



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

# Virtual Voice Assistant using Raspberry Pi

Sumit Hanbar<sup>1</sup>, Gaurav Gondake<sup>2</sup>, Pratik Gotakhinde<sup>3</sup>, Sourabh Dudhagi<sup>4</sup>

U.G. Student, Department of Electronics and Telecommunication, Kolhapur Institute of Technology, Kolhapur,  
Maharashtra, India<sup>1,2,3,4</sup>

**ABSTRACT:** The aim of this research to demonstrate the execution of “Virtual Assistant Using the Raspberry Pi” which will help to carry out number of tasks or commands and follow the instructions provided by a person or user. These task are relevance to the information or command given by the user, considering the area and locality of a person, and potential to get the facts, knowledge, data, statistics etc. with the help of various online sites or internet (such as to track the stock market prices, score of game, news, listening to music, watching videos and movies, checking the weather forecast, asking for products, checking groceries prices, notifications updates from different kind of social media applications, making calls, checking time, to get information from Wikipedia, Google etc. ...).

The main purpose of using Raspberry Pi as a prime hardware is to take the commands from the user in the form of voice. After taking the commands from the user in voice format, the commands can be converted to text with the use of speech to text engine. The resulting text later will be used to get the proper suitable information. Finally, when the requested information is available, it is then again converted to speech using “text to speech” and related information will be given to the user as output. In addition to this, with the help of some other additional modules implementation can be done to use the concept of keywords matching.

This can be a great use to the person or community who has any kind of temporary or permanent kind of vision loss. It will help them to easily connect to the rest of the world, they can have easy access to worldwide information, also it will be more convenient to use powerful tools like the internet with ease. This can be also used as a great way for entertainment or a source of education. As complementary to it, introducing IOT to this model we can fully automate any system from individual level to industrial level.

**KEYWORDS:** Virtual assistant, Raspberry Pi, Voice assistant, Text to speech, speech to text, gadget, automation, speech recognition

## I. INTRODUCTION

The main purpose of a voice-controlled system is to use the voice of a user as an input for the system, understand the meaning of the command process, and provide the relevant information in the form of voice output. Almost many voice-controlled systems require basic components such as speech to text converter, processing the query, and text to speech convertor.

Speech recognition is a technology that has an ability for computers to understand the language spoken by a person and to give the output according to input.

Text to speech is a process of changing machine recognized text into a language which easily can be understood by the user/speaker when the output text is read out loud.

The field of speech recognition is improving day by day and a number of innovations are being introduced. The improvements are because of high usage of big data and deep learning in this field. Machine learning is also having a positive impact due to the improvements done.

## II. SYSTEM ARCHITECTURE

### Present System-

The current approach limits the possibilities of technological improvements that could be deemed obsolete. There are some systems that process Natural Languages and put them to work using mobile phones. However, there is still a need for a dedicated gadget that can be relied upon. As a result, the current system is not very reliable. Certain virtual assistants are accessible as dedicated devices, which are substantially more expensive. As a result, it is no longer commercially viable. Despite the availability of gadgets such as Amazon Echo and Google Home, Furthermore, in order to accomplish the duties in the existing system, there must be some physical interaction with the gadgets to wake them up. It is possible to leverage open-source platforms like Jasper to construct voice-controlled applications. To

address all of the shortcomings with the present system, our proposal relies on a single python script that takes the least amount of configuration and setup and makes installation simple.

In the previous system, it had the drawback of only giving responses to predefined voices. Hence there is limitation in voices. Due to this, the user cannot get the complete information.

### Proposed system-

Our proposed system has capability which can overcome the drawback of the previous system. Our system design consists of a text to speech method. Following the command, anything the system receives as input will be output in the form of voice, i.e., speech.

## III. HARDWARE IMPLEMENTATION

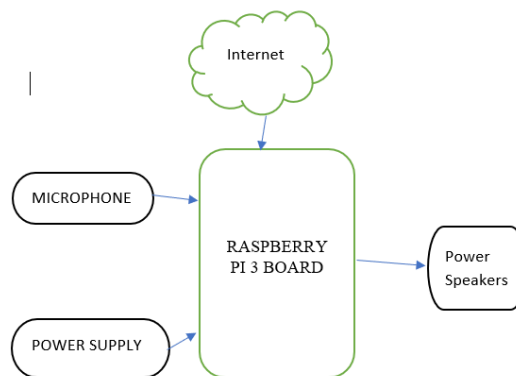


Fig. Block Diagram of System

### 1) Raspberry PI –

The Raspberry Pi is a credit-card-sized computer that was inspired by the BBC Micro in 1981 and was originally meant for education. It is a low-cost gadget that can be used to create a variety of applications. It is powered by a 1.2GHz quad-core ARM Cortex A53 processor. It has a 1 Gigabyte DDR2 RAM module with a 900MHz clock speed. It comes with built-in Wi-Fi and Bluetooth. The Raspberry Pi is at the heart of the voice command system, as it is engaged in every step of data processing and component connection. The Raspbian operating system is installed. The SD card is then inserted into the card slot to offer a backup. A steady 5V, 2.1 mA power source is required for the Raspberry Pi. This can be done with an AC power supply and a micro-USB charger, or with a power bank.

### 2) Speaker –

Speakers, after the user's query has been processed, the query's text result is translated to speech utilising an online text to speech converter. Now, this voice, which is an audio output, is conveyed to the user through the audio out speakers.

### 3) Microphone-

It is used to capture the sound's audio input. This audio input would be searched for keywords as it progressed through the system. These keywords are critical for the voice command system's operation, as our modules operate on the principle of searching for keywords and producing output by matching keywords, as illustrated in the block diagram.

### 4) SD card-

The SD card is an important component of the Raspberry Pi since it serves as the first storage location for the operating system and files. Many sorts of USB-connected accessories can be used to expand storage.

When the Raspberry Pi is 'switched on,' that is, connected to a power supply, a special piece of code known as the bootloader runs, reading more special code from the SD card that was used to start up the Raspberry Pi. It will not start if there is no SD card inserted. If you insert or remove an SD card while the Raspberry Pi is powered on, the data on the SD card is likely to be corrupted.

The SD card must be formatted, or written on, in a specific way so that the Raspberry Pi can read the data it requires to get up and running. If this is your first time, read the instructions or purchase a pre-formatted SD card. The

Raspberry Pi's card reader has a maximum throughput of 25 MB/s, and read and write speeds are unlikely to exceed 22 MB/s.

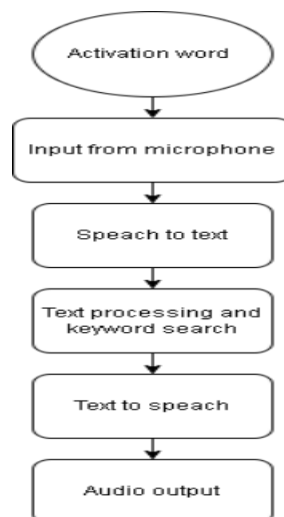
#### 5) Internet-

The voice command system's internet connection is provided via Ethernet. We need a continual connection to perform all of this because the system relies on online text to speech conversion, online query processing, and online speech to text conversion.

### IV. FLOW OF EVENTS IN VOICE COMMAND SYSTEM

Practically the system starts when ever we call it (when ever the word to activate the system is said ) basically if we divide the process in form of steps it would be

- 1- The word is heard by the systems by the microphone if the word is the activation word the system will respond and will ask what it can do for you for example activation word can be AXELA ,HEY GOOGLE.. etc
- 2- The device will hear through the microphone and is fed to the computer as an input for processing it further .
- 3- The voice is then fed to voice to text converter where the voice is converted to text for further recognition and processing by the computer
- 4- The text is parsed to search the keyword in the program
- 5- The system matches the keyword which are already declared and carry out the process defined regarding the keyword
- 6- But the output from the program is not in form of audio
- 7- So the program output should be converted to audio using text to speech convertor
- 8- The output is transmitted through the speaker connected to the raspberry pi module



### V. MODULES

#### Wikipedia Module

A keyword of “wiki” activates this module , if a request is made related to need of some information then request is made to wikipedia module,The information is provided in a form of short note summary of information  
If the information is not available an error message “sorry cant find information related to ... ”is passed as an output .

#### News module

A “News” keyword in a speech activates the module. Top headlines are searched using google news and passed as an output. It also search for a particular news which the used demands .if the system cant find the news related to the requested query a message as “News related to your topics is not available”



### **Weather**

When the user ask for weather details the module activates and ask for location if it is unable to detect the user location.the keyword “weather” activates the module , details such as temperature,humidity, Weather predictions etc is the output.it generates an error message on not able to retrieve information

### **SpeechRecognition:**

This module is responsible for speech recognition and processing Human speech , it recognize the word which triggers a response from the virtual assistance

### **pyttsx3:**

This module is the reverse of SpeechRecognition , it converts the text command to speech (Text -to-speech) , and works at the final stage of the process . it is cross platform and also work while not connected to internet

### **Webbrowser:**

This command is used to open browsers or apps such as youtube etc.

### **Joke**

This command responds with a joke, it works when a keyword “joke” is recognized it fetches random jokes and say it out

### **Unclear**

Much unclear speech can be input to the device to tackle unclear messages. This module is used to generate an error message if the input does not match with the declared keywords and answer with “Sorry I can't hear you, can you please repeat once more ”.

### **Other Modules**

Many additional module specific to the user requirement are added such as “Set an alarm”, “time”,etc. A keyword will respond with fetching the data from module and providing to the user

## **VI. RESULT**

The voice assistant takes command from the user and matches it with the query which is written in the program,if mathed it proceeds to for the next command and retrieve the information and send it to the user. As per the user requirement it can change the voice of the assistance using some id/command.

## **VI. CONCLUSION**

In the mobile sector, voice command systems such as Siri, Google Now, Cortana, and Alexa have huge potential in the future. Additionally, this lays the path for a future linked house using Internet of Things, voice command systems, and software computer vision.This system helps users to retrieve information through voice commands which makes their life easier.

## **VII. ACKNOWLEDGEMENT**

The authors extremely grateful to Dr.Nitin Sambare (HOD of Electronics and telecommunication department) , Prof. Paresh Sawant sir and respected teachers of Kolhapur institute of technology, college of engineering (Autonomous), Kolhapur for their guidance and endless support throughout the completion of research paper.

We also thank the board of directors for providing us the opportunity to present a research paper and project on Virtual voice assistant using Raspberry Pi, and to all friends and others who support us either mentally or physically.

## **REFERENCES**

1)Dahl, George E.; Yu, Dong; Deng, Li; Acero, Alex (2012). "Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition". IEEE Transactions on Audio, Speech, and Signal Processing. 20 (1): 30–42.



- 2)M. R and D. Subramaniyan, "Personal Assistant and Intelligent Home Assistant via Artificial Intelligence Algorithms(Raspberry PI/Pineapple)", *Impact: International Journal of Research in Engineering & Technology (IMPACT: IJRET)*, vol. 4, no. 6, pp. 9-13, 2016.
- 3)[SSW4-2001] *4th ISCA Tutorial and Research Workshop (ITRW) on speech synthesis (SSW4)*, Blair Atholl Palace Hotel, Perthshire, Scotland, August 29 - September 1, 2001, ISCA Archive,
- 4)Sutar Shekhar, Pophali Sameer, Kamad Neha, Prof.Devkate Laxman, " An Intelligent Voice Assistant Using Android Platform" , March 2015, IJARCSMS, ISSN: 2327782
- 5)Tung, Liam. "Google Home Mini Flaw Left Smart Speaker Recording Everything." ZDNet(October 11, 2017). <http://www.zdnet.com/article/google-home-mini-flaw-left-smart-speaker-recording-everything/>.



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

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



**www.ijirset.com**

Scan to save the contact details