



J.A.R.V.I.S - AN AI TECHNOLOGY BASED PERSONAL WINDOWS ASSISTANT SYSTEM

¹Susmita Das, ²Ankita Saha, ³Aparupa Chakraborty, ⁴Swagata Thakur, ⁵Sumanta Chatterjee

¹Student, ²Student, ³Student, ⁴Student, ⁵Assistant Professor

Department of Computer Science and Engineering,

JIS College of Engineering (An Autonomous Institution), Kalyani, India

Abstract: *JARVIS* draws its inspiration from virtual assistants like google assistant for android, and Siri for IOS. It has been designed to provide a user-friendly interface for carrying out a variety of tasks by employing certain well-defined commands. Users can interact with the assistant through voice commands. The approach of this paper is to propose *JARVIS* – it is a Digital Life Assistant which uses mainly human communication like assist you in any installation process, and open various applications. This software aims at developing a personal assistant for Windows-based systems. The main objective of the software is to perform the tasks of the user at certain commands, provided in speech. It can do Wikipedia searches for the user, by accessing websites like Google, YouTube, etc., in a web browser, and many more. This research paper aims to develop a personal assistant for desktops. Hence, a hassle-free experience using a single command. It has been a challenging task while working on our virtual assistant.

Index Terms–*JARVIS*, Artificial Intelligence, Speech Recognition, Synthesizer, Recognizer, Personal Windows Assistant.

I. INTRODUCTION

With the advancement of technology, all know about Siri, Google Assistant, and many other virtual assistants which are designed to aid the tasks of users on IOS, and Android platforms. But to our surprise, there is no such virtual assistant available for the paradise of developers i.e., the windows platform. In today's era, almost all tasks are digitalized. On desktops, the user has to work with a keyboard and mouse, but a blind person can't work on that, that's why we introduced *JARVIS*. Using *JARVIS*, the user only has to give the command, the task is done. In this voice assistant, the user has to say "Open YouTube" or "Open Google" and then the *JARVIS* will open those applications that are the task of a Virtual Assistant. Virtual Assistants are software programs that help users, make it easy with their day-to-day tasks, such as showing weather reports, creating reminders, etc. It can take commands via text (online chatbots) or by voice. Voice-based intelligent assistants need an invoking word or wake word to activate the listener, followed by the command. In this research paper, the wake word is 'Jiniya'. It has been designed to provide a user-friendly interface for carrying out a variety of tasks by employing certain well-defined commands. Users can interact with the assistant either through voice commands or using keyboard input. As a personal assistant, *JARVIS* assists the end-user with day-to-day activities like a general human conversation, searching queries in google, searching for videos, live weather conditions, and word meaning, and reminding the user about the scheduled events and tasks. Virtual assistants turning out to be smarter than ever. We aim to create more and more functionalities that can help humans to assist in their daily life and also reduce their efforts.

II. LITERATURE REVIEWS

In the year 2019, Deepak Shende, Ria Umahiya, Monika Raghorde, Aishwarya Bhisikar, Anup Bhange has published "AI Based Voice Assistant Using Python" on JETIR, in which, the principles of the functioning of voice assistants are described, its main shortcomings and limitations are given. The method of creating a local voice assistant without using cloud services is described, which allows to significantly expand the applicability of such devices in the future.[1] Jatu Naazneen Abdul Gaffar worked on "Virtual Assistant" building personal assistant software (a virtual assistant) is using semantic data sources available on the web, user generated content and providing knowledge from knowledge databases. The main purpose of an intelligent virtual assistant is to answer questions that users may have.[2] "Desktop Assistant AI Using Python" published in IJAR SCT by Abeed Sayyed, Ashpak Shaikh, Ashish Sancheti, Swikar Sangamner, Prof. Jayant H Bhargale, stated that the main task of a voice assistant is to minimize the use of input devices like keyboard, mouse, touch pens, etc. and which also helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user.[3] A paper on IEEE explore named "Artificial Intelligence-based Voice Assistant" authored by S Subhash; Prajwal N Srivatsa; S Siddesh; A Ullas; B Santhosh, stated that this

voice assistant will gather the audio from the microphone and then convert that into text, later it is sent through GTTS (Google text to speech). GTTS engine will convert text into audio file in English language, then that audio is played using play sound package of python programming Language.[4] A paper published in IJEAT named “*The Voice Enabled Personal Assistant for Pc using Python*”, proposed effective way of implementing a Personal voice assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions.[5] Research Paper on “*Desktop Voice Assistant*” published on IJRES, described that the method of creating a voice assistant without using cloud services, which will allow the expansion of such devices in the future. Virtual assistants are able to interpret human speech and respond via synthesized voices.[6] This paper “*JARVIS: An interpretation of AIML with integration of gTTS and Python*” presents JARVIS, a virtual integrated voice assistant comprising of gTTS, AIML[Artificial Intelligence Markup Language], and Python-based state-of-the-art technology in personalized assistant development, where it shows the adoption of the dynamic base Pythons pyttsx which considers intentionally in adjacent phases of gTTS and AIML, facilitating the establishment of considerably smooth dialogues between the assistant and the users. [7] A paper published in IJERT named “*Jarvis, Digital Life Assistant*” has discussed voice recognition algorithms which are important in improving the voice recognition performance. The technique was able to authenticate the particular speaker based on the individual information that was included in the voice signal. The results show that these techniques could use effectively for voice recognition purposes. [8] A paper named “*Jarvis - A Virtual Assistant based on Artificial Intelligence*” published on International Journal of Grid and Distributed Computing, by Dr.M. Sharada Varalakshmi, Dr.P. Lavanya, Sai Prakash Reddy, demonstrates the automation of the electronic devices through artificial intelligence and their long haul explore objective is to create astute frameworks that can bolster human learning.[9] VetonKëpuska, Gamal Bohouta, published a paper on IEEE Xplore, this research paper has used the multi-modal dialogue systems which process two or more combined user input modes, such as speech, image, video, touch, manual gestures, gaze, and head and body movement in order to design the Next-Generation of VPAs model. The new model of VPAs will be used to increase the interaction between humans and the machines by using different technologies, such as gesture recognition, image/video recognition, speech recognition, the vast dialogue and conversational knowledge base, and the general knowledge base.[10]

III. RESEARCH METHODOLOGY

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 opposite. It takes an audio stream as input and thus turns it into text transcription.

1. **Synthesizer:** A speech synthesizer is a computerized voice that turns a written text into a speech. It is an output where a computer reads out the word loud in a simulated voice; it is often called text-to-speech. It is not only to have machines talk simply but also to make a sound like humans of different ages and gender. A speech synthesizer is a computerized device that accepts input, interprets data, and produces audible language. A synthesizer is an electronic instrument that uses some form of digital or analog processing to produce audible sound. The counterpart of the voice recognition, speech synthesis is mostly used for translating text information into audio information and in applications such as voice-enabled services and mobile applications. Apart from this, it is also used in assistive technology for helping vision-impaired individuals in reading text content. There are many technologies involved in the production of speech with speech synthesizers. The two most definitive segments are how the user inputs information to be spoken and how the sounds for the words are actually interpreted and produced.

1.1. The first step to produce the speech is the composition of text to be spoken. In some cases, it is as simple as loading a computer text file into a software program. In other cases, a more complicated input system is required. There are many different input devices, but the most prevalent is a keyboard or other similar typing board (such as a touchscreen). Patients with severe mobility restrictions may instead use a joystick device. Special input devices are created that act as switches. These switches are programmed to accept and decipher the motions of the user, even blinking of the eyes. Essentially any muscular movement can be interpreted as a switch and programmed to produce language.

1.2. The second step is deciphering the input and producing the desired audio speech. Data is gathered or assembled through the input device until the user indicates that the information is complete. The computer then interprets and speaks the words, phrases, or sentences. Complicated logic is involved when translating written text into spoken language. For example, there are many words that are spelled the same, but pronounced differently in different contexts. The software must make that determination.

2. **Recognizer:** Speech recognition, or speech-to-text, is the ability of a machine or program to identify words spoken aloud and convert them into readable text. Rudimentary speech recognition software has a limited vocabulary and may only identify words and phrases when spoken clearly. Speech recognition uses a broad array of research in computer science, linguistics and computer engineering. Many modern devices and text-focused programs have speech recognition functions in them to allow for easier or hands-free use of a device. Speech recognition systems use computer algorithms to process and interpret spoken words and convert them into text. A software program turns the sound a microphone records into written language that computers and humans can understand, following these four steps: analyse the audio; break it into parts; digitize it into a computer-readable format; and use an algorithm to match it to the most suitable text representation. Speech recognition software must adapt to the highly variable and context-specific nature of human speech. The software algorithms that process and organize audio into text are trained on different speech patterns, speaking styles, languages, dialects, accents and phrasings. The software also separates

spoken audio from background noise that often accompanies the signal. To meet these requirements, speech recognition systems use two types of models: Acoustic models. These represent the relationship between linguistic units of speech and audio signal. Language models. Here, sounds are matched with word sequences to distinguish between words that sound similar.

Speech recognition (Fig 1) is an evolving technology. It is one of the many ways people can communicate with computers with little or no typing. A variety of communications-based business applications capitalize on the convenience and speed of spoken communication that this technology enables. Speech recognition programs have advanced greatly over 60 years of development. They are still improving, fueled in particular by AI.

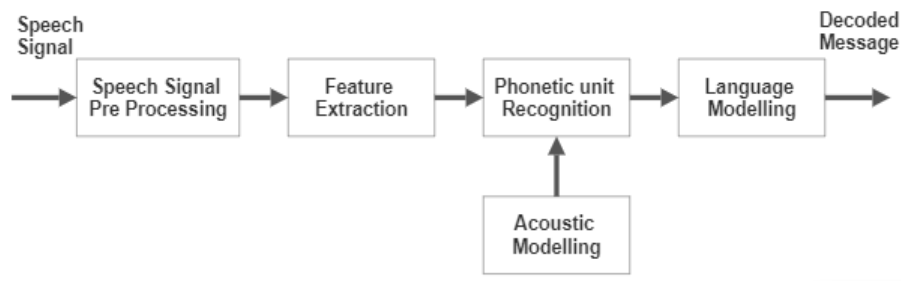


Fig1. Block diagram of Speech Recognition system

IV. WORKING PRINCIPLE

First of all, it will convert all text into speech with pyttsx3 library and for voice recognition we use sapi5. sapi5 is a Microsoft Speech API and we use runaway () function, runandwait function will make the speech audible in the system, if you don't write this command then the speech will not be audible to you. After that it take command from user, for converting spoken words to text we use Speech recognition library and for access microphone we use PyAudio module PyAudio is required if you want to use microphone input. Requests is a HTTP library for the Python programming language. The goal of the project is to make HTTP requests simpler and more human-friendlier. Pywikihow is an unofficial WikiWow python API. Uses BeautifulSoup to scrape WikiHow and return the data we want. The sys module in Python provides various functions and variables that are used to manipulate different parts of the Python runtime environment. It allows operating on the interpreter as it provides access to the variables and functions that interact strongly with the interpreter. The PlaySound library function plays a sound specified by the given file name, resource, or system event, we use this library for add intro audio. PyQt5 is the latest version of a GUI widgets toolkit developed by Riverbank Computing. It is a Python interface for Qt, one of the most powerful, and popular cross-platform GUI library. PyQt5 is a blend of Python programming language and the Qt library. From PyQt5 we import QtWidgets, QtCore, QtGui and from PyQt5.QtCore we import QTimer, QTime, QDate, Qt and from PyQt5.QtCore import QTimer, QTime, QDate, Qt. Web browser module is a convenient web browser controller. Pytube is a lightweight, dependency-free Python library which is used for downloading videos from the web.

Flowchart (Fig. 2): Run the program after that it will convert all the text into voice, after that it will take command from user, if the command is match with predefined command, then it will go further, if not match then it will again take command, when the command is match then it will work according to Instruction which is given by user and display the command result also.

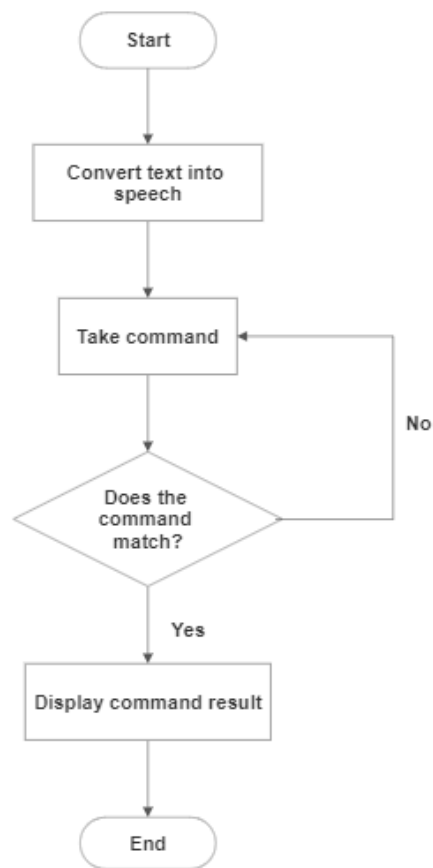


Fig2. Flowchart of the system

V. OUTPUTS AND DISCUSSIONS

GUI (Fig 3): It is used to interact with the user. GUI reflects the basic appearance of the application. There are many graphical user interfaces that you can use with the python programming language. For frontend design, you should be use PyQt5 tools.

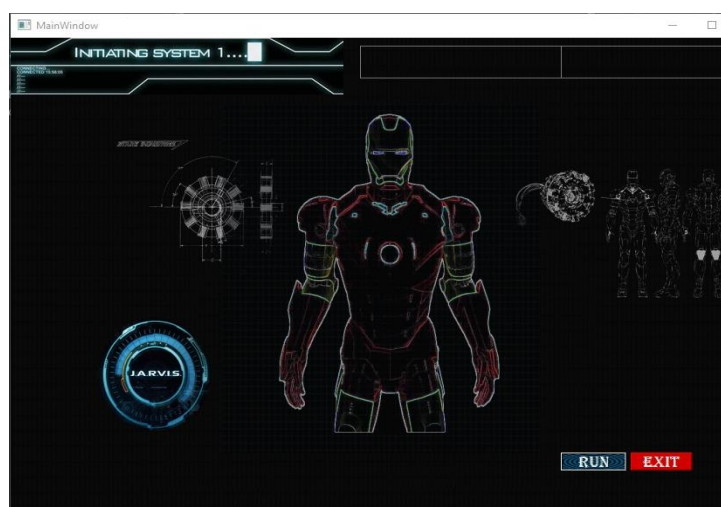


Fig3.GUI Interface

Wakeup Audio and interaction with jiniya (Fig 4): When you run the **JARVIS** then it will play the wakeup audio first. Playsound library is used for playing the wakeup audio. After playing the wakeup audio, Wishme () method is called.

Opening Notepad (Fig 5): Using “Opening Notepad” command, you can easily notepad without typing or clicking anything, it is saving our time. Using OS library, you can access any application which is present in your system.

```

C:\Users\Ankita\Desktop\p\Scripts\python.exe C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x00001D080C5D540>]
Hello Sir, Good Evening, its 07:57 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: hello ginia
hello ma'am, How are you?
Listening...
Recognizing...
user said: I am good what about you
i am also good ma'am, How can I help you?
Listening...
Recognizing...
user said: thank you genius for your service
thanks for using me, have a good day
Process finished with exit code 0

```

Fig4. Wakeup Audio

```

C:\Users\Ankita\Desktop\p\Scripts\python.exe C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x000023698>]
Hello Sir, Good Night, its 08:02 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: please open Notepad
okay, opening Notepad
Listening...

```

Fig5. Open Notepad

Fetching information directly from internet(Fig 6): For knowing information like what is python, how to install it, the command is “Please activate how to do mode”.

Mathematical calculation(Fig 7): For doing calculation the command is “Please activate mathematics mode”.

```

"C:\Program Files\Python310\python.exe" C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x00002F03739750>]
Hello Sir, Good Night, its 08:23 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: please activate how to do mode
How to do mode is activated
please tell, what you want to know?
Listening...
Recognizing...
user said: can you tell how to install Microsoft Visual Studio
Install Visual Studio Using Parallels Desktop on a Mac
1 - Purchase Parallels.
2 - Install Parallels.
3 - Check the system requirements.
4 - Choose a compatible Windows version.
5 - Purchase Windows.
6 - Download Windows.
7 - Start the installation.
8 - Click Continue.
9 - Enter the product key you were given when you purchased Windows.
10 - Pick your integration preference.
11 - Click Continue and choose a name and location.
12 - Wait and let Parallels do its magic.
13 - Start up and use Parallels.

```

Fig6. Fetching information directly from internet

```

"C:\Program Files\Python310\python.exe" C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x0000018D3250A470>]
Hello Sir, Good Night, its 08:24 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: please activate mathematics mode
Mathematics mood is activated
Say what you want to calculate
Listening
534 + 3
your result is
537
Listening...
Recognizing...
user said: ok genius thank you for your service
thanks for using me, have a good day
Process finished with exit code 0

```

Fig7. Mathematical Calculation command

Access Google automatically(Fig8): You can open Google and search anything automatic without typing.

Access YouTube automatically(Fig9): You can open YouTube and search anything automatic without typing.

```

C:\Users\Ankita\Desktop\p\Scripts\python.exe C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x00002A800000000>]
Hello Sir, Good Night, its 08:08 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: please open Google
opening google in microsoft edge
ma'am, what should I search on google
Listening...
Recognizing...
user said: what is the Jarvis in Python
Listening...
Recognizing...
Say that again please...
Listening...

```

Fig8. Access Google automatically

```

"C:\Program Files\Python310\python.exe" C:/Users/Ankita/Desktop/index/jarvis.py
[<pyttts3.voice.Voice object at 0x000002717E800AB0>]
Hello Sir, Good Night, its 08:20 PM
Hi!, I'm Jiniya, How can I help you?.....
Listening...
Recognizing...
user said: open YouTube
ma'am, what should I search on youtube
Listening...
Recognizing...
user said: please play a song for me
which song you want to listen
Listening...
Recognizing...
user said: see you again
Listening...
Recognizing...
Say that again please...
Listening...
Recognizing...
Say that again please...
Listening...

```

Fig9. Access YouTube automatically

Detect live location(Fig10): Using JARVIS you can detect your live location using IP address.

Temperature cheeking(Fig11): You can cheek the current temperature.

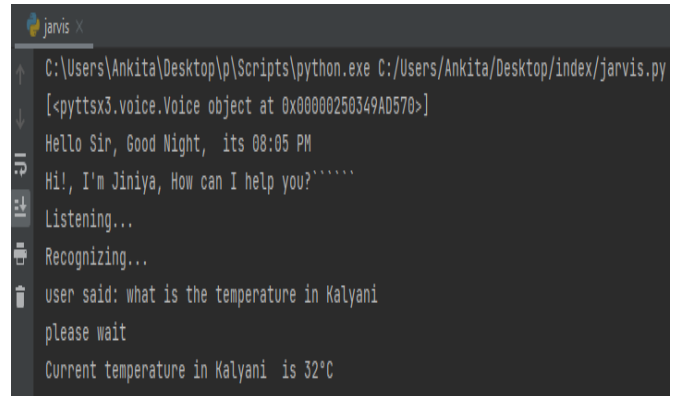


```

C:\Users\Ankita\Desktop\p\Scripts\python.exe C:/Users/Ankita/Desktop/index/jarvis.py
[<pytttsx3.voice.Voice object at 0x000001CB778FD7E0>]
Hello Sir, Good Night, its 08:07 PM
Hi!, I'm Jiniya, How can I help you?''''''
Listening...
Recognizing...
user said: do you know where I am right now
wait, let me check
157.40.69.67
i am not sure, but i think we are in Kolkata city of India country
Listening...

```

Fig10. Detect live location



```

C:\Users\Ankita\Desktop\p\Scripts\python.exe C:/Users/Ankita/Desktop/index/jarvis.py
[<pytttsx3.voice.Voice object at 0x00000250349AD570>]
Hello Sir, Good Night, its 08:05 PM
Hi!, I'm Jiniya, How can I help you?''''''
Listening...
Recognizing...
user said: what is the temperature in Kalyani
please wait
Current temperature in Kalyani is 32°C

```

Fig11. Temperature checking

Through this voice assistant, we have automated various services using a single line command. It eases most of the tasks of the user like searching the web, retrieving weather details, opening applications and doing calculation. We aim to make this research paper a complete server assistant and make it smart enough to act as a replacement for a general server administration.

VI. CONCLUSION AND FUTURE SCOPE

This research paper is based on the concept of artificial intelligence, areas of artificial intelligence, and its techniques. The field of artificial intelligence gives the ability to the machines to think analytically, using concepts. Artificial Intelligence will continue to play an increasingly important role in various fields. We conclude that further research in this area can be done as there are very promising and profitable results that are obtainable from such techniques, while scientists have not yet realized the full potential and ability of artificial intelligence. This technology and its applications will likely have far-reaching effects on human life in the years to come. This review has not attempted to detail all the literature in the area but to report mainly the most recent work. In this paper, we presented a technology to utilize Artificial Intelligent techniques. Artificial Intelligence technology can provide intelligent models and voice-text assistants which are much simpler than traditional physical models.

The plans include integrating JARVIS with mobile using React Native to provide a synchronized experience between the two connected devices. Further, in the long run, JARVIS is planned to feature auto-deployment supporting elastic beanstalk, backup files, and all operations which a general Server Administrator does. We show that it is more effective than considering only the area for which weather forecasting is done. In the future, we have plans to utilize some other features like sending emails, chatting on WhatsApp or any other social sites through voice, accessing YouTube, health recommendations based on symptoms and reminding the user about the scheduled events and tasks, and low-cost Internet of Things (IoT) devices, such as temperature and humidity sensors.

VII. REFERENCES

- [1]Deepak Shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhange, "AI Based Voice Assistant Using Python," Journal of Emerging Technologies and Innovative Research(JETIR), Published in: Volume 6 | Issue 2 | Year February – 2019, e-ISSN: 2349-5162, Country: Nagpur, Maharashtra, India.
- [2] Jatu Naazneen Abdul Gaffar, "Virtual Assistant",under the guidance of Prof. Abhijit Palse, University of Mumbai, DTSS College of Commerce, published on <https://www.academia.edu/>, 2019, Mumbai, India.
- [3] Abeed Sayyed, Ashpak Shaikh, Ashish Sancheti, Swikar Sangamnere, Prof. Jayant H Bhangale, "Desktop Assistant AI Using Python", International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 6, Issue 2, June 2021.
- [4] S Subhash; Prajwal N Srivatsa; S Siddesh; A Ullas; B Santhosh, "Artificial Intelligence-based Voice Assistant", published in: 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), Date of Conference: 27-28 July 2020, Date Added to IEEE Xplore: 01 October 2020, INSPEC Accession Number: 20023658, DOI: 10.1109/WorldS450073.2020.9210344, Publisher: IEEE, Conference Location: London, UK.
- [5] V. Geetha, C.K. Gomathy, Kottamasu Manasa Sri Vardhan, Nukala Pavan Kumar, "The Voice Enabled Personal Assistant for Pc using Python", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-10 Issue-4, April 2021.
- [6]Vishal Kumar Dhanraj, Lokeshkriplani, Semal Mahajan,"Desktop Voice Assistant", International Journal of Research in Engineering and Science (IJRES) ISSN (Online): 2320-9364, ISSN (Print): 2320-9356 www.ijres.org Volume 10 Issue 2 | 2022 | PP. 15-20, Feb 2022.

- [7] Ravivanshikumar Sangpal, Tanvee Gawand, Sahil Vaykar, Neha Madhavi, "JARVIS: An interpretation of AIML with integration of gTTS and Python" Published in: 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), DOI: 10.1109/ICICT46008.2019.8993344 Publisher: IEEE, Conference Location: Kannur, India.
- [8] ShrutiKa Khobragade, "Jarvis, Digital Life Assistant", published on International Journal of Engineering Research and Technology (IJERT), Paper ID : IJERTV2IS1237, ISSN (Online) : 2278-0181 Pune, India. Posted in Volume 02, Issue 01 (January 2013).
- [9] Dr.M. Sharada Varalakshmi, Dr.P. Lavanya, Sai Prakash Reddy, "Jarvis - A Virtual Assistant based on Artificial Intelligence", International Journal of Grid and Distributed Computing, Published 2020-11-01, Issue Vol. 13 No. 2 (2020)
- [10] VetonKëpuska, Gamal Bohouta, "Next-Generation Of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa And Google Home)", Published in: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), DOI: 10.1109/CCWC.2018.8301638, Date Added to IEEE Xplore: 27 February 2018, Conference Location: Las Vegas, NV, USA.
- [11] Dr.VasudhaVashisht, Ankit Rai, Dikshant Sharma, Ansh Sharma, AjinuEapenMathew, " Perceptive Personal Voice Assistant", Indian Journal of Research in Applied Sciences Engineering (IJRASE), 25th April 2022.
- [12] R Sivapriyan; N Sakshi; T Vishnu Priya, "Comparative Analysis of Smart Voice Assistants", Published in: 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Date Added to IEEE Xplore: 25 January 2022, DOI: 10.1109/CSITSS54238.2021.9683722, Location: Bangalore, India.
- [13] Tae-Kook Kim, "Short Research on Voice Control System Based on Artificial Intelligence Assistant", Published in: 2020 International Conference on Electronics, Information, and Communication (ICEIC), Date Added to IEEE Xplore: 02 April 2020, DOI: 10.1109/ICEIC49074.2020.9051160, Location: Barcelona, Spain.
- [14] Peng Cheng; Utz Roedig, "Personal Voice Assistant Security and Privacy—A Survey", Published in: Proceedings of the IEEE (Volume: 110, Issue: 4, April 2022), Page(s): 476 – 507, DOI: 10.1109/JPROC.2022.3153167, Funding Agency: 10.13039/501100001602-Science Foundation Ireland.
- [15] VadaboyinaAppalaraju; V Rajesh; K Saikumar; P. Sabitha; K Ravi Kiran, "Design and Development of IntelligentVoice Personal Assistant using Python", Published in: 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Date Added to IEEE Xplore: 09 March 2022, DOI: 10.1109/ICAC3N53548.2021.9725753, Location: Greater Noida, India.