Speech Recognition Based Infobot System Using Raspberry Pi

Meghana M N[1], Kushal Gowda M B[2], Lekhana Y B[3], Nagaraj P D[4], Prakash Stevagh G[5]
[1]*Assistant Professor, Department of ECE, MIT, Thandavapura*
[2,3,4,5]*Department of ECE, MIT, Thandavapura*

***Abstract*—The project aims to create an innovative interactive robotic assistant tailored for the Electronics and Communication branch. It will offer detailed profiles of faculty members and insights into the student body. The robot's distinctive feature is its interactive functionality, which warmly welcomes and engages users with friendly hand-waving gestures. This enhances accessibility to information and fosters greater engagement within the academic community. By integrating robotics with information dissemination, the project seeks to revolutionize how information is accessed and shared in academia, promising elevated accessibility, efficiency, and user experience.**

***Keywords*— Infobot, Speech Recognition, APIs**

I. INTRODUCTION

Introducing our pioneering project: an advanced robotic assistant designed specifically for the Electronics and Communication branch. By combining cutting-edge robotics with comprehensive information dissemination, we aim to simplify access to faculty and student profiles while creating a more engaging academic atmosphere.

This innovative creation welcomes users with a friendly handwave, initiating an inter active and inviting experience. Through its seamless integration of technology, our robotic assistant pledges to redefine how information is accessed and shared within our academic community.

As we embark on this transformative journey, utilizing robotics to enhance accessibility and connectivity within the Electronics and Communication domain, let's recognize the potential of our project to revolutionize communication and collaboration. Together, we're shaping a brighter future for all stakeholders involved.

II. RELATED WORK

In[1],The proposed work, titled "Implementation of Google Assistant on Raspberry Pi," was conducted at Politehnica University of Timisoara in Timisoara, Romania, in 2018. Authored by Septimiu Mischie, Liliana Mâ iu-Iovan, and Gabriel G p resc, the project focuses on developing a spoken dialogue system or voice-controlled assistant. These systems are designed to respond to various voices, regardless of accent, and can execute multiple commands while providing answers to simulate natural conversations. Such a system comprises several key components, including automatic speech recognition (ASR), spoken language understanding (SLU), dialogue management (DM), knowledge database (KDB), natural language generation (NLG), and text-to-speech synthesis (TTS) systems. The objective is to integrate these components into a cohesive system that runs on Raspberry Pi, enhancing accessibility and usability for users. This endeavor aims to explore the potential of voice-controlled assistants in facilitating natural and intuitive interactions with technology, ultimately advancing the field of human-computer interaction.

In[2],At GLA University in Mathura, Chaumuhan, Uttar Pradesh, India, Piyush Vashistha, Juginder Pal Singh, Pranav Jain, and Jitendra Kumar authored a project titled "Raspberry Pi based voice-operated personal assistant" in 2019. The proposed system aims to address the limitations of existing systems by creating a standalone personal assistant activated solely through the user's voice .performs various tasks such as reading text from images and controlling actions through voice-based commands. This system serves as a model for a range of applications.

In [3], At Gandhi Institute of Technology and Management in Hyderabad, Telangana, India, P Srinivas,T Sai Teja, CH Bhavana, R Likhith, and K

Satish Kumar authored a project titled "Raspberry Pi based Personal voice assistant using Python." The proposed system is a voice assistant that utilizes speech recognition to understand human speech and provide relevant responses in speech form. Deployed on a mini computer known as Raspberry Pi, this system offers a hands-free user experience and serves as a personal assistant, assisting users with everyday tasks. Portable and versatile, this device is suitable for usage in workplaces, homes, and various other environments.

In [4], At Alamuri Ratnamala I nstitute of Engineering and Technology in Sahapur, Asangon District Thane, Maharastra, India, Ms. Akshata Sarang and Mr. Shital Agrawal developed a project titled “Student Information INFOBOT system for Colleges” in 2020.The proposed system features a customer login and signup process as the initial interaction point for users, ensuring a personalized experience. Built with a focus on Natural Language Processing (NLP), the INFOBOT Responding Structure effectively addresses and assesses suspicions or dissent. Furthermore, the system seamlessly handles search requests within the database, enhancing its ability to provide relevant information. Moreover, equipped to respond to and address user complaints, the Chat BOT fosters a responsive and customer-centric interaction environment.

III. PROPOSED METHOD

The proposed system addresses challenges in accessing essential college information by integrating robotics and information dissemination. Using Raspberry Pi and ChatGPT API, we've developed a user-friendly platform for parents, newcomers, and visiting officials to easily access student profiles, course details, and academic performance data. By centralizing information and incorporating interactive features, we aim to enhance transparency and communication within the academic community. Through ongoing refinement, our system prioritizes accessibility and usability, ensuring stakeholders can efficiently navigate and utilize the platform to meet their informational needs.

3.2.1 Raspberry Pi 3B+:The Raspberry Pi 3B+ stands out as a highly capable single-board computer, boasting impressive specifications and versatile features. Powered by a 1.4GHz64-bitquad-coreARMCortex-A53processor, it delivers enhanced performance, making it suitable for a wide range of applications. With 1GB of LPDDR2 SDRAM, the Raspberry Pi 3B+ ensures smooth multitasking and efficient execution of compute-intensive tasks, providing users with a seamless computing experience.

Fig2:RaspberryPi3B+

In terms of connectivity, the Raspberry Pi 3B+ offers a comprehensive range of options to suit various needs. It comes equipped with built-in 2.4GHz and 5GHz wireless LAN, Bluetooth 4.2/BLE, and Gigabit Ethernet, enabling seamless wireless and wired communication .Additionally, its four USB 2.0 ports, HDMI output supporting resolutions up to 1080p, and GPIO pins for hardware interfacing further enhance its connectivity and flexibility for different projects.

Moreover, the Raspberry Pi 3B+ supports optional Power- over-Ethernet (PoE) via a separate PoE HAT, simplifying deployment in industrial settings where power management is crucial. Its compact form factor and compatibility with a wide range of operating systems make it a versatile platform suitable for diverse applications, including media centers, IoT projects, educational endeavors, and more.

Overall, the Raspberry Pi 3B+ offers a powerful and accessible solution, empowering enthusiasts, educators, and professionals to bring their ideas to life with ease.

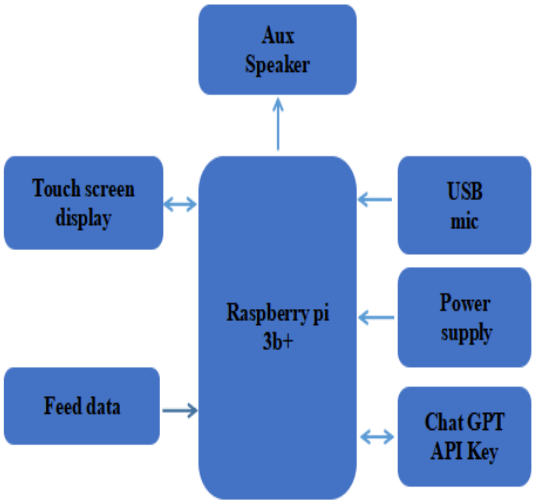


Fig 1: Block Diagram of Infobot System

3.2.2 AUX Speaker: The aux speaker, an integral component of our project's robotic assistant, fulfills a crucial role in audio output systems. Its primary function is to reproduce sound from connected audio sources, enabling the transmission of spoken responses, alert notifications, and other audio cues essential for effective communication with users. Selected for its clear sound reproduction and compact design, the aux speaker features a 3.5mm audio jack for compatibility with various devices, including smart phones, tablets, and single-board computers like the Raspberry Pi.

Fig 3: AUX Speaker

Key specifications of the aux speaker include its ability to deliver high-quality audio output, ensuring crisp and intelligible sound reproduction for an enhanced user experience. Operating on low power, it is energy-efficient and suitable for integration into battery-powered or low- power devices like our robotic assistant. Additionally, the aux speaker offers a balanced frequency response and impedance matched to the output capabilities of the audio source, ensuring optimal performance across a wide range of audio content types. Its durable construction further enhances reliability, making it suitable for diverse environments and applications within the academic realm of Electronics and Communication.

3.2.3 USB Mic: The USB microphone plays a crucial role in our project's audio input system, facilitating seamless interaction between users and the robotic assistant. Its primary function is to capture audio input accurately, enabling the assistant to receive spoken commands and queries with clarity. Equipped with a sensitive microphone element, the USB microphone ensures clear and intelligible communication, capturing sound accurately.

Fig 4: USB Mic

Microphone seamlessly integrates with various devices, including the Raspberry Pi used in our setup. Its direct connectivity to the Raspberry Pi's USB port simplifies setup and eliminates the need for additional audio interface hardware. Moreover, many USB microphones come with built-in noise reduction and echo cancellation algorithms, enhancing the quality of recorded audio by minimizing background noise and reverberation. This ensures the assistant can effectively interpret and respond to user commands, even in noisy environments. Despite its compact design, the USB microphone delivers high-quality audio recordings, enabling the robotic assistant to accurately capture spoken commands and queries. Overall, the USB microphone serves as a vital component, enabling effective audio input and interaction within the realm of Electronics and Communication.

3.2.4 Memory Card: The memory card plays a pivotal role in our project, serving as the primary storage medium for the Raspberry Pi. Its essential function includes storing the Raspberry Pi OS, system files, and user data necessary for the project's operation. Acting as the boot device, the memory card initiates the system startup process by containing the boot loader and boot configuration. Additionally, it facilitates the installation and storage of the Raspberry Pi OS, enabling users to set up and configure the system according to project requirements. Furthermore, the memory card provide sample storage capacity for user data, documents, media files, and application data, supporting data management and manipulation within the project environment. Its expandable storage capabilities allow for scalability, ensuring adaptability to future project expansions or modifications.

Fig 5: Memory Card

3.2.5 IPad (IKALL): The iKall iPad is a budget-friendly tablet offering a 7 to 10-inch HD display, typically powered by a quad-core processor with 1GB to 3GB of RAM and 8GB to 16GB of internal storage, expandable via micro SD. It runs on Android OS, providing access to Google Play Store.

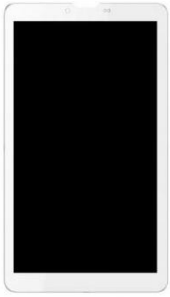


Fig 6: IPAD

The tablet features basic front and rear cameras (around 2MP or 5MP), moderate battery life, and standard connectivity options including Wi-Fi, Bluetooth, and sometimes 4G LTE. While its affordable price makes it accessible for basic tasks like web browsing, media consumption, and light gaming, the lower-end specs might limit performance with demanding apps and multitasking, and build quality may not be as robust as higher-end tablets.

3.2.6 Ethernet cable: An Ethernet cable is a type of network cable used for wired connections between devices such as computers, routers, and switches, enabling high-speed data transfer. These cables are essential for establishing Local Area Networks (LANs) and are commonly used in both home and office environments. Ethernet cables come in various categories, including Cat5e, Cat6, Cat6a, and Cat7, each offering different performance levels in terms of speed and bandwidth. Cat5e supports speeds up to 1 Gbps, while Cat6 and higher can support speeds up to 10 Gbps over shorter distances. The cables are typically constructed with twisted pairs of wires, encased in a protective sheath to reduce electromagnetic interference. Ethernet cables terminate in RJ45 connectors, making them compatible with a wide range of networking equipment. They offer a reliable and secure means of networking, often preferred for their stability and consistency compared to wireless connections.

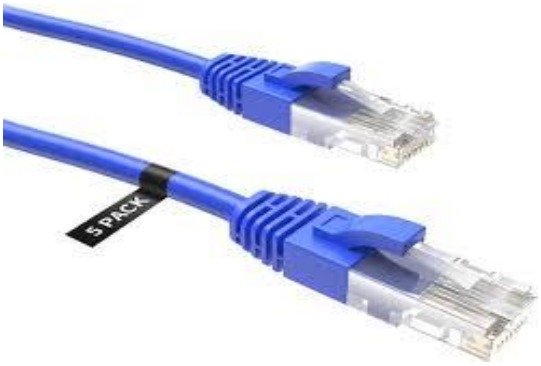


Fig 7: Ethernet Cable

3.2.7 HDMI Cable: An HDMI (High-Definition Multimedia Interface) cable is a standard connection for transmitting high-quality audio and video between devices such as TVs, monitors, computers, gaming consoles, and Blue-ray players. Introduced in 2003, HDMI cables support high-definition video resolutions, including Full HD (1080p), 4K, and 8K, along with multiple audio channels for excellent sound quality

Fig 8: HDMI Cable

Available in versions like HDMI 1.4, 2.0, and 2.1, these cables offer varying capabilities, with newer versions supporting higher resolutions and frame rates, enhanced audio, and features like ARC (Audio Return Channel), CEC (Consumer Electronics Control), and Ethernet over HDMI. Their ability to carry both audio and video signals through a single cable, coupled with backward compatibility, makes HDMI cables essential for modern digital connectivity.

3.2.8 USB to Micro B converter: A USB to Type B converter is an adapter that allows you to connect devices with a USB Type B port, commonly found on printers,

scanners, and some external hard drives, to a computer or other device with a standard USB Type A port. The converter typically has a USB Type A plug on one end and a USB Type B plug on the other, facilitating data transfer and device communication. These converters are essential for maintaining compatibility between older peripheral devices and modern computers, ensuring seamless connectivity for data transmission and peripheral control.

Fig 9: USB to Micro B Converter

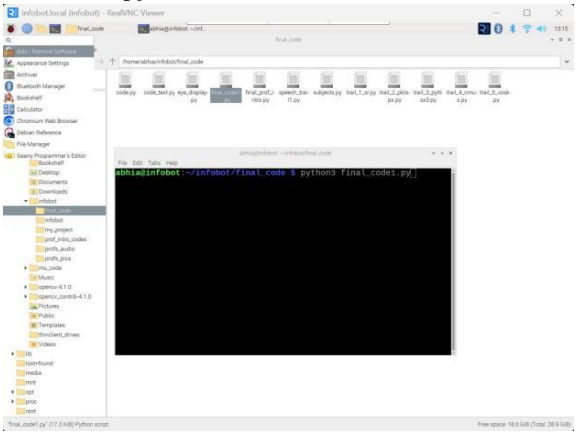
3.2.9 HDMI Video Capture: An HDMI video camera is a type of camera equipped with an HDMI output, allowing it to transmit high-definition video and audio directly to an external display, recorder, or live streaming device. This feature is particularly useful for professional videography, live broadcasting, and video production, as it enables real-time monitoring and high-quality recording. HDMI video cameras support resolutions such as 1080p (Full HD) and 4K, ensuring crisp and detailed video output. They often include additional features like optical zoom, image stabilization, and various shooting modes to enhance video quality and versatility. By connecting an HDMI video camera to compatible devices, users can achieve superior video playback and recording capabilities, making them ideal for content creators, filmmakers, and live streamers.

Fig 10: HDMI Video Capture

IV. RESULTS

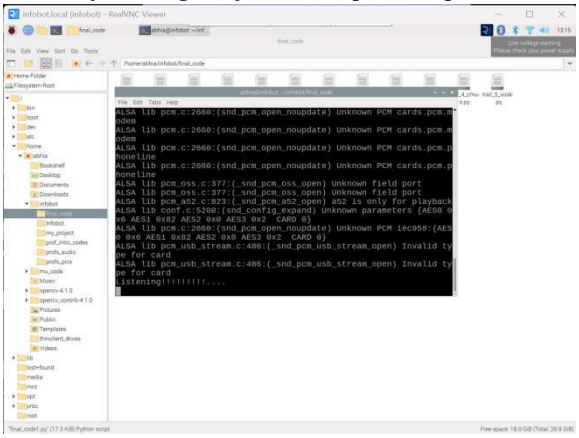
- Initiating the program to execution

The file where the program is present is opened in terminal. The program is initiated by running the prompt python3 final_code.py here final_code.py is the name of the program file. To execute the program, I accessed the file containing the code in the terminal. This file is named final_code1.py. Within the terminal, I initiated the program by running the command python3 final_code1.py. The python3 command tells the terminal to use the Python 3 interpreter to execute the code within the specified file, which in this case is final_code1.py.



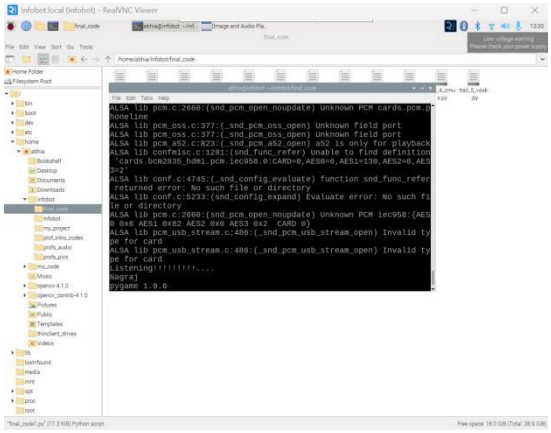
- Execution

Upon successful execution of final_code1.py, the program initializes the microphone. This process configures the microphone and prepares it to receive audio input. Once initialization is complete, the program typically displays a message on the screen indicating it's "listening" or ready to accept your spoken commands. This visual cue lets you know the program is actively waiting for your microphone input.



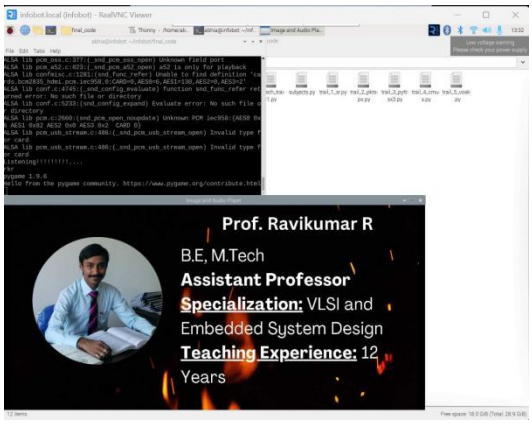
- Identification of speech and displaying it

The process begins with your Speech Input, where you provide instructions through the microphone. This spoken information is captured by the microphone and converted into a digital signal during Audio Capture. Next comes Speech Recognition, where the program analyzes this digital signal to identify the underlying speech patterns. These patterns are then translated into text format, essentially Text Conversion, which acts as a transcription of your spoken words. Depending on the program's design, the converted text might undergo further Text Processing. Finally, in Display Output, the program presents the recognized text on the screen, providing a visual confirmation of your spoken input.



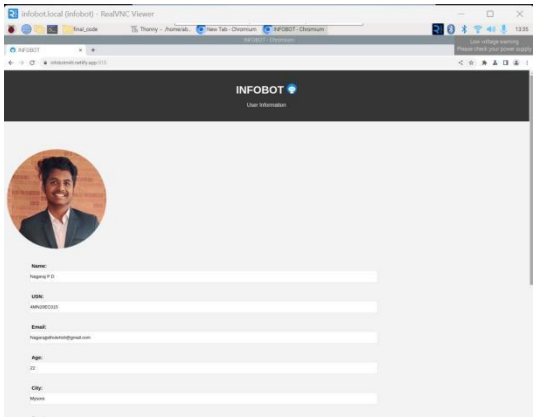
- Execution for professors data

The program handles professors too! Similar to students, it uses speech recognition to analyze your voice input and identify the professor's name. This name is then matched against an internal database. If found, the program displays the professor's name for confirmation.



- Output 2

Beyond just displaying the professor's name, the program can delve deeper into their profile if designed with additional features. This might include highlighting the professor's Specialization, showcasing their academic expertise. Additionally, details about their Teaching Experience could be displayed, providing insights into their teaching style and past courses. Finally, it's crucial to rely on a reliable and up-to-date data source to guarantee the accuracy of the displayed professor information.



REFERENCE

1. Emanuela Haller and Traian Rebedea, “Designing a Chat-bot that Simulates an Historical Figure”, IEEE Conference Publications, July 2013.
2. Maja Pantic, Reinier Zwitserloot, and Robbert Jan Grootjans, “Teaching Introductory Artificial Intelligence Using Asimple Agent Framework”, IEEE Transactions On Education, Vol. 48, No. 3, August 2005.
3. X. Cheng, Y. Bao, A. Zarifis, W. Gong, and J. Mou, “Exploring consumers’ response to text-based chatbots in e- commerce: the moderating role of task complexity and chatbot disclosure,” Internet Research, 2021.
4. L. Müller, J. Mattke, C. Maier, T. Weitzel, and H. Graser, “Chatbot acceptance: A latent profile analysis onindividuals’ trust in conversational agents,” in Proceedings of the 2019 on Computers and People Research Conference, 2019, pp. 35–42.
5. Bill Lubanovic. (2014). Introducing Python :Modern Computing in Simple Packages.
6. 2nd Edition. O'Reilly Media. (pp. 109-121). [2] Emad S. Othman, (2017),
7. Wadhwani, Preeti, and Saloni Gankar, (2017), “Intelligent Virtual Assistant Market Size & Share: Global Report 2024.” Global Market Insights, Inc., www.gminsights.com/industry-

analysis/intelligentvirtual-assistant-iva-market.
8. Uday Kamath et al, (2019), Deep Learning for NLP and Speech Recognition, Springer, (pp. 141–174).
9. Dhanush Kumar S, Lavanya S, Madhumita G, Mercy Rajaselvi V, (2018), "Journal on Speech to Text Conversion." International Journal of Advance Research, Ideas and Innovations in Technology, 4(2), (pp. 867-872).
10 Stephanie Diamond. (2014). Dragon NaturallySpeaking for Dummies. 4 th Edition. Wiley Publishing. (pp. 11-14).
11. Pearl Cathy, (2017), Designing Voice User Interfaces: Principles of Conversational Experiences, O'Reilly.
12. Li Deng and Dong Yu, (2015), Automatic Speech Recognition: A Deep Learning Approach, Edition-1, Springer-Verlag London, (pp. 221-231).
13. Smart sheet. (2020). How Voice Assistants Are Changing
14. En.wikipedia.org. (2020). Voice user interface. [online] Available at: https://en.wikipedia.org/wiki/Voice_user_interface [Accessed 26 Feb. 2020].
15. Components101.com. (2020). Pi Camera Module Pinout, Features & How to Use. [online] Available at: https://components101.com/misc/pi-camera-module [Accessed 26 Feb. 2020].
16. Pyowm.readthedocs.io. (2020). PyOWM — pyowm documentation. [online] Available at: https://pyowm.readthedocs.io/en/latest/ [Accessed 26 Feb. 2020].
17. McGregor Colin, (2006), "The Linux Infrared Remote Control (LIRC) Project." Linux Journal, www.linuxjournal.com/article/8811. [14] Monk Simon, (2012), Programming the Raspberry Pi: Getting Started with Python, McGraw-Hill, (pp. 1-159).
18. Manley William. (2014). "Tester: Creating LIRC Config Files - Part 1: Theory." Stb.