ijRASET

International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Literature Survey Paper on Natural Language Processing (NLP) in Voice Assistants

Aayush Neupane

UG (BE) Student, Department of CSE, Dr Ambedkar Institute of Technology, Bengaluru, India, 560056

Abstract: People have always wanted to talk to computers and eventually succeeded to fulfill their wish by developing the virtual assistants. Virtual assistants in computer are software that talk to the human beings, understand their language and perform the given tasks. In order for a virtual assistant to be able to communicate to human beings, they levy natural language processing. This survey paper looks upon the previous researches done on the same, the use of NLP in voice assistants, machine learning algorithms used, to enable voice activated commands, and the ability of voice assistants to learn and adapt to individual user preferences. The paper thoroughly discovers the capabilities, limitations, application, efficiency and productivity of the voice assistants, flowcharts, algorithms in the design, and use of data, collection and storage of data for user behavior prediction, and the limitations of voice assistants. It highlights the potential benefits and challenges associated with the integration of this technology into various domains as well.

All the figures and even keywords and sentences, paragraphs, findings and results have been extracted from various papers and ORIGINAL CREDITS have been given to the deserving authors.

Original Credits for the contents:

- *Voice Based Personal Assistant Prachi, Abhishek Kumar Singh, Mohd Akmam SCSE, Galgotias University, Greater Noida, India, 2021 – Paper 1*
- *Virtual Assistant using NLP Techniques, G Rushivardhan, Mrs K Santoshi, Department of Information Technology, GMR Institute of Technology, Rajam, India, October 2022 – Paper 2*
- *Personal Desktop Voice Assistant Sakshi R Jain, Prof Feon Jason Jain University, Bengaluru, March 2023 – Paper 3*
- *Fig1: https://techvidvan.com/tutorials/natural-language-processing-nlp*
- *Fig 2: https://studentsxstudents.com/all-about-virtual-voice-assistants-natural-language-processing-and-speech-recognition-ae04f854bc59*
- *Fig: https://www.researchgate.net/figure/Data-flow-diagram-of-speech-recognition_fig2_287429405*

*The intention of this literature survey is to provide the related and required information about Natural Language Processing in Virtual Assistants, proposed in 3 different papers, for the new researchers, and not intended for any commercial use and neither I do claim any authority over the original materials of the original authors. *

Keywords: Voice assistant, NLP (Natural Language Processing), Lexical Analysis, Morphological Analysis, Syntactic and Semantic Analysis, APIs, Machine Learning

I. INTRODUCTION

People have wanted to talk to computers almost from the moment the first computer was invented. Science fiction is full of computers that can hold a conversation, from HAL 9000 and the Star ship Enterprise's computer to Marvin the Paranoid Android and KITT the car. Just a few decades ago, the idea of holding meaningful conversation with a computer seemed futuristic, but the technology to make voice interfaces useful and widely available is already here. Several consumer-level products developed in the last few years have brought inexpensive voice assistants into everyday use, and more features and platforms are being added all the time. Users can do everything from asking simple informational questions to playing music and dialing their phone or turning lights on and off via voice control.

A. Natural Language Processing (NLP)

Natural language processing (NLP) is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyse large amounts of natural language data. The goal is a computer capable of "understanding" the contents of documents, including the contextual nuances of the language within them. The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves.

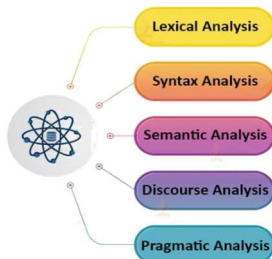International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

Fig 1: Working of Natural Language Processing.
(Original Credits: https://techvidvan.com/tutorials/natural-language-processing-nlp)

B. *Working Of Virtual Assistant*

Fig 2: Working of Virtual Assistants
(Original Credits: https://studentsxstudents.com/all-about-virtual-voice-assistants-natural-language-processing-and-speech-recognition-ae04f854bc59)
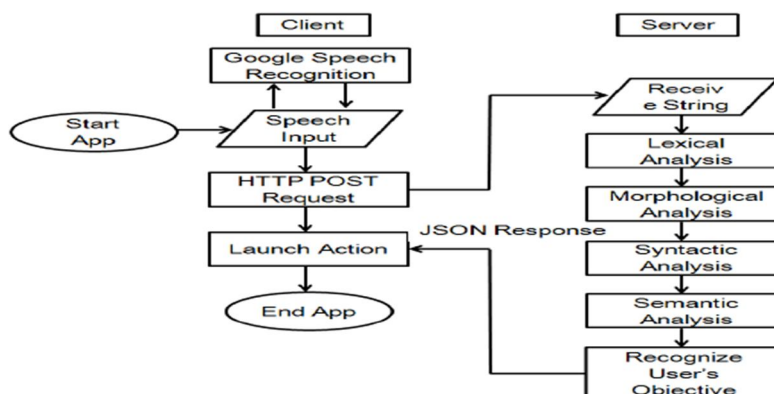

Fig 3: Flowchart of Working of Voice Assistant
(Original Credits: https://www.researchgate.net/figure/Data-flow-diagram-of-speech-recognition_fig2_287429405)

II. LITERATURE SURVEY

A. *Voice Based Personal Assistant Prachi, Abhishek Kumar Singh, Mohd Akmam SCSE, Galgotias University, Greater Noida, India, 2021*

This research paper is for solving real world problems with the help of advanced AI (Artificial Intelligence) and ML (Machine Learning) techniques. The voice based personal assistant technology attracts most of the people around the world and people do interaction in many ways like via mobile phone, laptop, computer, etc. This project is mainly focusing on broader parts of windows services or applications which can be automated for making system life easy. The sole purpose of this product is to give more comfort of the user with the help of Automated features of speech recognition.

The Most well-known utility of the iPhone is "SIRI" which enables a quick consumer to talk with the quick consumer cellular with speech and it additionally answers to the voice instructions of the consumer. Same form of function exists likewise via way of means of Google that is "Google Voice Assistant" that's used in Android Phones.

1) *Problem Solved*

a) *Voice Based Recognition and Related Understanding:* Virtual Assistant are going to be very often useful in our daily life because of its reach in smart-phone ir-respective of operating system whether its "SIRI"," GOOGLE ASSISTANT", and" CORTANA" many of us own at least once. There is main challenge in VBPA (voice based personal assistant) is that people's voices vary and they speak in different languages along with different accent which sometimes create issue to reorganize voice. In voice base recognition, all needed to know is that it was the way regularly the virtual assistants could perceive the words users were saying to found more. A survey has been done. In this survey users tried speech recognition from various gadgets and it also changing level of foundation noise found out that Google and Siri knew their users well according to the graphs. There was a couple of misconceptions, similar to when someone asked Siri that "Will I need an umbrella for next two months can you suggest?" And voice based personal assistant gave user the date for one week later, so it is clear that voice-based assistant not good at understanding. Google Voice based personal assistant is very good at understanding natural languages, Alexa is good in music related activities and their thing, But Alexa is not comfortable with basic yet simple questions on the other side Cortana is not that good in basic conversation.

b) *Human Free Collaboration:* The issue that our survey found out that for some users, hands free connection is the real utilization condition. We thought that in this specific situation, interrupt ability was a huge burden. They were puzzled when they were requested by the intelligent virtual assistants draw in outwardly with the screen, or to select alternatives by tapping the touch screen as opposed to through utilizing dialog. This appeared to interfere with the without hands involvement of IVAs and was considered especially complicated in circumstances, for example like when driving. We suggest that keeping up dissertation as the fundamental information and yield all through the connection should be a need in the future plans of IPAs with a specific end goal to guarantee that hand free association is supported completely and that events are not caught up by a cooperation methodology move. Hands free technology is the future where people can utilize their maximum time of their lives. Speaking of desktop and mobile applications, the ones whose interface is powered by AI and is processed by Chabot, they have the ability of providing perfect delivery and quick access of the required data, at the call of the users.
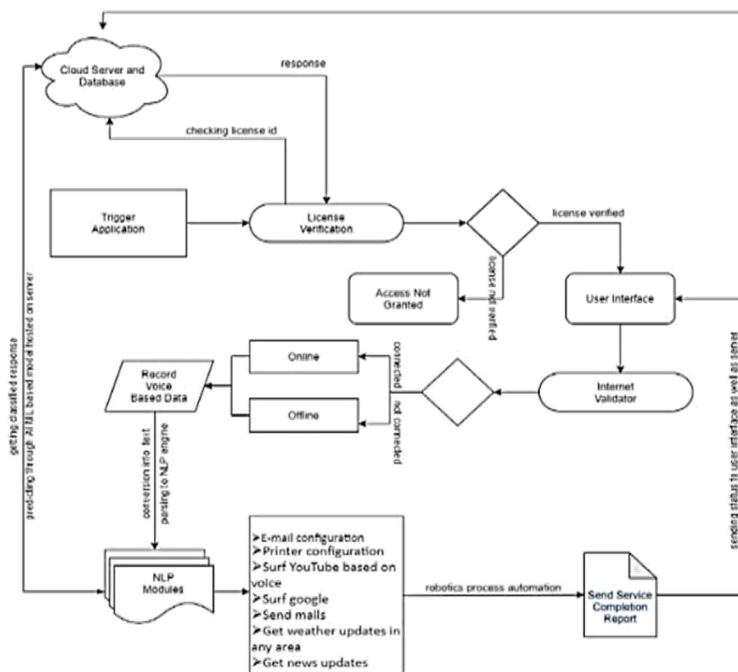
2) *Methodology*

Fig 4: Methodology of Voice Assistant

(Original Credits: Voice Based Personal Assistant Prachi, Abhishek Kumar Singh, Mohd Akmam SCSE, Galgotias IUniversity, Greater Noida, India, 2021)

a) *Cloud Server:* Cloud Server or Database helps in to check the license id of the user which tells the system if the user is a legit user or not. The input of User's speech will be converting into texts from the special domain which are specially organized on the pc network server at the knowledge center from the electro-acoustic transducer is briefly hold on within the system which is sent to cloud for recognizing the speech. The equivalent text is then received and sent to the central processor.

b) *API:* An API is a technology that allows applications to talk to each other. In other term, an API is like a messenger that delivers any request of the user to the provider who handles the user's request and then delivers the response back to the user.

c) *Backend:* The python backend script receives the output from the voice reputation module after which it identifies whether or not the command or the voice output is an API Call, Context Extraction of data, or it's just a System Call. The output is then dispatched lower back to the python backend module to represent the expected output to the user. Speech recognition is the process of converting the voice of a user into text. This is commonly used in voice based personal assistants like Google, Cortana, etc. Python provides an API which called Speech-Recognition to allow programmer to convert the user's voice into text for further processing.

d) *Internet Validator:* It is used to let the system know if the system is allowed to do online/offline work or both.

e) *NLP Modules:* NLP helps in speech recognition and converting into text. TTS conversion which talk about the skill of computers to read text properly or not. A Text to Voice Engine that can able to translates written text to a voice representation, then converts the voice representation to waveforms that is the output which we get as the audio. Text to Voice engines with the support of different languages and vocabularies that are available to programmer with the help of third-party publishers like APIs and Text to Voice python modules.

3) Challenges

a) The Types of Speech Utterance

Speech recognition are classified consistent with what sort of utterance they need ability to acknowledge. They are classified as:

- *Isolated Word:* They are some phrase recognizers commonly calls for every vocable to own quiet (loss of an audio signal) on every aspect of the pattern window. It accepts an unmarried phrase at a time.
- *Connected Word:* It is nearly like a remoted word; however, it permits separate utterances to "run-together" which includes a minimal pause in among them.
- *Continuous Speech:* It lets in the customers to talk obviously and in parallel the computer will decide the content.
- *Spontaneous Speech:* it is the speech that's herbal sounding and isn't always rehearsed.

b) Types of Speaker Model

The Speech recognition system is divided into the two main categories:

- *Speaker Dependent Models:* These systems are designed for a single selected speaker. They are easier to develop and more accurate but they are not flexible.
- *Speaker Independent Models:* These systems are designed for a diversity of speakers. These systems are difficult to develop and fewer accurate but they are flexible.

c) Types of Vocabulary

The vocabulary size of the speech recognition systems affects the processing requirements, accurateness and density of the system. In voice recognition system(dot): voice-to-text. The types of vocabularies can be classified as follows:

- Small vocabulary includes solo letters. Like: a, b, c, d and etc.
- Medium vocabulary means two or three letter words. Like: The, who, why, will
- Large vocabulary means more letter words. Like: voice, speech, machine, weather etc.

4) Outcomes of the Paper

DOT is a web application designed to ease our daily life work by saving out time by doing our work on Voice Commands and has the capability to understand what the user is saying without internet connection. DOT has various functionalities which makes it fun for the user like to open the YouTube video or to configure the mail or to doing fast google research etc. It only able to work on human voice Text-to-Speech mechanism convert it into commands and gives desired answers to the user on the basis of the user's query that is being asked to make a call or to perform any acts and operations.

DOT also gives the same output as other does and along with it gives greeting to the user who says like Hello or Hi so that user will feel more comfort and feels free to relate with the voice based personal assistant. This application is also able to reduce any kind of manual work which is required in the user's life which are hectic or feel bored. The DOT can easily do all the tasks. The entire system works on the user's voice input rather than the text which is a one-step advancing in the world of applications. Hence, DOT is the application which fully operates on the Voice Command of the user.

(Original Credits: Voice Based Personal Assistant Prachi, Abhishek Kumar Singh, Mohd Akmam SCSE, Galgotias University, Greater Noida, India, 2021)

B. Virtual Assistant using NLP Techniques, G Rushivardhan, Mrs K Santoshi, Department of Information Technology, GMR Institute of Technology, Rajam, India, October 2022

A Virtual Assistant is software that can have Natural Language Conversations with people. The modelling of dialogue is one of the key tasks in Artificial Intelligence, Voice Recognition and Natural Language Processing. Making a good Virtual Assistant has been the most difficult challenge since the advent of Artificial Intelligence.

Although Voice Assistants are capable of a variety of activities, their main responsibility is to recognize human speech and react properly.

There are still some issues with developing data-driven systems despite the fact that there are now many Voice Assistant platforms available because a substantial amount of data is needed for their creation. Consequently, implementing these Virtual Assistants with Python libraries (like NLTK, SpaCy, Polyglot, Text Blob, Flair) may be accomplished. Moreover, to provide a better platform, Web Connectivity is also can be done to evaluate the Voice Assistant on a Web-based platform which will help in analyzing Human Voice Assistant Interactions can be used with voice commands only.

Due to which Voice Assistants will learn how to interact with humans. In this paper the making and the working of voice assistant is summarized, its and limitations are also given.

1) Problem Solved

Voice Assistant keeps learning the sequence of questions asked to it related to its context which it remembers for the future. So, when the same context is mentioned, it starts a conversation with you asking the relevant questions.

It performs the Arithmetic Calculations based on voice commands and giving back the computer solution through the voice. Search Internet based on user voice input and giving back the reply through a Voice Assistance. Results produced were 98 percent accurate to the input. The voice recognition was made more accurate.

2) Methodology

a) Speech Synthesizer Algorithm

The steps involved in this Algorithm:

- NLP Core Engine processes the input given by the user so that it can be sent to the speech synthesizer to respond.
- Microsoft Speech Synthesizer is used which consists of STT. STT is speech to text, sounds or voice received from user is converted to Text for processing of information by bot. Then it responds according to input given by the user.

b) Data Flow Sequence Algorithm

The steps involved in Data Flow Sequence:

- Initialize Device
- Task Manage
- Service Manager: Analyses the commands and matches them with servers
- Execute Command: When the matching found for commands, run the py script and gives   response.
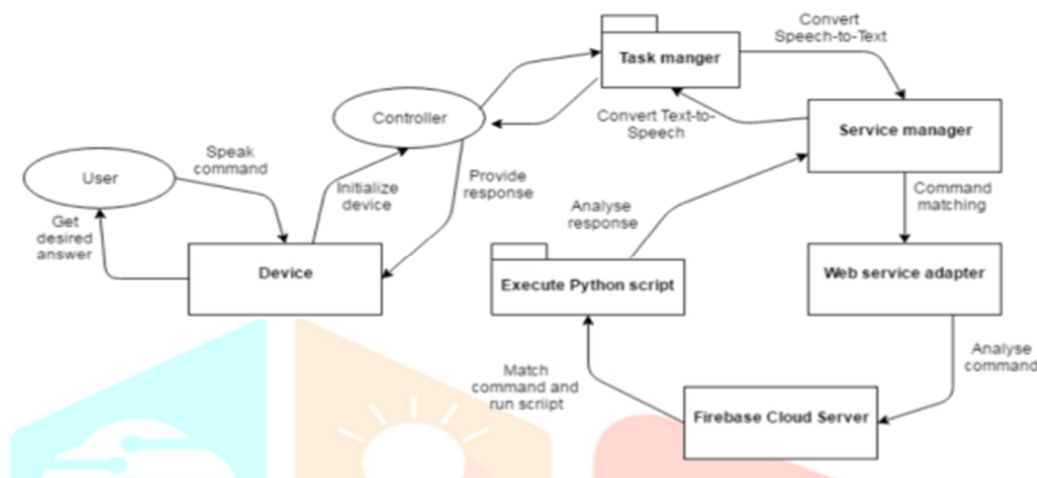
Fig 6: Data Flow Sequence

(Original Credits: Virtual Assistant using NLP Techniques, G Rushivardhan, Mrs K Santoshi, Department of Information Technology, GMR Institute of Technology, Rajam, India, October 2022)

c) Porter Stemming Algorithm

It is the process for removing the suffixes from words in English. Removing suffixes automatically is an operation which is especially useful in field of information retrieval.

- Gets rid of plural and -ed or -ing suffixes
- Turns terminal y to i
- Maps double suffixes to singles ones: -ization , -ational etc…
- Deals with suffixes -full,-ness etc.

3) Challenges

Less accuracy in recognizing voice in noisy environment.

4) Outcomes

S.no	Model	Accuracy %
1	Speech Synthesizer	97.3
2	Data Flow Sequence	98
3	Core and interface Accessing	93.2
4	Porter Stemming	79.6

Fig 7: Table with results accuracy

(Original Credits: Virtual Assistant using NLP Techniques, G Rushivardhan, Mrs K Santoshi, Department of Information Technology, GMR Institute of Technology, Rajam, India, October 2022)

C. Personal Desktop Voice Assistant Sakshi R Jain, Prof Feon Jason Jain University, Bengaluru, March 2023

The term "virtual assistant" refers to a software agent that can carry out tasks or provide services on behalf of a person. Virtual assistants that may be accessed through online chat are occasionally referred to as "chatbots" in general or in relation to just those. The term "virtual assistant" (VA) refers to computer-simulated settings that can mimic physical presence in both the actual world and made-up universes. In order to construct an intelligent Virtual Personal Assistant (VPA), new technology could be used in a number of ways, with a focus on user-based data. The goal of this project is to provide technical information about virtual assistant technology, including its advantages and disadvantages in many contexts.

The project focuses on virtual assistant types and structural elements of a virtual assistant system. This research paper explores the development and application of Personal Desktop Voice Assistants in various domains. The study focuses on the use of natural language processing and machine learning algorithms to enable voice-activated commands, and their ability to learn and adapt to individual user preferences. The paper reviews the current state-of-the-art in Personal Desktop Voice Assistants, including their capabilities, limitations, and potential applications. It examines the impact of these technologies on productivity, efficiency, and accessibility for individuals with disabilities. The research also considers the ethical and privacy implications of Personal Desktop Voice Assistants, including data collection, storage, and usage. It explores the need for transparency, consent, and accountability in the development and deployment of these technologies. The paper presents a case study on the integration of Personal Desktop Voice Assistants in healthcare, highlighting their potential to improve patient outcomes, reduce healthcare costs, and enhance patient satisfaction. Overall, this research paper provides a comprehensive overview of Personal Desktop Voice Assistants, their current state of-the-art, and future directions. It highlights the potential benefits and challenges associated with the integration of this technology into various domains, and the need for responsible and ethical development and deployment.

1) Problem Solved

Built a virtual voice assistant that will enable users to interact with emerging technologies, manage their devices, and utilize technology for learning. It serves as a voice assistant for visually impaired people and is a cutting-edge system. By utilizing distinct custom layouts and speech to text, this solution improves system quality while enabling visually challenged users to access the desktop's most crucial functionalities. The user's speech will be the basis for all actions taken by the system. The system assists the user based on voice note, meaning that it follows instructions provided by the user. Because the user cannot see the action going place on the desktop, the system speaks out if the user needs to receive a response.

a) The blind applicant will also sense independence. • Because the system is a machine, it will execute without error.

b) Your smartphone will be controlled solely by voice commands, and the assistant will recognize the situation and respond to the user appropriately.

 Although many seniors are unable to utilize desktop computers, they can still benefit from this.

These assistive technologies will enable users who are blind or visually handicapped to learn from, compete with, and interact with their sighted counterparts.

2) Methodology

A personal desktop voice assistant is a software application that is designed to understand and respond to voice commands provided by the user. The following is a high-level system design description for a personal desktop voice assistant. • User Interface: The user interface of the personal desktop voice assistant should be intuitive and easy to use. The user should be able to interact with the voice assistant through natural language commands.

a) Speech Recognition: The system should have a robust speech recognition module that can accurately convert the user's voice commands into text. The speech recognition module should also be able to distinguish between different users and adapt to their speech patterns.

b) Natural Language Processing (NLP): The system should have an NLP module that can interpret the user's text commands and extract the relevant information. The NLP module should also be able to identify the user's intent and provide appropriate responses.

c) Knowledge Base: The system should have a knowledge base that contains information on a wide range of topics. The knowledge base should be regularly updated to ensure that the voice assistant can provide accurate and up-to-date information.

d) Machine Learning: The system should use machine learning algorithms to continuously improve its performance. The machine learning algorithms can be used to improve speech recognition accuracy, NLP performance, and user interaction.

e) APIs: The system should be able to integrate with other applications and services through APIs. This will enable the voice assistant to provide more comprehensive responses to user requests.

f) Privacy and Security: The system should be designed to protect user privacy and ensure that user data is secure. The system should only collect data that is necessary to provide the voice assistant's services, and user data should be encrypted and stored securely.

g) User Personalization: The system should be able to personalize the user experience based on the user's preferences and previous interactions. The system should also be able to learn from user feedback and adapt to their preferences over time.

h) *Action Fulfilment:* Once the intent of the user's request is identified, the system will execute the necessary action. For example, if the user requested to play a song, the system will find the song and play it through the desktop speakers.

i) *Response Generation:* Finally, the system will generate a response to confirm that the requested action has been completed. The response could be as simple as a confirmation message, or it could be more detailed, providing additional information related to the user's request.

Overall, a personal desktop voice assistant should be designed to provide a seamless and natural interaction between the user and the system. The system should be reliable, accurate, and secure, and should continuously learn and improve to provide better services to the user.
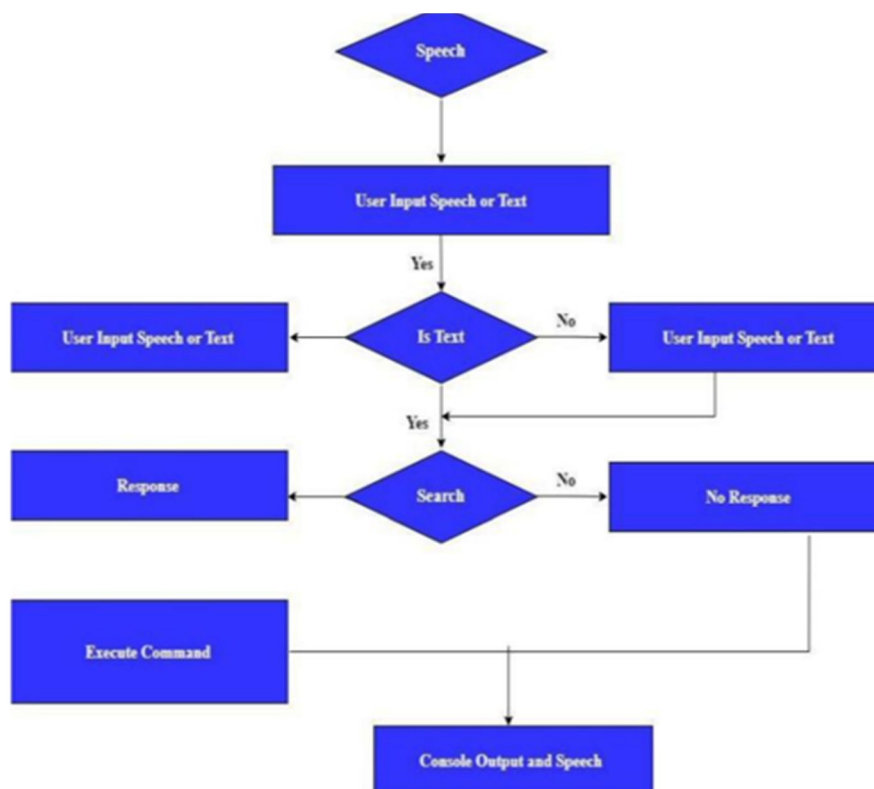

Fig 7: Flowchart of Desktop Assistant

(Original Credits: Personal Desktop Voice Assistant Sakshi R Jain, Prof Feon Jason Jain University, Bengaluru, March 2023)

3) *Challenges*

a) *Concerns About Data Security:* Despite the fact that people are using voice assistants more frequently, there is still a lot of worry about the information these devices collect and the businesses that create the apps that run on them. Customers are concerned about the data's storage methods, viewers, and eventual disposition. If marketers don't handle these data and privacy issues, they won't be able to access these prospects or their data.

b) *Disconnected Exchange:* Another drawback is that compared to other platforms, voice assistants as a channel offer fewer enriching interactions. The choices include visual interactions versus speech material alone, which generally means recycling current content. This might make some of the more significant interactions that marketers can have elsewhere less effective.

c) *Reliance on Gadget Manufacturers:* You are at the mercy of the manufacturers of devices, such as the corporations that build wearable tech, cars, and appliances, as a marketer. Before getting started, you should carefully investigate the device manufacturers you wish to collaborate with to achieve long-term success.

d) *Investing in voice-activated Apps and Skill Sets:* The cost of creating the voice app for this channel can be high. Building an internal skill set tailored towards the intricacies of voice assistants may take a lot of time if you participate in this channel. Consequently, it's crucial to weigh the advantages and disadvantages of using voice assistant channels.

4) Outcomes

When the user speaks the wake word, the assistant wakes up and starts listening for the user's request. The user then speaks their request, which is captured by the assistant's microphone and converted to text using speech recognition technology. The assistant analyses the user's request to understand the intent behind it and maps it to a specific action or set of actions. The assistant then executes the mapped action(s) and generates a response to the user's request like setting reminders or appointments, sending emails, controlling home automation devices, playing music, performing web searches, and checking the weather or news updates. Finally, the assistant speaks the response to the user.

(Original Credits: Personal Desktop Voice Assistant Sakshi R Jain, Prof Feon Jason Jain University, Bengaluru, March 2023)

III. CONCLUSION

From the analysis conducted in these papers, different algorithms are used. Comparative study performed among the various techniques like speech synthesizer, data flow sequence, core and interface accessing, porter stemming. The Data Flow Sequence Algorithm has most accuracy for providing the required output as required for the user.

The complexity and accuracy of voice recognition technology and voice assistant software have grown exponentially in the last few years. Currently available voice assistant products from Apple, Amazon, Google, and 86 M. B. HOY Microsoft allow users to ask questions and issue commands to computers in natural language. There are many possible future uses of this technology, from home automation to translation to companionship and support for the elderly. However, there are also several problems with the currently available voice assistant products. Privacy and security controls will need to be improved before voice assistants should be used for anything that requires confidentiality. Librarians should monitor these products and be ready to provide assistance to their patrons with these devices. They should also explore the possibilities for providing library materials via voice assistants as the technology matures. Selecting a perfect gift on the birthdays or wedding anniversaries of our beloved ones is too difficult today. Guess what? Your virtual personal assistant can do that for you. Even one of the hottest gift ideas in this season is virtual assistants with a voice-enabled user interface. These assistants are very handy devices, and they are becoming increasingly popular in our daily lives. But what empowers these virtual personal assistants. There has been a lot of research going around as to what tech can be employed to improve virtual assistance. It turns out that IoT app development extends an enormous number of possibilities to realize the true digitization in personal assistance.

REFERENCES

[1] Voice Based Personal Assistant Prachi, Abhishek Kumar Singh, Mohd Akmam SCSE, Galgotias University, Greater Noida, India, 2021
[2] Virtual Assistant using NLP Techniques, G Rushivardhan, Mrs K Santoshi, Department of Information Technology, GMR Institute of Technology, Rajam, India, October 2022
[3] Personal Desktop Voice Assistant Sakshi R Jain, Prof Feon Jason Jain University, Bengaluru, March 2023

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊘ (24*7 Support on Whatsapp)