

# Creating Desktop Speech Recognition Using Python Programming.

**Mrs.A.M.Sermakani.,M.E.,(Ph.D)<sup>1</sup>, J.Monisha<sup>2</sup>, G.Shrisha<sup>3</sup>, G.Sumisha<sup>4</sup>**

Associate Professor,Department of Information Technology,S.A Engineering College,Chennai<sup>1</sup>

UG Scholar Department of Information Technology,S.A Engineering College,Chennai<sup>2-4</sup>

**Abstract:** As a personal assistant, this project is built with AI technologies, Artificial intelligence technologies are beginning to be actively used in human life, this is facilitated by the appearance and wide dissemination of the Internet of Things (IOT). Autonomous devices are becoming smarter in their way to interact with both a human and themselves. The advancement in technology over time has been unmeasurable. In such an era of advancement if people are still struggling to interact with their machine using various input devices then its not worth it. For this reason, many voice assistants were developed and are still being improved for better performance and efficiency. The main task of a voice assistant is to minimize the use of input devices like keyboard, mouse, touch pens, etc. This will reduce both the hardware cost and space taken by it.

## 1.INTRODUCTION

The development of artificial intelligence (AI) systems that are able to organize a natural human-machine interaction (through voice, communication, gestures, facial expressions, etc.) are gaining in popularity. One of the most studied and popular was the direction of interaction, based on the understanding of the machine by the machine of the natural human language. It is no longer a human learns to communicate with a machine, but a machine learns to communicate with a human, exploring his actions, habits, behavior and trying to become his personalized assistant.

The work on creating and improving such personalized assistants has been going on for a long time. These systems are constantly improving and improving, go beyond personal computers and have already firmly established themselves in various mobile devices and gadgets. One of the most popular voice assistants are Siri, from Apple, Amazon Echo, which responds to the name of Alex from Amazon.

These types of virtual assistants are very useful for old age, blind & physically challenged people, children, etc. by making sure that the interaction with the machine is not a challenge anymore for people. Even blind people who couldn't see the machine can interact with it using their voice.

The understanding and executing commands are reached a new level like the virtual assistant of the iron man named Jarvis. This is although fictional yet this is what that can be achieved using virtual assistants. All you need to do is give a command to the assistant and the rest will be performed by the assistant. With the help of voice-activated virtual assistants, there will be no need to write long codes to perform a task, the system will do so for us. The machine will work in three modes- supervised, unsupervised or reinforcement learning depending upon the usage for which the assistant is developed.

## 2.RELATED WORK

The intelligent assistant applies his own specific methods and approaches for development, which in turn affects the final product. One assistant can synthesize speech more qualitatively, another can more accurately and then without additional explanations and corrections perform tasks, others are able to perform a narrower range of tasks, but most accurately and as the user wants. Obviously, there is no universal assistant who would perform all tasks equally well.

The set of characteristics that an assistant has depends entirely on which area the developer has paid more attention. Since all systems are based on machine learning methods and use for their creation huge amounts of data collected from various sources and then trained on them, an important role is played by the source of this data, be it search systems, various information sources or social networks. The amount of information from different sources determines the nature of the assistant,

which can result as a result. Despite the different approaches to learning, different algorithms and techniques, the principle of building such systems remains approximately the same. Figure 1 shows the technologies that are used to create intelligent systems of interaction with a human by his natural language. The main technologies are voice activation, automatic speech recognition, Teach-To-Speech, voice biometrics, dialog manager, natural language understanding and named entity recognition.



VOICE TECHNOLOGY	BRAIN TECHNOLOGY
Voice activation	Voice biometrics
Automatic speech recognition (ASR)	Dialog management
(Teach-to-speech (TTS))	Natural language understanding (NLU) Named entity recognition (NER)

### 3. LITERATURE SURVEY:

Abhay Dekate (2016) presented in the Modern Era of fast moving technology we can do things which we never thought we could do before but, to achieve and accomplish these thoughts there is a need for a platform which can automate all our tasks with ease and comfort. Thus we need to develop a Personal Assistant having brilliant powers of deduction and the ability to interact with the surroundings just by one of the materialistic form of human interaction i.e. Human Voice. Dr. Kshama V. Kulhalli (2017) et al. proposed the Most famous application of iPhone is "SIRI" which helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user. 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. But our Proposed System has capability to work with and without Internet Connectivity. Veton Kėpuska (2018) proposed one of the goals of Artificial intelligence (AI) is the realization of natural dialogue between humans and machines. In recent years, the dialogue systems, also known as interactive conversational systems are the fastest growing area in AI. Deny Nancy (2019) presented in the Modern Era of fast moving technology we can do things which we never thought we could do before but, to achieve and accomplish these thoughts there is a need for a platform which can automate all our tasks with ease and comfort.

Deepak Shende (2019) et al. presented artificial intelligence technologies are beginning to be actively used in human life, this is facilitated by the appearance and wide dissemination of the Internet of Things (IOT). Autonomous devices are becoming smarter in their way to interact with both a human and themselves. Kishore Kumar R1 (2018) et al. presented to develop an economically effective and performance wise efficient virtual assistant using Raspberry Pi for home automation based on the concepts of Internet of Things, Speech Recognition, Natural Language Processing and Artificial Intelligence. People who are using it can give voice inputs and the device itself responds through voice commands by itself.

Rutuja V. Kukade (2018) et al. proposed there are various communication barriers for people who are blind, and they have to face various challenges. In this paper, we have discussed the implementation of a personal virtual assistant which can take the human voice commands to perform tasks which otherwise would need the dependence on others. Isha S. Dubey (2019) et al. proposed about a different combination of a reading machine (OCR), virtual assistant and Domotics system using Raspberry-Pi which will be a combination of a great system. M. A. Jawale (2019) et al. proposed in today's world, many artificial intelligence applications developed using programming languages like Python, R and so on. Tushar Gharge (2019) et al. presented the problem of user while developing a computer program. Developing a computer program is not an easy task it needs hardware resources which user have to handle. While continuously typing the code there may be possibility of injuries to the fingers of the user. To avoid the problems we are designing a system in which the computer program can be developed through the voice.

### 4. EXISTING SYSTEM:

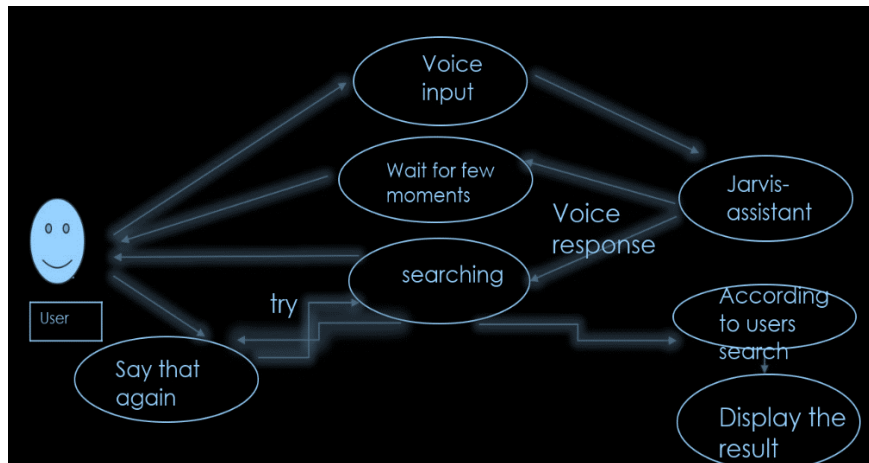
Desktop personal assistant is built with AI technologies, here Jarvis assists the end-user with day-to-day activities like general human conversation, searching queries in various search engines like Google, searching for videos in youtube, retrieving images, files in local drive, play songs in desktop by the use of imported modules like pyttsx3 is an offline module that is used for text to speech conversion in python, DateTime module is imported to support the functionality of date and time. web browser, Wikipedia, Google those are imported to fetch the user result. The user statements/commands are analysed with the help of Artificial intelligence to give an optimal solution.

**5. PROPOSED SYSTEM:**

The proposed system will have the following functionality:

- (a) The system will keep listening for commands and the time for listening is variable which can be changed according to user requirements.
- (b) (b) If the system is not able to gather information from the user input it will keep asking again to repeat till the desired no. of times .
- (c) The system can have both male and female voices according to user requirements.
- (d) Features supported in the current version include playing music, emails, texts, search on Wikipedia, opening anything on the web browser, etc.
- (e)

**flow diagram:**

**6. IMPLEMENTATION:****STEP1:**

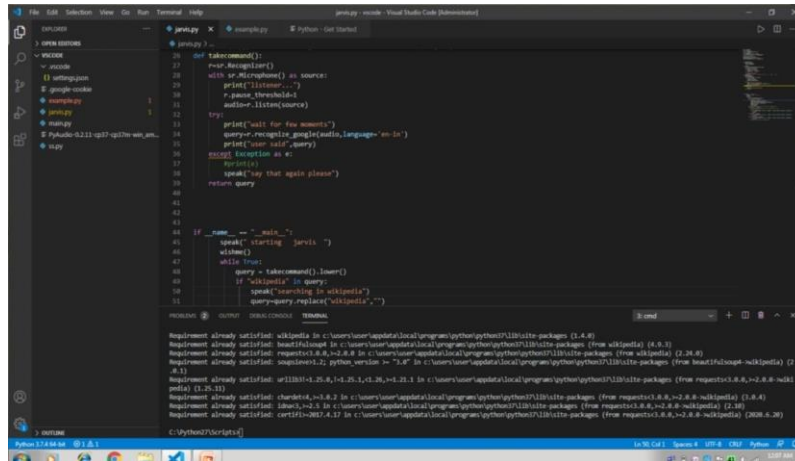
**(import packages for this program):**

- Import pytsx3
- Import datetime
- Import speech\_recognition as sr
- Import Wikipedia
- From googlesearch import search
- From youtube\_search import youtubearch
- Import urllib.request
- Import os
- Import webbrowser

**PIP INSTALLATION IN THIS PROGRAM WHICH IS ESSENTIAL:**

- PIP INSTALL PYLINT
- PIP INSTALL PYTTX3
- PIP INSTALL SPEECH RECOGNITION
- PIP INSTALL WIKIPEDIA
- PIP INSTALL GOOGLE
- PIP INSTALL YOUTUBE-SEARCH
- PIP INSTALL WEBBROWSER

**Step2:**Definition of functions

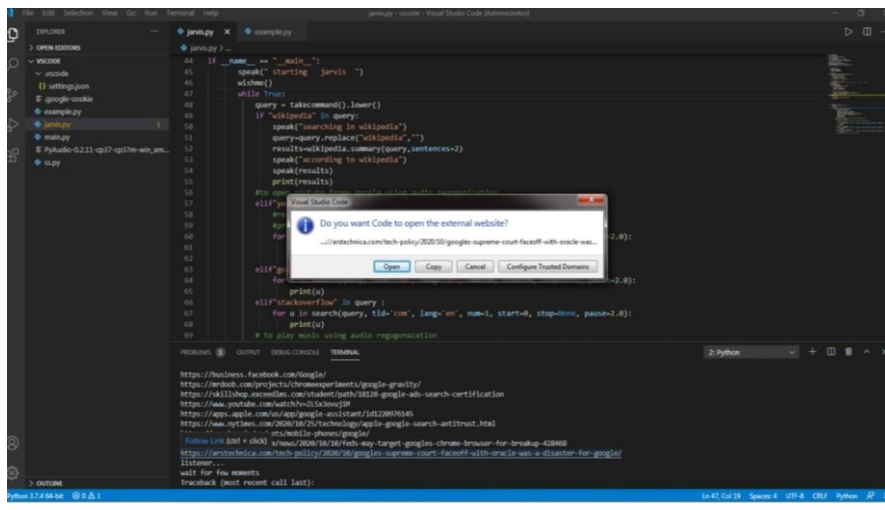


```
import requests
import random
import sys

def takeCommand():
    if takeCommand() == "quit":
        sys.exit()
    elif takeCommand() == "help":
        print("I am Jarvis AI")
    elif takeCommand() == "search":
        query = takeCommand().lower()
        if "wikipedia" in query:
            speak("searching in wikipedia")
            results = wikipedia.summary(query, sentences=2)
            speak("according to wikipedia")
            speak(results)
            print(results)
        elif "youtube" in query:
            speak("opening youtube")
            results = youtube.search(query, max_results=5)
            for i in results:
                print(i)
            break
        elif "google" in query:
            speak("opening google")
            results = search(query, tld="com", lang="en", num=1, start=0, stop=None, page=2)
            for u in results:
                print(u)
            break
        elif "stackoverflow" in query:
            speak("opening stackoverflow")
            results = search(query, tld="com", lang="en", num=1, start=0, stop=None, page=2)
            for u in results:
                print(u)
            break
        else:
            speak("I don't know about that")
    else:
        speak("I don't know about that")

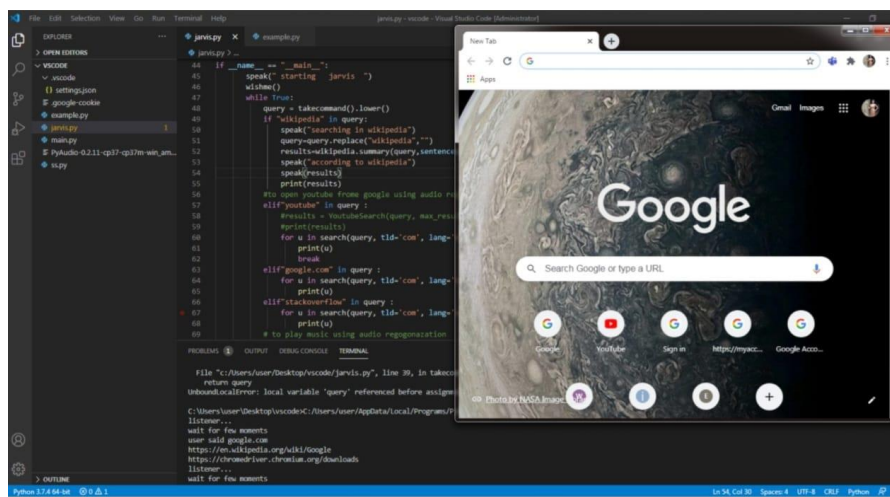
if __name__ == "__main__":
    speak("starting Jarvis")
    while True:
        query = takeCommand().lower()
        if "wikipedia" in query:
            speak("searching in wikipedia")
            results = wikipedia.summary(query, sentences=2)
            speak("according to wikipedia")
            speak(results)
            print(results)
        elif "youtube" in query:
            speak("opening youtube")
            results = youtube.search(query, max_results=5)
            for i in results:
                print(i)
            break
        elif "google" in query:
            speak("opening google")
            results = search(query, tld="com", lang="en", num=1, start=0, stop=None, page=2)
            for u in results:
                print(u)
            break
        elif "stackoverflow" in query:
            speak("opening stackoverflow")
            results = search(query, tld="com", lang="en", num=1, start=0, stop=None, page=2)
            for u in results:
                print(u)
            break
        else:
            speak("I don't know about that")
```

Step3:Run the program



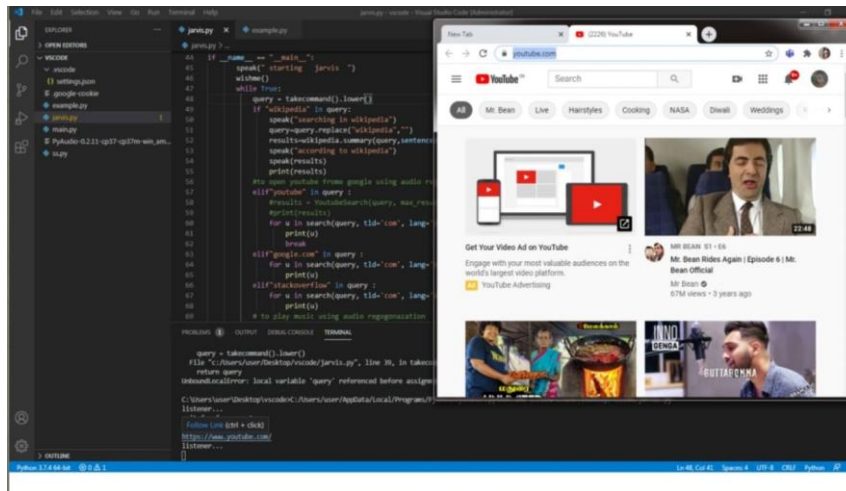
```
https://business.facebook.com/google/
https://webhook.com/jarvis/chromeos/taunt/google-gravity/
https://killshop.ameedias.com/student/path/3823-google-ad-search-certification
https://www.youtube.com/watch?v=3k10v19
https://apps.apple.com/us/app/google-assistant/id128979345
https://www.nytimes.com/2020/10/25/technology/apple-google-search-artifact.html
https://www.foxnews.com/tech/2020/10/18/tech-may-target-google-chrome-browser-for-breakup-62868
https://arxiv.org/abs/2008.10188/tech-may-target-google-chrome-browser-for-breakup-62868
listener...
wait for few moments
frackack (last recent call last):
```

Step4:According to user command output of Google

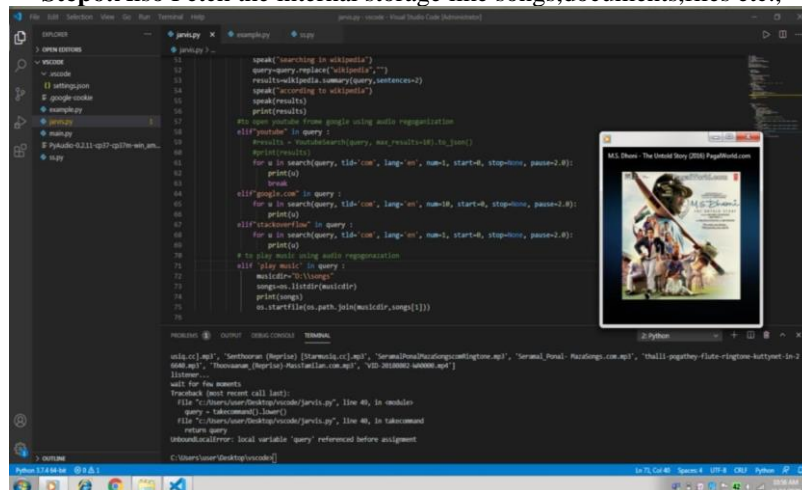


```
File "C:/Users/Abhishek/Desktop/vscode/jarvis.py", line 39, in takeCommand
return query
UnboundLocalError: local variable 'query' referenced before assignment
C:/Users/Abhishek/Desktop/vscode/C:/Users/Abhishek/AppData/Local/Programs/Python/Python38-64/Python.exe
listener...
wait for few moments
user said google.com
https://om.wikipedia.org/wiki/google
https://chromeosuser.chromium.org/downloads
listener...
wait for few moments
```

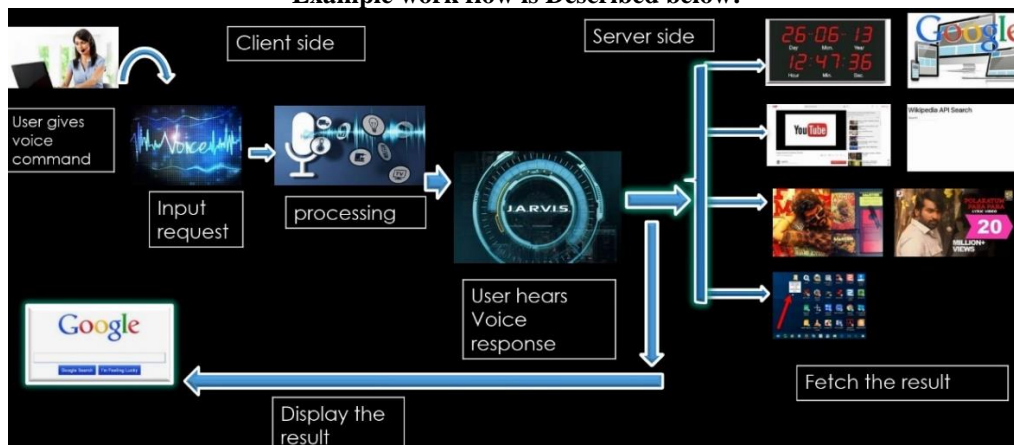
Step5:According to user command output of youtube



Step6: Also Fetch the internal storage like songs, documents, files etc.,



### Example work flow is Described below:



### FUTURE SCOPE :

The virtual assistants which are currently available are fast and responsive but we still have to go a long way. The understanding and reliability of the current systems need to be improved a lot. The assistants available nowadays are still not reliable in critical scenarios. The future of these assistants will have the virtual assistants incorporated with Artificial Intelligence which includes Machine Learning, Neural Networks, etc. and IoT. With the incorporation of these technologies, we will be able to achieve new heights. What the virtual assistants can achieve is much beyond what we have achieved till now. Most of us have seen Jarvis, that is a virtual assistant developed by iron man which is although fictional but this has set new standards of what we can achieve using voice-activated virtual assistants



**CONCLUSION:**

In this paper, we discussed the design and implementation of a Digital Assistance. The project is built using open source software modules with visual studio backing which can accommodate any updates in the near future. The modular nature of this project makes it more flexible and easy to add additional features without disturbing current system functionalities. This examination speaks to an initial phase in investigating the potential job of remote helpers in programming improvement ventures finished by virtual groups. This project will be help for visually impaired and physically challenge people. Instead, we will see a fragmented marketplace emerge. It will be a market where you are might into using default AI providers depending on the hardware purchase. This will lead to consumer friction and third party solutions to remove incumbent solution.

**REFERENCES**

- [1] Subhas mani Kaushal, megha Mishra (IRJET) volume:07 Issue:09 sep 2020
- [2] Dhiraj Pratap singh, Deepika Sherawat, Sonia school of computer science and engineering, Galgotias University volume 26, Issue 6, 2020.
- [3] Kei Hashimoto, Junichi Yamagishi, William Byrne, Simon King, Keiichi Tokuda, "An analysis of machine translation and speech synthesis in speech-to-speech translation system" proceedings of 5108978-1-4577-0539- 7/11/\$26.00 ©2011 IEEE.
- [4] B. Marr, The Amazing Ways Google Uses Deep Learning AI.
- [5] Cortana Intelligence, Google Assistant, Apple Siri.
- [6] Hill, J., Ford, W.R. and Ferreras, I.G., 2015. Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*, 49, pp.245-250.
- [7] Abhay Dekate, Chaitanya Kulkarni, Rohan Killedar, "Study of Voice Controlled Personal Assistant Device", *International Journal of Computer Trends and Technology (IJCTT)* – Volume 42 Number 1 – December 2016.
- [8] Deny Nancy, Sumithra Praveen, Anushria Sai, M. Ganga, R.S. Abisree, "Voice Assistant Application for a college Website", *International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019.
- [9] Deepak Shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhangre, "AI Based Voice Assistant Using Python", *Journal of Emerging Technologies and Innovative Research (JETIR)*, February 2019, Volume 6, Issue 2,
- [10] M. A. Jawale, A. B. Pawar, D. N. Kyatanavar, "Smart Python Coding through Voice Recognition", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-10, August 2019.
- [11] Tushar Gharge, Chintan Chitroda, Nishit Bhagat, Kathapriya Giri, "AI-Smart Assistant", *International Research Journal of Engineering and Technology (IRJET)*, Volume: 06 Issue: 01, January 2019. 10. Veton Kępuska, "Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google
- [12] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, "An Intelligent Chatbot using Natural Language Processing". *International Journal of Engineering Research*, Vol. 6, pp. 281-286, 1 May 2017.