands enable users to communicate with their computer, making chores more comfortable and effective. The AI desktop voice assistant can manage tasks, comprehend and reply to user inquiries, offer information, and operate several computer

applications through its conversational interface. By providing a smooth and intuitive computer interface, this technology seeks to improve user experience and productivity.

REFERENCES

- 1). DESKTOP VOICE ASSISTANT Gaurav Agrawal*1, Harsh Gupta*2, Divyanshu Jain*3 , Chinmay Jain*4 , Prof. Ronak Jain*5
https://en.wikipedia.org/wiki/Desktop_Assistant,
- 2). Personal Desktop Voice Assistant - Research Paper 1Sakshi R Jain, 2Prof Feon Jason
https://ijarcce.com/wpcontent/uploads/2023/03/IJARC_CE.2023.12324A.pdf
- 3) Desktop Voice Assistant Nikhil Patel1, Trupti Landge1, Radhika Tiwari1, Arjun Verma1, Prof. Shabana Pathan2
<https://www.irjet.net/archives/V7/i4/IRJETV7I4187.p df>
- 4) AI-based Desktop Voice Assistant <https://ieeexplore.ieee.org/document/10146699>
- 5) Desktop Voice Assistant
Authors: Ujjwal Gupta, Utkarsh Jindal, Apurv Goel, Vaishali MalikDOILink:
<https://doi.org/10.22214/ijraset.2022.42390> Certificate: [View Certificate https://www.ijraset.com/research-paper/desktop-voice-assistant](https://www.ijraset.com/research-paper/desktop-voice-assistant)
- 6) AI-BASED DESKTOP VOICE ASSISTANT Shubham Thorbole1 , Anuradha Pandit2 ,Gayatri Raut3 , Tejas Sirsat4 Department of E&TC, SKNCOE,SPPU, Pune <https://www.jetir.org/papers/JETIRFX06069.pdf>
- 7) JARVIS” AI DESKTOP VOICE ASSISTANT USING SPEECH RECOGNITION
[HTTPS://PUBS.AIP.ORG/AIP/ACP/ARTICLE- ABSTRACT/2721/1/070031/2904605/JARVIS-AI- DESKTOP- VOICE-ASSISTANT-USING- SPEECH?REDIRECTEDFROM=FULLTEXT](https://pubs.aip.org/aip/acp/article-abstract/2721/1/070031/2904605/JARVIS-AI-DESKTOP-VOICE-ASSISTANT-USING-SPEECH?REDIRECTEDFROM=FULLTEXT)



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379

doi[®]
CROSS **ref**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details