

Empowering Students and Enabling Accessibility- A survey on Personal Voice Assistant for Students Education and Enhanced Accessibility

Shreya Yadav¹, Gauri Dhingra², Sarang Gupta³ and Betty Paulraj⁴

¹⁻⁴Department of Computer Science and Engineering, Amity School of Engineering and Technology, Uttar Pradesh, Noida, India

Abstract— Personal voice assistants have revolutionized how people interact with technology by enabling effortless access to a variety of services by making the user experience natural and more effective using Artificial Intelligence (AI) and Natural language processing (NLP) technologies to understand spoken instructions and reply the information with the most appropriate information or actions that are appropriate to the current context. The voice assistant can now be the ultimate time saver thanks to its wide range of functionalities, which range from basic tasks like setting alarms and playing music on YouTube to more advanced ones like understanding text through paragraphs and responding to any questions inside it using NLP. Personal voice assistants have shown great promise in changing human-computer interaction and promoting a more comfortable, personalized, and natural experience with computing.

Index Terms— Voice assistant, Speech Recognition, Python, Text-To-Speech (TTS)

I. INTRODUCTION

The way we live, work, and communicate with the world surrounding us has changed as a result of technological advancements in the current digital age. Among these revolutionary developments, personal voice assistants have become indispensable helpers that make chores easier, offer us knowledge, and provide us unmatched convenience. But what if a voice assistant could go beyond practicality and serve as a lifesaver for people with special needs? Enter Jinn, a revolutionary personal voice assistant that has been meticulously developed to meet the demands of both students and people with special needs. Jinn offers a wide range of features that make it an invaluable tool for users from all walks of life.

Jinn helps in closing the accessibility gap by providing ways for people who are visibly impaired to easily access technology making books, documents, and internet material more accessible than ever before by enabling blind users to interact with written text through speech recognition and image-to-text processing. Jinn also emerges as a crucial study partner for students. They can instruct Jinn to "remember" critical notes, lectures, and deadlines thanks to its capacity for storing as well as retrieving information. Students have access to a world of knowledge because of Jinn's seamless connection with OpenAI's enormous knowledge collection, which makes research and learning more effective and interesting. Jinn goes above and beyond accessibility to improve daily living. Users are given navigation support while traveling thanks to its seamless integration with Google Maps. Users may make calls and send messages hands-free thanks to WhatsApp's fully automated capabilities, redefining communication

for everyone, especially when traveling. Additionally, Jinn's connection with YouTube provides a tailored and engaging entertainment experience, playing videos, making recommendations, and even summarizing information as necessary.

II. SCOPE

Human beings have created a place in the world where computers can help human beings and assist them in their activities. A student might need some assistance to become smarter than their colleagues or a disabled person with some assistance may become self-reliant. This leads to the development of artificial intelligence and voice assistants, hence its demand.

The scope of our project is to help students and people with disabilities live more easily by enabling them to complete a variety of daily tasks with only a single voice command. The voice assistant can converse with you and uses speech-to-text and text-to-speech modules to read out the website's content. To deliver the outcomes the user has requested, the voice help that has been built communicates with the applications that are intended. Our Voice Assistant tries to provide a direct answer with little effort but with pertinent knowledge. However, when asked about a complete insight the voice assistant can provide the user with the same content. We set out to create a system that would make it possible for every kind of user to use the internet in a way that is special, simple, and convenient for everyone. Visually impaired or people with disabilities and students can easily access their phones and laptops using the features provided by a voice assistant. The aim is to create a personal voice assistant named "Jinn" that enables users to interact with new technology, control their devices, and learn through technology. It is the most advanced voice assistant for people with visual impairments with an approach that improves the quality of the system while allowing the visually impaired user to access features of the desktop through unique custom layout and speech-to-text.

The purpose of this paper is to demonstrate the importance of use of voice assistants in everyday life and explore how voice assistants can be made accessible to students and people with disabilities.

III. LITERATURE REVIEW

In recent years, Voice assistants have advanced from simple voice recognition systems to strong digital allies with advanced features. This study of the literature offers an overview of the current research and advancements in the realm of voice assistants, with a particular focus on those that have advanced functionality and accessibility features, like the Jinn voice assistant. The best technique of recognition is speech recognition, which makes humans identical and makes it easier for computers to recognize them. This has a great reputation and aids with autonomous speech recognition. Dynamic Time Warping (DTW), HMM, and other voice recognition methods are some of the most used.[3]

The most difficult and actively researched problem in the fields of artificial intelligence and human language technology is the understanding of human language by robots. Old information retrieval (IR) techniques are becoming obsolete due to the ongoing growth of information available in digital format. This makes search engines more precise, which can only be accomplished by enabling them to deal with documents' semantics (at least in part) and not just view them as a collection of words. Search engines must extract brief, precise text fragments that answer a user's question as part of the question answering (QA) assignment. [4]

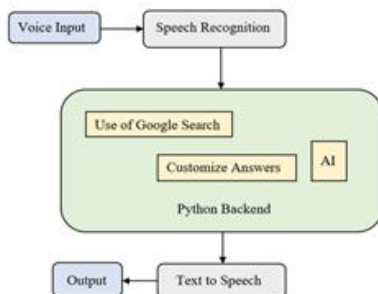


Figure 1: System Architecture [9]

Figure 1 given above shows the basic system architecture flow that is used for building a personal voice assistant using python.

The use of digital assistants with speech recognition capabilities has seen significant advancements in recent years, particularly with the rise of smart devices such as smartwatches, health bands, speakers, Bluetooth earphones, laptops, televisions, and mobile phones. As more and more devices integrate voice assistants, there is a growing need for techniques that can enhance the performance of voice-activated search. To achieve better results, digital assistants need to be trained using Machine Learning and other important techniques such as Artificial Intelligence, Big Data access and management, and the Internet of Things. With the help of these techniques, our devices can learn to adapt to their users' needs and provide personalized results. [5]

Natural language processing (NLP) enables the creation of a ground-breaking assistance. In paper [6], it has been discussed how NLP can enable assistants to become knowledgeable enough to understand instructions in any native language, enabling all members of society to gain from its benefits.

It is understandable why voice assistants are getting more and more common as they become more relevant to provide excellent yet simple customer service. Around the world, 15% of people are disabled in some way, and of those, 2-4% struggle significantly to operate. For people with disabilities, using a website might be an activity that appears simple to most people but is actually very challenging. In contrast to brick and mortar businesses, where accessibility can be achieved by adding ramps for wheelchairs or braille interfaces, the internet is a primarily visual medium, making it susceptible to numerous "accessibility blockers" that might hinder different sorts of websites.[1]

The Mini-Mental State Examination (MMSE) and the Robertson Dysarthria Profile (specifically, the ability to repeat words) are predictive of how well users with disabilities will be able to communicate with a voice assistant, according to the best model (AIC=130.11). Users with physical, linguistic, and cognitive disabilities can efficiently communicate with voice assistants if they have some residual cognitive and language ability [7].

Among the notable VPs are Alexa, Cortana, Siri, and Google Assistant. While each of the four IVAs could answer roughly 17.35% of daily queries, Google Assistant was shown to be the most successful, answering 59.80% of them. Some were disappointed with Siri's response rate, which was only 45%, which was lower than expected. Conversely, Cortana was only able to respond to one-third of the questions. Alexa was clearly outperformed in this competition, answering only 7.91% of the time. Without a question, that was a terrible defeat. As per our survey, only 25% of regular daily users are IVAs, indicating a very low retention rate. Siri, Cortana, Alexa, and GoogleAssistant are all getting better over time at a steady pace. Long-term benefits of IVAs are more than they first appear to be. For a significant portion of the disabled population, who frequently experience cognitive impairments that make it difficult for them to speak and compose entire sentences, a personal assistant might not be a viable alternative. While every tested device has room for improvement, new research can be conducted and the potential of IVAs assessed in a range of unexplored domains, such as sales, banking, business, counselling, and education [8].

IV. PROPOSED METHODOLOGY

We developed an efficient method for handling the personal voice assistant for this project. Many built-in routines in the Speech Recognition library will aid the assistant in comprehending the command given to it and responding vocally utilizing the Text to Speech feature. The underlying algorithms translate the voice into text when the user issues a voice command to the assistant. The voice input is initially converted into text by the speech recognition module. The result of the user's inquiry is then returned when the text has been processed. The final output is voice, which is a conversion of the processed query result. Because the system must first listen to the user, STT takes the longest. The voices of different users vary. While some are simple to hear, others might be challenging to identify. Our overall execution time originates from this. The orders are carried out after the voice has been transformed to text, and the user is then given the outcome.

Making a speech out of the query result is the final stage. The speech is the product at the end. STT takes the longest to complete since the system must first listen to the people. The voices of different users vary. While some are simple to hear, others are difficult to distinguish. Our entire execution time is determined by this phase. Once the voice has been transformed to text, we can begin carrying out commands and reporting back to the users on the outcomes.

This is possible due to the functionality found in a variety of libraries. Additionally, the helper has been able to accomplish all of its tasks with the assistance of multiple APIs. These APIs have been used for calculations, news extraction, web searches, and other purposes. We have made requests to the API and the API has provided us with the relevant response. APIs such as NewsAPI are useful to fetch the latest news headlines. We have employed the JSON package to process JSON data from the website into string format, which allows us to extract news from

the web and pass it to a function to process it further. Additionally, we have used libraries such as Random and a variety of other libraries corresponding to different technologies. The OS library was used to generate operating system related functions, such as turning off or turning on a system. [11]

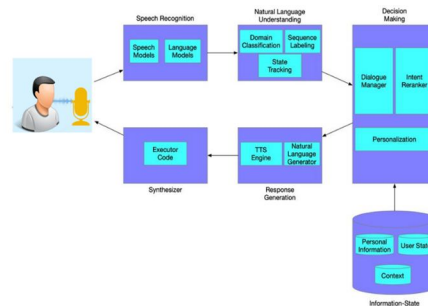


Figure 2: Components of a Voice Assistant

List of functions and features to be implemented in voice assistant and will be performed by it:

- Sending and receiving of mails: Jinn can send emails with complete body or open email messages.
- Searching and surfing on internet/search engines, reading out content from websites: Searching for websites across the internet to reach anywhere on the web with just a simple voice command.
- Desktop applications: Jinn can perform opening and closing operations of various desktop applications, and executing application tasks like setting alarm on clock, playing music on music player, etc.
- Fetching news from the internet: Jinn can bring the latest and most popular news from around the world to keep us updated with what's happening around the globe with the help of NewsAPI.
- Using google translate for multilingual languages: Jinn can support a wide range of languages including multiple regional languages with the help of google translate.
- Notifying current temperature, date and time: Jinn gives reminders with a variety of date and time features, as well as providing us with regular weather updates.

Advanced Features:

- **Fully automated YouTube videos:** Providing full control of YouTube videos like pausing, playing, forwarding, etc. When a user says "Jinn, play" or "Jinn pause," the voice assistant recognises the command and plays or pauses the YouTube video. Furthermore, "Jinn" recognises commands to move forward or backward in the video. It can also increase or decrease volume.
- **Full automation of WhatsApp:** Sending messages to individuals or group chats, making voice and video calls through WhatsApp by just a voice command. The assistant navigates to the chosen conversation and takes the message. The user speaks orders such as "Jinn, make a voice call to [contact name]" or "Jinn, start a video call with [contact name]."
- **Remember/Reset function:** Jinn is made to remember and store some information for reference in future. This information can also be deleted with time. Jinn is asked to remember certain points and tasks which can be accessed when asked upon. Jinn is given memory to actively maintain and store facts, dialogues, preferences, etc.
- **Chatbot:** A text to text system which provides information about anything using a question answer method. The chatbot's functionality begins with understanding and interpreting the user's input using natural language processing (NLP) techniques. The query is dissected by the system to determine the purpose and extract key items or keywords.
- **PDF QnA:** Multiple PDFs can be uploaded and queried upon giving accurate answers of the question asked from the PDF. In this way the reader won't have to read the entire PDF to get an answer to the question. This can be very beneficial for the students.
- **Integrating Open AI:** When a user interacts with the voice assistant, the voice recognition system encodes and processes their voice command. The generated text is then forwarded to OpenAI for analysis. OpenAI's model of language understands the input, determines the intent, and uses extensive training on a variety of textual data in order to produce a concise and content-rich return.
- **Image to Text/Speech:** Jinn can do captioning of an image and provide a brief summary of any image given to it [14].

Some of the modules to be used during the implementation of the project are listed below:

pyttsx3: A cross-platform text-to-speech library without platform restrictions is called pyttsx. The capacity to make use of this library offline is its key advantage for text-to-speech conversion. Linux, macOS, and Windows are just a few of the operating systems which Pyttsx3 may operate on. By manipulating variables like speech rate (words per minute) and volume level, you can influence the voice production. It supports different text-to-speech engines, including the following:

- **sapi5:** Used in case of Windows-based systems.
- **gTTS:** Google Text-to-Speech engine.

Speech Recognition: This library offers users access to speech recognition engines and APIs, making it adaptable and appropriate for use in a range of settings, from voice assistants to services for transcription. It is capable of comprehending speech in numerous tongues and supports many different languages.

Pywhat kit: This is a simplistic library which permits quick interaction with the browser. It involves carrying out inquiries on the internet, delivering emails, running computations, and even more. It is an adaptable structure that makes it possible for users to carry out these operations employing an easy-to-understand Python interface.

Selenium: Interaction with web browsers is automated utilizing the selenium package of Python. It enables you to oversee and automate web browsers, engage with web content, and conduct out a number of operations like web scraping, web testing, and web application automation. Selenium is capable of being used for web scraping by programmatically loading websites, collecting HTML data, and storing it for subsequent processing or evaluation.

JSON: Notation for JavaScript objects. JSON is a compact standard for storing and transferring data. Data is transferred from the server to a page on the web using JSON.

gTTS: Your voice request for a query is transformed to text utilizing Google's text-to-speech engine. When you apply a look-up function to locate the response to an inquiry or command, gTTS translates the result into an audio file in MP3, WAV, or Ogg Vorbis formats. The Google Translate API is interfaced with via the contents of this package. It provides multilingual support and AI generated voice manipulation.

OS: Python's OS module enables users an opportunity for interacting into the operating system. It permits users to execute operations concerning the manipulation of files and directories, procedures, and variables.

Pyaudio: PortAudio is a C++ library that can play audio, alter audio information, and record sounds through a microphone or additional sound input devices. Speech synthesis, voice recognition, audio recording, and real-time audio manipulation are all widespread applications for PyAudio.

Smtplib: It is an integrated Python package that allows you to send electronic mail using the SMTP. One can connect via an SMTP server and deliver emails to a number of receivers using smtplib.

Open AI: It is a Python module that makes it trivial to connect to the OpenAI API from Py-based applications. It provides an already-defined collection of classes for API resources that dynamically initialise itself from API responses, making it functional with a broad spectrum of OpenAI API versions.

V. RESULT

By carrying out routine chores or basic human actions, the voice assistant lessens the workload of users and can eliminate any reliance on third parties to complete the same duties. Phases of the system will include: data or a question is entered during the input phase using text or voice, text to voice interpretation, Data is processed and stored, and the output voice from the polished text is sent to the Jinn console. The data generated at each stage can then be utilized to search for trends and analyze them in order to utilize them later. The conclusion drawn from a literature review and analysis of persisting systems is that our offered system will not only make it easier for users to interface with other systems and modules, but also help them stay more organized.

All generations of individuals, including those with special needs, will benefit from this project. Our personal voice assistant is simple to use and reduces the need for manual labor to complete a variety of tasks. Because the voice assistant system is modular, new features can be added without affecting existing system functionality.

Among the notable VPs are Alexa, Cortana, Siri, and Google Assistant. Voice recognition, contextual comprehension, and human contact are among the problems that IVAs have. Every IVA varies with time. While each of the four IVAs could answer roughly 17.35% of daily queries, Google Assistant was shown to be the most successful, answering 59.80% of them. Some were disappointed with Siri's response rate, which was only 45%, which was lower than expected. Conversely, Cortana was only able to respond to one-third of the questions. Alexa was clearly outperformed in this competition, answering only 7.91% of the time [16]. Jinn is projected to provide much higher accuracies over a large number of search queries.

Stability: The stability of a system is determined by the output it produces. If the output is limited to the specified input, the system is considered stable. Additionally, if the system operates on all of the poles, it is considered stable.

TABLE I: COMPARISON BETWEEN JINN AND OTHER PROJECTS

Paper Name	Functionalities	Limitations	Comparison
ARA- A Voice Assistant for Disabled Personalities [1]	Speech to text conversion, translation, access to any result required over the internet, date and time, web search and helping them to find a particular resource.	Performs only a limited number of tasks with specified voice commands. Limited to high-speed networks	Jinn provides a wide range of functionalities for all differently abled people. Most features have offline support.
Personal Desktop Voice Assistant - Research Paper [2]	wake up detection, make private conversations, search on the web, play videos and music, and special functions for blind.	Data security concerns, disconnected exchange, reliance on gadget manufacturers, high cost of voice activated apps and skill sets	Our jinn provide a wide range of functionalities for not only the visually impaired but also for students. it also has a more accurate comprehension of user's voice.
SARA: A Voice Assistant using Python [9]	Reading Newspapers, Weather reports, updating emails, playing music, setting alarm and searching on web	High cost of weather API, lot of errors faced while running application.	High threshold of non-paying API requests as well as smooth running of functionalities.
Research Paper on Desktop Voice Assistant [10]	Weather report, playing music, opening apps, managing emails, showing date and time	Similar with basic Prototype of voice assistants and lacks multidimensional function support	Jinn is a multi-utility voice assistant catering to a large variety of needs and requirements specific to students and visually impaired people.
The Voice Enabled Personal Assistant for Pc using Python [11]	Taking a note, YouTube Searches	It's only functionality is to search. It does not fully automate controls of YouTube.	Jinn can play, pause, mute, unmute, rewind, forward, increase/ decrease playback speed, and navigation of YouTube searches, etc. The Selenium module is used to provide a complete automated YouTube.
Voice Assistant Automation System for Applications [15]	Web Searches by WOLFRAMALPHA, opening of applications, YouTube searches	Increased cost rate as WOLF RAMALPHA is a paid API. Results of web search for the queries are also not very concise.	Use of googlescrap gives more precise information for any given query increasing Jinn's efficiency. Cost is not very high.

A. Advantages

Personal Desktop Voice Assistant offers a number of key benefits, including the ability to:

Time and Cost Saving: By automating tasks and providing quick responses to queries, our personal assistant can save users valuable time and reduce the cost spent on PVA resources.

Offline access: The main benefit of using Python Text to Speech module is that it can be used offline. Hence, it can perform tasks like opening a notepad, music player and other desktop applications which cannot be opened using a personal assistant like Google without internet connection.

Advanced features: Jinn has personalized features like extracting information from an uploaded pdf, Image captioning, etc. which are unique to it.

Hands-Free Operation: Personal assistants enable hands-free operation, which is especially valuable in situations where manual interaction is impractical or unsafe, such as while driving or cooking.

Always listening personality: The "always listening" feature of voice assistants refers to their ability to constantly monitor for a specific wake word or phrase even when the device is in a standby or idle mode.

B. Limitations

There are a few drawbacks to using Personal Desktop Voice Assistant, such as:

Data security: Despite the increasing popularity of voice assistants, many individuals remain concerned about the data collected by these devices. The manner in which the data is stored, interpreted, and ultimately discarded is to scrutinize. The "always listening" feature can raise privacy concerns as these devices may inadvertently record sensitive or private conversations.

Challenges in Interpretation by Voice Assistants: Voice assistants face challenges in accurately comprehending spoken language due to ambiguity, accents, and context, resulting in occasional misunderstandings. Misinterpreting homophones, emotions, and cultural nuances further complicates their performance.

VI. CONCLUSION

In the ever-evolving realm of technology, Jinn stands out as a model of innovation and inclusivity, creating a new benchmark for personal voice assistants. The impressive features of this digital companion, including speech recognition, image-to-text conversion, and seamless interaction with vital programs like Google Maps, WhatsApp, and YouTube, demonstrate its boundless potential. The most significant achievement of Jinn, however, is its consistent dedication to improving accessibility and productivity for students and people with disabilities. By giving the visually impaired tools to interact with written text, Jinn unlocks a world of knowledge and education that was previously concealed behind walls. Students can access a wealth of information and support from Jinn, who acts as an important study buddy. These aren't merely features; they're game-changing capabilities that might fundamentally alter how people with special needs use technology.

Additionally, Jinn's integration with OpenAI and its continuously growing knowledge base guarantee that it will always be a smart and flexible voice assistant prepared to fulfil the various needs of its customers. Jinn symbolizes a future in which technology is envisioned as a potent weapon for empowerment, levelling the playing field and improving the standard of living for people who might have particular difficulties on a daily basis.

VII. FUTURE ENHANCEMENTS

Gazing towards the future, it becomes clear that Jinn has limitless potential to develop and broaden its influence on people and communities all over the world. Jinn can be made more emotionally intelligent by integrating schematics, facial expressions and more complex visual data. Jinn can be a personalized study partner by providing personalized learning pathways, adaptive study advice with academic institutions to develop AI-driven, immersive learning experiences. Privacy protection and data security will continue to be of utmost importance as Jinn becomes a vital part of users' life. Enhancing user control over data, adopting strong encryption, and keeping ahead of possible cybersecurity risks should be the main goals of future growth. Integration of Jinn with smart cities and houses to act as a central control point for managing multiple IoT devices, giving simple access to surroundings.

REFERENCES

- [1] Sharma, Khushboo & Bahal, Disha & Sharma, Aman & Garg, Ankita & Verma, Neeta. (2022). REVIEW ON ARA- A VOICE ASSISTANT FOR DISABLED PERSONALITIES. International Journal of Engineering Applied Sciences and Technology. 7. 172-175. 10.33564/IJEAST.2022.v07i01.026.
- [2] Sakshi R Jain, Prof Feon Jason. Personal Desktop Voice Assistant - Research Paper. International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Impact Factor 7.918. Vol. 12, Issue 3, March 2023 DOI: 10.17148/IJARCCE.2023.12324
- [3] Rohit K. Roy, Megha D. Deshmukh, Rajat K. Biswas, & Milind V. Lande. (2023). An Overview of Voice Enabled Personal Assistant Using Python. International Journal of Innovations in Engineering and Science, 8(1). doi:10.46335/ijies.2023.8.1.3.
- [4] Rosso, Paolo & Hurtado Oliver, Lluís & Segarra, Encarna & Sanchis, Emilio. (2012). On the Voice Activated Question Answering. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 42. 75 - 85. 10.1109/TSMCC.2010.2089620.
- [5] Indranil Basu, Saumyadeep Bhattacharyya, Arpan Mondal, Babin Maitra, Roshni Joardar, Koushik Pal, & Koushik Pal. (2023). Python-Based Desktop Assistant with Voice Recognition. International Journal of Advanced Research in Science, Communication and Technology, 398-403. doi:10.48175/ijarsct-10618.
- [6] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, An Intelligent Chatbot using Natural Language Processing, International Journal of Engineering Research , Vol.6 , pp.281-286, 2017.
- [7] Masina F, Orso V, Pluchino P, Dainese G, Volpato S, Nelini C, Mapelli D, Spagnolli A, Gamberini L. Investigating the Accessibility of Voice Assistants With Impaired Users: Mixed Methods Study. J Med Internet Res. 2020 Sep 25;22(9):e18431. doi: 10.2196/18431. PMID: 32975525; PMCID: PMC7547392.
- [8] Tulshan, Amrita & Dhage, Sudhir. (2019). Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa: 4th International Symposium SIRS 2018, Bangalore, India, September 19-22, 2018, Revised Selected Papers. 10.1007/978-981-13-5758-9_17.
- [9] Ayush Chinchane , Aryan Bhushan , Ayush Helonde , Prof. Kiran Bidua, "SARA: A Voice Assistant Using Python", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538.
- [10] Vishal Kumar Dhanraj (076), Lokeshkriplani (403), Semal Mahajan (427), "Research Paper on 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.

- [11] 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, Volume-10 Issue-4, April 2021.
- [12] Rajat Sharma, Adweteeya Dwivedi. "'JARVIS" - AI Voice Assistant". <https://www.ijsr.net/archive/v11i5/SR22503183839.pdf>. (accessed August 29, 2023)
- [13] Pooja C. Goutam, Monika S. Jalpure, Akshata S. Gavade, Pranjali Chaudhary, Prof. A.V. Gundavade. "VOICE ASSISTANT USING PYTHON". <https://ijcrt.org/papers/IJCRT22A6338.pdf>. (accessed July 1, 2023).
- [14] S. Thakare, A. Kamble, V. Thengne and U. R. Kamble, "Document Segmentation and Language Translation Using Tesseract-OCR," 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), Rupnagar, India, 2018, pp. 148-151, doi: 10.1109/ICIINFS.2018.8721372.
- [15] Bhaskar Balasaheb Patil, Srinivas Joshi, Soumya R. Shivashimpi, Sagar, Dr. Sadhana P. Bangarashetti. "Basaveshwar Engineering College, Bagalkote, Karnataka. "Voice Assistant Automation System for Applications". ISSN: 2454-132X Impact Factor: 6.078 (Volume 8, Issue 4 - V8I4-1224) <https://www.ijariit.com>.
- [16] Ammari, Tawfiq & Kaye, Jofish & Tsai, Janice & Bentley, Frank. (2019). Music, Search, and IoT: How People (Really) Use Voice Assistants. ACM Transactions on Computer-Human Interaction. 26. 1-28. 10.1145/3311956.
- [17] Abdolrahmani Ali. Mukkath Rishin, Kuber Ravi, Branham Stacy M., "Blind People Are Power Users: An Argument for Centering Blind Users in Design of Voice Interfaces".