Survey on Personalized Voice Assistant

Dr. S Srikrishnan[1]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
dean.cse.adishankara.ac.in

Ashwin Pushparaj[2]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
ashwinpushparaj2002@gmail.com

Advaith Suresh[3]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
advaithsuresh07@gmail.com

Akash Krishnan Nair[4]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
akashkuzhiyathu@gmail.com

Teena George[5]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
teena.cs@adishankara.ac.in

Sharika T R[6]
Department of Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology Kalady , Kerala
sharika.cs@adishankara.ac.in

*Abstract -***This survey reviews the current state of personalized voice assistant systems that utilize natural language processing, speech recognition, and AI techniques. It examines methodologies for developing intelligent assistants that can securely access local data, learn preferences, and control IoT devices via voice. The paper analyzes the capabilities, advantages, and limitations of existing voice assistant solutions. It also explores their integration with web technologies and discusses challenges like privacy, speech accuracy, and language modeling constraints. The goal is to provide an overview of personalized voice assistant technologies, their potential impact, and future research directions toward enabling seamless voice-based interactions.**

Keywords – Personalized voice assistant, NLP, Speech Recognition

I. INTRODUCTION

Voice assistants are AI-powered software programs that understand natural language, recognize speech, and respond via synthesized speech. They allow hands-free, voice-controlled access to information and services by interpreting voice commands. Intelligent personal assistants are software that can help users complete simple tasks by having conversations with them. Smart assistants are physical devices with intelligent speakers that listen for a wake word before assisting with certain activities, like Amazon Echo, Google Home, and Apple HomePod. Virtual digital assistants are computer programs that understand written and spoken natural language input to help users. Chatbots mimic human conversation. Many companies use virtual assistants and chatbots for customer service to answer simple questions and connect to a human if needed. Voice assistants are becoming more important in mobile apps because they provide highly engaging user experiences, more so than other interfaces. They are used in various sectors like healthcare, education, entertainment, and automotive for hands-free control.

II. RELATED WORKS

Paper [1] discusses the development of a voice assistant called 'Sara' using Python. It uses speech recognition and natural language processing to take voice commands and perform tasks. Key features include playing music, web searches, opening applications, Wikipedia searches etc. Python libraries like SpeechRecognition, PyAudio, gTTS etc are used. The system architecture consists of speech recognition to convert audio to text, Python backend to analyze text and call appropriate functions, API calls, content extraction etc. Compared to other existing voice assistants, Sara currently does not have features like weather updates, taking voice input without activation, or language translation.

In paper [2], the proposed virtual assistant uses supervised learning, unsupervised learning and reinforcement learning techniques. It can be trained using deep learning and machine learning. The assistant is customized for Windows OS and can open programs, search Wikipedia, send emails etc. Important Python libraries used include SpeechRecognition, pyttsx3, Wikipedia, webbrowser, PyAudio etc.

[3] is a personal voice assistant that takes the user commands as input and performs tasks based on the user commands.It provides more natural and efficient interaction with support for multiple voice commands in the same utterance. This assistant has a unique face recognition technique through which only the authorized user can provide the command to the assistant and can perform their various tasks on the system. The proposed system reaches out to help our society by making their work easier as this system can tell the news, search what you want, send email by only your voice command, play game with you, set reminders, tell the location, forecast weather, can tell horoscope of you and endless number of tasks can be done by this .Since the system uses facial recognition to authenticate users, there are privacy implications regarding the storing and use of biometric data.

The paper [4] Proposes a home automation system using Internet of Things (IoT), natural language processing (NLP), and machine learning.Allows control of home appliances like lights and fans via two input modes - chatbot text interface and voice assistant.Input is processed by Raspberry Pi to determine user intent using NLP techniques like tokenization and lemmatization.Based on extracted intent, Raspberry Pi controls attached appliances via relays.Chatbot provides text input via Telegram API and custom keyboard. Voice assistant captures speech input.Output responses are given to user via text or speech.Overall allows flexible input modes for remote control of appliances.

The paper[5] discusses developing a personal AI desktop assistant using Python that can understand voice commands and perform basic tasks.The assistant uses speech recognition to convert audio commands to text and identify keywords to determine appropriate actions.It incorporates modules like Pyttsx3 for text-to-speech conversion, Speech Recognition for audio transcription, web browser control, and OS commands.Proposed features include playing music, accessing email/texts, Wikipedia searches, launching apps, voice greetings based on time.The architecture comprises speech input, audio-to-text conversion, command comparison, and voice/action output.

[6] is a voice-controlled personal assistant for computers built using Python. It can understand voice commands and complete tasks like searches, news reading, translation, weather updates, and more.Speech recognition and natural language processing are used to analyze audio commands and trigger different functions based on keywords. APIs and libraries provide capabilities like WolframAlpha computation.JARVIS can perform tasks like Wikipedia searches, news reading, weather updates, translations, math solving, taking photos, and emptying the recycle bin through voice commands.The goal was to create a hands-free assistant to help various users like busy multitaskers, visually impaired, or motor impaired.

The paper[7] presents the development of a voice assistant called "Poodle" using Raspberry Pi.
It uses speech recognition to take audio input commands and text-to-speech to provide audio output.The system architecture includes wake words, microphone for input, command processing, and speaker output.Components used are Raspberry Pi, MicroSD card, power supply, USB hub, microphone, and speaker.It supports automating tasks like controlling lights as well as features like music playback, launching apps etc.Blynk app integration allows controlling appliances through the assistant. Alexa integration also provided.The system aims to provide a low cost, customizable, and extensible voice assistant platform.Benefits highlighted include hands-free control, time-saving, remote access to systems, convenience.Future work proposed includes improving natural language processing and adding more smart features.

Article [8] A recent innovation in mobile technology is a voice assistant application tailored for Serbian language speakers. This app enables voice control capabilities on mobile phones through integration with an open-source speech recognition system. To develop a robust Serbian acoustic model, the creators leveraged a pronunciation database of around 70,000 samples. They also augmented the data by artificially adding varying levels of noise, which improved the model's accuracy in noisy real-world conditions. Extensive testing was conducted, using a test set of approximately 4,500 pronunciation samples and a 14,000 word vocabulary. This voice assistant represents the first application designed specifically for voice interaction in Serbian. By supporting natural verbal communication between user and device, it provides Serbian speakers with hands-free convenience and accessibility. The developers' techniques of training the acoustic model on noisy data significantly enhanced its performance

In [9] Firefox Voice is an open source, customizable voice assistant implemented as a browser extension for the Firefox web browser on desktops. It is built using web technologies like JavaScript and React.js, with the interface integrated into the browser. Users can activate Firefox Voice through a hands-free wake word, keyboard shortcut, or toolbar icon. It then displays a tooltip over the current tab, listens to requests, and sends audio to an online speech recognition engine. The transcription is parsed locally in the extension to determine intent. Firefox Voice performs relevant actions and can respond by voice when needed. Being part of the browser allows it to leverage web technologies for flexibility and extensibility. The open source nature also allows community contributions to improve functionality.

According to a paper[10] , a personal assistant system with voice recognition intelligence is proposed that can operate with or without an internet connection. The system can receive inputs from users in the form of voice or text, process them, and return results in various forms such as performing

actions or presenting search results to the end user. Furthermore, this proposed system can change the way of interaction between end users and mobile devices. The system is designed so that the end user can access all the services provided by the mobile device using the user's voice command.

The article [11] discusses the development of an intelligent voice assistant system that can interact with people in their local language to help simplify their daily lives. The system is designed with rural life and its basic needs in mind.This article also tends to explore new possibilities.The system is recommended to connect to the Internet using any WiFi network.It is always on and listens for the wake word to wake the device.Once it hears this, the device collects the wake word voice commands and sends a signal to a cloud-based natural speech recognition service called The Voice Service, which interprets them and return the appropriate response.The system may include a set of microphones that can pick up your voice from across the room, even above

music and other environmental noises**.**

According to article[12], A voice-controlled intelligent personal assistant is implemented using a Raspberry Pi. It uses speech-to-text to convert voice commands to text. Natural language processing (NLP) interprets the text to understand the user's intent. Based on the derived meaning, text-to-speech conversion produces a spoken response. The device can provide information like weather, time, or access music applications in response to vocal requests.

The Raspberry Pi enables building a custom voice assistant cost-effectively. Speech recognition and NLP allow natural conversational interactions. By leveraging text-to-speech, the system can respond in a human-like voice. The capabilities can be customized by enhancing the NLP model and integrating different information sources and services. An open source stack enables community collaboration to expand functionality. Overall, the device demonstrates how modern AI techniques can create an intelligent voice-controlled personal assistant.

Table I. Comparison of Voice Assistant related papers

Ref No	Year	Methodology	Features	Disadvantage
[1]	2022	It uses speech recognition and natural language processing to take voice commands and perform tasks. Python libraries like speech recognition ,PY audio, GTTS etc are used.	Key features include playing music, web searches, opening applications, Wikipedia searches etc.	Limited conversational ability and contextual awareness - the system has restricted domains and may not handle complex queries well.
[2]	2023	The proposed virtual assistant uses supervised, unsupervised, and reinforcement learning techniques.	It's customized for Windows OS, can open programs, search Wikipedia, send emails, etc.	The assistant's functionality is restricted to application data. It has limited contextual awareness and struggles with complex queries.
[3]	2020	This system uses a unique face recognition technique to ensure that only authorized users can access the assistant and perform various tasks on the system.	The system can perform a wide range of tasks such as telling the news, searching the web, sending emails, playing games, setting reminders, and more.	The paper does not discuss data protection, the assistant's capabilities are basic, and hardware requirements are not provided..

[4]	2020	The paper proposes a home automation system using IoT, NLP, and machine learning.	It controls appliances via a chatbot text interface and voice assistant, with input processed by Raspberry	The system is limited to controlling a few appliances, has performance limitations due to Raspberry Pi hardware, may struggle with speech recognition in noisy environments, has potential security vulnerabilities, and requires technical expertise for setup.
[5]	2023	It incorporates modules like Pyttsx 3 for text-to-speech conversion, Speech Recognition for audio transcription, web browser control, and OS commands.	Proposed features include playing music, accessing email/texts, Wikipedia searches, launching apps, voice greetings based on time.	The system has limited features like basic tasks but lacks support for complex ones. Customization details are unclear. Voice recognition accuracy and privacy/security concerns exist.
[6]	2021	JARVIS is a voice-controlled assistant for computers, built using Python.	It performs tasks like searches, news reading, and weather updates through voice commands.	Closely linked to only google platforms so limited to google web searches
[7]	2021	Poodle is a Raspberry Pi-based voice assistant that uses speech recognition and text-to-speech for interaction.	It automates tasks, controls appliances, and offers extensibility.	Limited capabilities, Hardware constraints ,Speech recognition issues, Security concerns and Difficulty handling errors.
[8]	2015	A new mobile application for Android phones allows Serbian speakers to control their devices by voice. This Serbian language voice assistant app provides hands-free operation and verbal communication capabilities through integration with speech recognition software.	As the first voice assistant app created specifically for Serbian, this innovation enables more seamless verbal interaction for native speakers. Core functionalities include voice command capabilities, understanding natural language requests, and a large vocabulary of recognizable Serbian terms. The speech recognition integrates robust noise cancellation, allowing accurate operation even in noisy environments.	Limited vocabulary - The test vocabulary only contained around 14,000 words. A more expansive vocabulary would likely be needed for robust real-world use. Speaker dependence - The acoustic models were trained on a specific set of speakers. Performance could degrade for new speakers not in training data
[9]	2021	The voice assistant application was developed using modern web technologies. Its codebase is entirely written in JavaScript, taking advantage of its cross-platform capabilities	This voice assistant integration enables voice commands and hands-free web browsing within the Firefox desktop application. By implementing it as a browser extension, this allowed seamless integration with the existing Firefox codebase	constraints around vocabulary, noise robustness, language modeling, diversity of test data, flexibility of conversational capability, and technical practicality
[10]	2017	This paper proposes a voice-controlled personal assistant system for mobile devices that can work with or without internet connectivity.	It takes voice or text input and performs actions or provides search results through voice output.	The system has limited capabilities, relying on specific commands rather than natural language understanding. It does not provide advanced search or conversational abilities.
[11]	2020	This paper proposes an intelligent voice assistant system for rural areas that interacts in the local language to simplify daily life.	It connects via WiFi and is always listening for a wake word to activate and interpret voice commands using cloud services.	Relies on Amazon Alexa skills, limiting flexibility. Testing seems limited based on few example responses. Deployment and maintenance in rural areas may present unexplored challenges.

[12]	2019	An intelligent voice-controlled personal assistant is implemented using a Raspberry Pi as the main hardware platform. The device leverages speech recognition and natural language processing to interpret voice commands and determine user intent.	Based on the derived meaning, it provides relevant information or performs actions through synthesized speech responses. The Raspberry Pi provides a flexible and cost-effective way to prototype a custom voice assistant with modern AI capabilities.	Minimal conversational ability - The system appears to only respond to specific commands rather than having natural conversational interactions. Limited vocabulary - No details provided on the speech recognition vocabulary size and capabilities. May not understand diverse ways users could phrase requests.

III. CONCLUSION

Based on the literature survey conducted, it is evident that voice assistants have become an integral part of our daily lives. They have revolutionized the way we interact with technology and have made our lives easier. However, it is also important to note that there are some disadvantages associated with voice assistants. These include privacy concerns, security issues, and the potential for voice assistants to be hacked.

Despite these disadvantages, voice assistants have become more intelligent, more intuitive, and more user-friendly. They have made our lives easier and more convenient, and they have opened up new possibilities for human-machine interaction. As the technology continues to evolve, we can expect to see even more exciting developments in the future. In conclusion, voice assistants have come a long way since their inception. They have become more intelligent, more intuitive, and more user-friendly. They have made our lives easier and more convenient, and they have opened up new possibilities for human-machine interaction.

IV. REFERENCES

[1] SARA, an Assistant Using Python. International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653;ICValue:5.98;SJImpactFactor:7.5383DOI:https://doi.org/10.22214/ijraset.2023.4951 9

[2] Mrs. Patil Kavita Manojkumar, Aditi Patil, Sakshi Shinde, Shaktiprasad Patra, Saloni Patil "AI-Based Virtual Assistant Using Python: A Systematic Review" Department of Information Technology, Zeal College of Engineering and Research, Pune-41, Maharashtra, India
DOI :https://doi.org/10.22214/ijraset.2023.49519

[3] BRAIN – THE A.I. (Personal Voice Assistant) Mr. Yash Agarwal, Mr. Ranjeet Rai, Mr. Punit Kumar Chaubey

[4] Bhavyasri Kadali, Neha Prasad, Pranaya Kudav, and Manoj Deshpande, "Home Automation Using Chatbot and Voice Assistant" Department of Computer Engineering A.C. Patil Collegeof Engineering. DOI:https://doi.org/10.1051/itmconf/20203201002

[5] Rabin Joshi, Supriyo Kar, Abenezer Wondimu Bamud and Mahesh T R, (2023). Personal A.I. Desktop Assistant,2(2), 54-60. ISSN: 2583 5343 DOI: 10.59461/ijitra.v2i2.58

[6] Jash Vora, Deepak Yadav, Ronak Jain & Jaya Gupta "JARVIS: A PC Voice Assistant"Department of Computer Engineering, A. P. Shah Institute of Technology, G.B.Road, Kasarvadavli, Thane (W), Mumbai-400615, University of Mumbai.

[7] Dhrati Gupta , Abhishek Dixit , Akshat Mitra , Mr. Anuj Kumar "POODLE: Voice Assistant Using Raspberry Pi"UG Final YearStudents, Department of Electronics & Communication Engineering, RKGIT College, U.P., India 4Assistant Professor, Department of Electronics & Communication Engineering, RKGIT College, U.P., India

[8] Branislav Popović,Edvin Pakoci, Nikša Jakovljević, Goran Kočiš, Darko Pekar "Voice Assistant Application for Serbian language"

[9] Julia Cambre Alex C.Williams AfsanehRazi Ian Bicking "Firefox Voice: An Open and Extensible Voice Assistant Built Upon the Web"

[10] Dr. Kshama V.Kulhalli Vice-Principal & HOD IT D.Y. Patil College of Engineering and Technology, "Personal Assistant with Voice Recognition Intelligence" Kolhapur-416006 Dr.Kotrappa Sirbi Department of ComputerScience and engineering, KLE's Dr. M.S.Sheshagiri College of Engineering & Technology, Belgaum, India Mr. Abhijit J. Patankar Research Scholar, Computer Science and Engineering, VTU, Belgaum, Karnataka. International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017)

[11] Akshay Pramod Patil Suyash SubhashSamant Mohit Raju Ramtekkar "Intelligent Voice Assistant"

[12]Pooja Singh , Pinki Nayak , Arpita Datta , Depanshu Sani , Garima Raghav , Rahul Tejpal "Voice Control Device using Raspberry Pi"