A Novel, Voice-Activated Smart Home Assistant Using a Large Language Model for Nutrition Assistance

Jay Bhardwaj, Hugo Carducci, Kenneth Clark, Nicholas Kennedy, Dylan Smith,
Tim Thai, Emily Francis, Ramana Reddy
Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, West Virginia, United States
jb00046@mix.wvu.edu, hwc0002@mix.wvu.edu, krc00002@mix.wvu.edu, nmk00001@mix.wvu.edu,
dds00002@mix.wvu.edu,tnt00001@mix.wvu.edu, ef00006@mix.wvu.edu, ramana.reddy@mail.wvu.edu

2024 IEEE Global Humanitarian Technology Conference (GHTC) | 979-8-3503-7734-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/GHTC62424.2024.10771529

***Abstract*—Malnutrition is a prevelant issue in older adults, as it can lead to adverse health effects such as heart disease, high blood pressure, and diabetes. To address this concerning issue, we propose a novel smart home assistant using Google's Gemini to provide nutritional guidance to a user. Results find that the proposed system can provide the user more in-depth answers - specifically in the case of recipe assistance - than Amazon Alexa and Google Nest. This indicates the potential for the proposed system to assist in combating malnutrition, by providing conversational answers to smart home inquiries.**

***Index Terms*—LLM, Smart Home Assistant, Voice Assistant, Nutrition**

I. Introduction

Around 25% of older adults suffer from some form of malnutrition [1]. For this reason, promoting a healthy, balanced diet is imperative for establishing physical health in older adults. Sticking to a healthy diet can be difficult when relying on one's own internal accountability, especially if one is unsure exactly how to stick to a healthy diet. Considering that around 27% of older adults live alone [2], the challenge of maintaining a healthy diet becomes even greater.

Moreover, according to the P&S Intelligence report, there is a projected market increase in the smart healthcare space from $8.7 billion in 2019 to $96.2 billion in 2030 [3]. In addition to the expected growth in the smart healthcare space, the COVID-19 pandemic has forced the medical field to re-evaluate options for the most effective healthcare practices. With remote and contactless care emerging as a viable solution, the medical field has adapted to using applications of Remote Patient Monitoring (RPM). RPM has been promoted by hospitals since the pandemic to provide at-home care that can be supported by smart systems [4].

Funding for this research provided by West Virginia University's Statler College of Engineering, via the Statler Fellowship award.

To address the concerning issue of malnutrition in older adults, while taking advantage of the growth in home healthcare, this paper proposes a novel, conversational, smart home assistant. This assistant aims to help the elderly access assistance with nutritional guidance while living alone, to diminish the chance of adverse health risks that come with malnutrition. To accomplish this, the assistant leverages a large language model, Google's Gemini to interpret user requests for nutritional assistance, and provides in-depth responses.

The rest of this paper is organized as follows. Section II discusses the current state of research regarding voice-enabled smart home assistants. Section III presents our novel home care assistant architecture and dataflow. Section IV presents an evaluation of our prototype implementation, and compares the proposed framework to commercially available home assistants. Section V provides a discussion of the proposed framework, including ethical considerations of the framework and anticipated future work.

II. Related Work

A. Voice-Enabled Smart Home Systems

Smart home systems are typically comprised of one or more devices to sense and control the environment. These devices are coordinated by some central hub computer [5]. Commercial smart home assistants such as Amazon's Echo and Google's Nest also include a voice-activated home assistant that listens for their respective wake words - such as "Hey Alexa" or "OK Google". Following the invocation of the wake word, a home

assistant then transcribes the speaker's words and calls upon the API to understand the user's intent.

The review of current research on smart home systems consisted of conducting a search for papers for the terms "Smart Home" or "IOT" and "Voice Assistant" in the title of the document. Papers were limited to the years 2020-2024 for the most up-to-date research.

Of the reviewed papers - 6 papers utilize Google Assistant [6]–[11], 6 papers utilize Amazon's Alexa [10]–[15], 1 paper utilizes Apple's Siri [11], and 3 papers implement a custom method for integrating a voice-enable home assistant [16]–[18]

In looking at systems with custom voice assistants, [16] utilizes Rhasspy - an offline voice assistant system for controlling smart homes, [17] utilizes the javascript framework Artyom.js, and [18] utilizes the Python library SpeechRecognition.

B. Literature Gaps and Contributions

One of the major drawbacks of existing home assistants is that they follow a question/answer model, where the user asks a question to the home assistant, the assistant replies with an answer, and the interaction then ends. We envision a home assistant where a conversation with the assistant can persist across user invocations of the assistant. For this reason, we are proposing an architecture that incorporates a Large Language model into a Smart Home Assistant for the user.

The main novel aspect of the proposed research is the integration of OpenAI's local voice recognition model Whisper and Google's Gemini.

III. Methodology

The proposed system prioritizes user experience with intuitive setup processes and a first-time setup wizard for guidance, alongside seamless integration with new smart devices and updates requiring minimal user effort. Reliability is ensured through error-checking mechanisms and unit testable functions, complemented by a crash recovery system that automatically restarts the app and logs crash reports. User preferences are stored in a structured database, with version control to prevent data loss. Robust security measures include data privacy safeguards, encryption, access controls, and regular vulnerability assessments, with end-to-end encryption employed for data transfers and storage. Advanced natural language processing facilitates natural conversations, enabling the assistant to handle basic web inquiries, set reminders, and interact with smart devices. The assistant maintains context from previous interactions, navigable via voice commands, with integration with local services providing relevant information and location-based recommendations for enhanced user experience. Speech recognition ensures coherent responses, while adjustable volume and speed settings cater to user preferences, with continuous adaptation based on collected user data for personalized interactions.

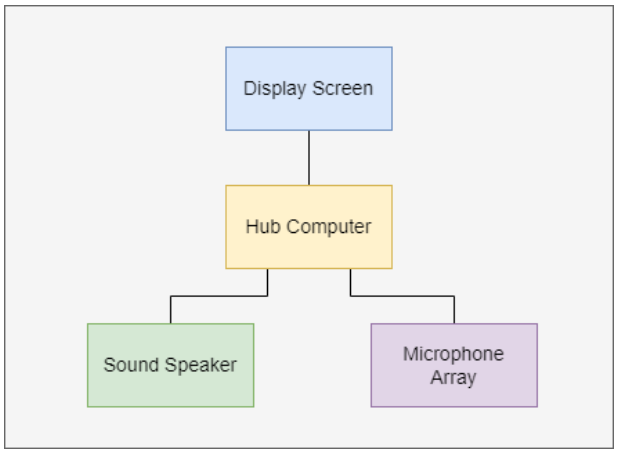

Fig. 1. Hardware Architecture

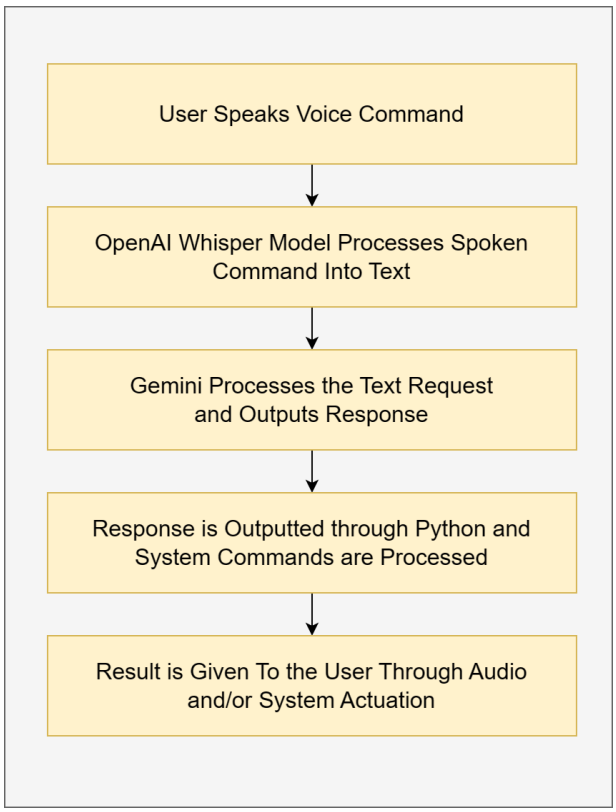

Fig. 2. Proposed Method

The hardware architecture, shown in Figure 1, consists of a hub computer, combined with a Microphone array and speakers, to demonstrate a local voice assistant for a smart home system.

Figure 2 shows the overall method that the proposed

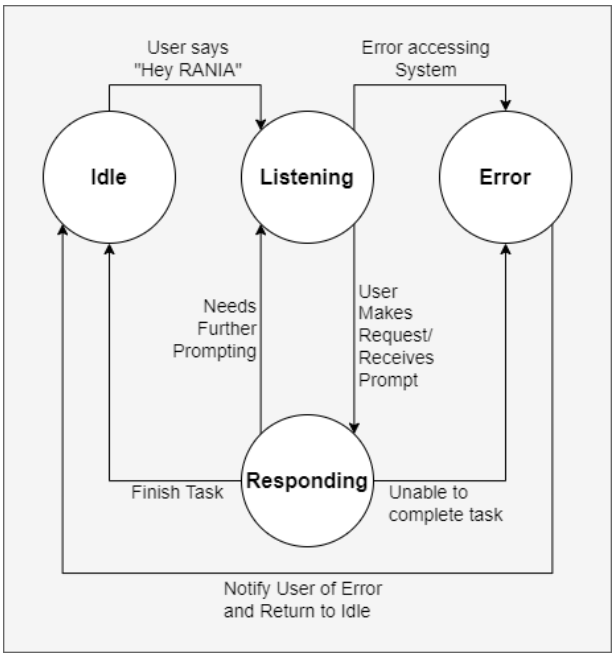

Fig. 3. State Transition Diagram of the System

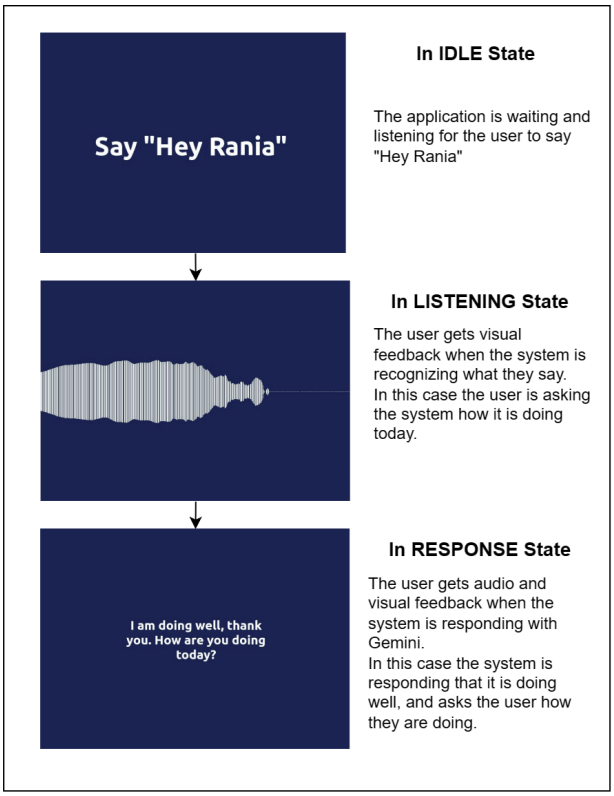

Fig. 4. User Interface of of the Chat Bot Assistant

home assistant will utilize. A local server will run OpenAI's "Whisper" speech to text model to transcribe user voiced input. This transcription is then processed by Gemini. Whatever is returned from the LLM is sent back to the local server and is fed through a local text to speech model for voiced output. See Table I for a list of components used in the project and the reasoning for using each component.

TABLE I
SYSTEM COMPONENTS

Component	Description
Gemini LLM	Google DeepMind's powerful Gemini LLM (successor to PaLM 2)
Whisper Speech-to-Text	OpenAI's Whisper Speech-to-text Model (STT) to translate user input
PocketSphinx	Carnegie Mellon University's Pocket-Sphinx is a continuous speech recognition engine
Coqui TTS	Responsible for performing text to speech services for the response
React and Electron	React and Electron frameworks to build an intuitive, easy to understand user interface
Raspberry Pi 4 Model B	A small, affordable, personal mini-computer that runs the voice recognition application and stores data locally
OpenBLAS	Able to achieve much faster performance on tasks that involve intensive linear algebra calculations, specifically speech to text processing in our case

Figure 3 shows the state transitions of the system. The system will start in an IDLE state where it waits for the key words to be spoken by the user. In our case, this project is being incorporated into a larger suite of home care devices dubbed the Resident's Aware Network for Intelligent Assistance (RANIA) System. For this reason, the key words are "Hey RANIA". Following this the system will transition to a Listening state and will translate the users speech to text in order to respond. Following the Listening state the command or question will be stored and either a response will be initiated or an error will result if the message or command cannot be understood. The response state sends the message to Gemini and uses text to speech to relay the answer to the user, unless Gemini needs more prompting in that case it will give feedback to prompting the user. If a response is given back the task is finished and RANIA will go back to idle.

IV. RESULTS

A photograph of the prototype system is provided in Figure: 5. We utilize a Raspberry Pi 4 Model B as the hub computer, along with a ReSpeaker Mic Array V2.0 for the microphone. In regard to speech recognition with OpenAI's Whisper model, there are multiple sizes of model to choose from. We use the smallest model, the

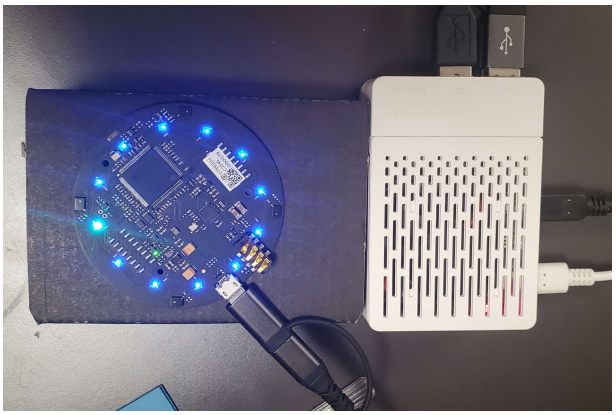

Fig. 5. Photo of Hardware Setup

'tiny' model to test speech recognition in the system. In our preliminary test of the system, we consider someone who is living alone. This person wants to follow a nutrition regiment provided by their doctor.

The proposed system is being compared to two common commercially available smart home assistants: Amazon Echo and Google Nest. All assistants were given the following prompt:

"Can you give me one recipe for breakfast that doesn't include eggs, but includes protein and has roughly 700 to 1000 calories?"

We evaluated the smart home assistants both quantitatively by the time (in seconds) it takes from end of user speech to assistant output and qualitatively by the perceived usefulness of the assistant's output. See Table II for a comparison of smart home assistants.

A. Amazon Alexa Response

When being asked to provide us with a recipe, Alexa takes around 3.5 seconds to provide the user with an answer.

"I don't know, but I would suggest something like yogurt."

The Amazon Alexa provided an honest response about not having a detailed answer, and suggested a quick, simple answer. It is worth noting that no supplementary Alexa "skills" were installed to provide extra nutritional assistance, which may be able to provide more in-depth answers to a user.

B. Google Nest Response

When being asked to provide us with a recipe, Google Assistant takes roughly 4 seconds to provide the user with an answer. "I don't know, but I found this on the web." The assistant is incapable of providing us with a detailed answer independently, and has to rely on supplying the user with internet resources (recipe websites, etc.) to answer the question. It may or may not be able to read the user the website, which is a disadvantage to visually impaired people.

C. Proposed Home Assistant Response

When being asked to provide us with a recipe, RANIA takes roughly 15 seconds to provide the user with a detailed answer.

The full output of the proposed home assistant system is as follows:

Powerhouse Breakfast Bowl

Serves: 1

Total Calories: 850

Ingredients:

1 cup rolled oats

2 cups almond milk

1/4 cup chia seeds

1/4 cup protein powder (whey or plant-based)

1/4 cup chopped walnuts

1/4 cup chopped almonds

1/4 cup dried cranberries

1/2 cup fresh berries (such as blueberries, strawberries, raspberries)

1 tablespoon honey

Instructions:

1. In a medium saucepan, combine the oats, almond milk, chia seeds, and protein powder. Bring to a boil over medium heat, stirring frequently. Once boiling, reduce heat to low and simmer for 5 minutes, or until the oats are cooked through.

2. Remove the saucepan from the heat and stir in the walnuts, almonds, cranberries, and fresh berries.

3. Divide the oat mixture evenly among two bowls. Drizzle each bowl with honey.

Nutritional Facts:

Calories: 850

Protein: 30 grams

Carbohydrates: 100 grams

Fat: 30 grams

Fiber: 15 grams

Sugar: 25 grams

V. Discussion

A. Ethical Considerations

Ethical considerations play a crucial role in ensuring developing technologies serve the greater good while minimizing harm, especially with the deployment of artificial intelligence models. The RANIA Smart Home Assistant makes use of a "large language model" (LLM) which is a type of deep learning model that is trained on great amounts of language data to achieve language generation as well as language comprehension. While

Home Assistant	Response Time	Relevance	Accuracy	Clarity	Actionability
Google Nest	4 seconds	Google Nest's audio response is not relevant to the question asked. It may be more relevant based on if the Google Nest is able to read the website it pulls information from	The audio response is not accurate to the prompt	The audio response is unclear to the requestor	The response is not actionable. The requestor can check the Google Nest if it has a display to read the website if necessary.
Amazon Echo	3.5 seconds	Yogurt is a common breakfast option in the United States	Yogurt does not include eggs, has a significant amount of protein, but does not fulfill the calorie requirements in the request	While the Echo is "unsure" of its response, it clearly defines yogurt as a suggestion	Yes, Alexa suggests yogurt, which aligns with the request and is a simple task to fulfill
Proposed Assistant	15 seconds	The recipe provided is relevant to the question asked	The recipe does not include eggs as an ingredient. The requestor must be sure to use a protein powder without eggs. The recipe claims 850 calories and 30 grams of protein per serving	The recipe description is thorough. The protein powder ingredient does not emphasize to check for egg protein in the protein powder.	The recipe is thorough and followable.

TABLE II
QUALITATIVE EVALUATION OF HOME ASSISTANTS

the primary goal of RANIA is to enhance the living standards of the elderly, it also brings several ethical challenges that need to be addressed to ensure the creation of a smart home assistant that is not just technologically advanced and user-friendly, but also ethically sound and secure.

In the case of employing an LLM based chat bot for home nutrition assistance, it is imperative that the system does not provide incorrect or harmful advice to a user - especially to avoid situations like Google's Gemini suggesting people to use glue to keep cheese on pizza. Information and recipes should be cross-checked with other sources to rule out misinformation.

It is also important to make sure the chat bot does not recommend recipes with ingredients that the requestor is allergic or sensitive to. For this reason, it is imperative to keep a running list of food allergies and sensitivities of people in the household, to avoid recommending recipes that may be harmful to the requestor.

Additionally, our project is incorporated into a wider-scaled home care system that manages multiple home care devices. In our prototype, we put a protective measure in place so that if the user asks about medical advice, the system responds with the suggestion to consult a doctor.

VI. CONCLUSION

The preliminary results highlight the weakness in the case of commercially available smart home assistants, when it comes to requesting nutritious meal assistance. Our preliminary tests seem to indicate that the Raspberry Pi 4 is not the ideal computer for this application. Utilizing the current hardware, the response time is too slow to be acceptable. More powerful hardware is recommended when implementing the proposed method of chat bot. With more powerful hardware than a Raspberry Pi, we believe this system shows promise for acting as a home assistant.

For future work, this assistant will be modified to be able to enact changes in the environment for the smart home system it is part of.

REFERENCES

[1] Naydeen M., Ahmed S., Omar R., Omar B. "Malnutrition-related mortality trends in older adults in the United States from 1999 to 2020." BMC Med. 2023 Nov 7;21(1):421. doi: 10.1186/s12916-023-03143-8. PMID: 37936140; PMCID: PMC10631109.

[2] Administration for Community Living, "2021 Profile of Older Americans", U.S. Department of Health and Human Services, Available from: https://acl.gov/aging-and-disability-in-america/data-and-research/profile-older-americans.

[3] Piekarz, Dan. "The Future of Smart Home Care Monitoring Technologies for Patients and the Elderly." *LinkedIn*, Available at: https://www.linkedin.com/pulse/future-smart-home-care-monitoring-technologies-patients-dan-piekarz/.

[4] Markets, Research and. "Global Smart Home Healthcare Market Report 2023: Increasing IoT Adoption Fuels Focus on Smart Healthcare Solutions - Growth, Trends, COVID-19 Impact, and Forecasts to 2028." *GlobeNewswire News Room*, Sept. 01, 2023. Available at: https://www.globenewswire.com/en/news-release/2023/09/01/2736207/28124/en/Global-Smart-Home-Healthcare-Market-Report-2023-Increasing-IoT-Adoption-Fuels-Focus-on-Smart-Healthcare-Solutions-Growth-Trends-COVID-19-Impact-and-Forecasts-to-2028.html.

[5] B. L. R. Stojkoska and K. V. Trivodaliev, "A review of Internet of Things for smart home: Challenges and solutions", Journal of Cleaner Production, vol. 140, pp. 1454—1464, January 2017.

[6] H. Isyanto, A. S. Arifin, and M. Suryanegara, "Design and Implementation of IoT-Based Smart Home Voice Commands for disabled people using Google Assistant," in *2020 International Conference on Smart Technology and Applications (ICoSTA)*, February 2020, pp. 1–6. https://ieeexplore.ieee.org/document/9079344/ doi: 10.1109/ICoSTA48221.2020.1570613925.

[7] C. Soni et al., "Multi-Language Voice Control IOT Home Automation Using Google Assistant and Raspberry Pi," in *2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, January 2022, pp. 1–6. https://ieeexplore.ieee.org/document/9752606/ doi: 10.1109/ACCAI53970.2022.9752606.

[8] A. Seekoli et al., "Sophisticated IOT Home Control Utilizing the Things Voice IOT Framework with Google Assistant," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, September 2023, pp. 1969–1973. https://ieeexplore.ieee.org/document/10398063/ doi: 10.1109/IC3I59117.2023.10398063.

[9] M. Gupta et al., "IoT Based Voice Controlled Autonomous Robotic Vehicle Through Google Assistant," in *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, December 2021, pp. 713–717. https://ieeexplore.ieee.org/document/9725526/ doi: 10.1109/ICAC3N53548.2021.9725526.

[10] A. K. Kaiborta and S. Samal, "IoT based Voice Assistant for Home Automation," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, January 2022, pp. 165–172. https://ieeexplore.ieee.org/document/9716533/ doi: 10.1109/ICSSIT53264.2022.9716533.

[11] H. Isyanto, A. S. Arifin, and M. Suryanegara, "Performance of Smart Personal Assistant Applications Based on Speech Recognition Technology using IoT-based Voice Commands," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, October 2020, pp. 640–645. https://ieeexplore.ieee.org/document/9289160/ doi: 10.1109/ICTC49870.2020.9289160.

[12] R. Kumar et al., "Power Efficient Smart Home with Voice Assistant," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, July 2020, pp. 1–5. https://ieeexplore.ieee.org/document/9225612/ doi: 10.1109/ICCCNT49239.2020.9225612.

[13] C. Jimenez et al., "Alexa-Based Voice Assistant for Smart Home Applications," *IEEE Potentials*, vol. 40, no. 4, July 2021, pp. 31–38. https://ieeexplore.ieee.org/document/9764005/ doi: 10.1109/MPOT.2020.3002526.

[14] J. S. Chavis et al., "A Voice Assistant for IoT Cybersecurity," in *2021 IEEE Integrated STEM Education Conference (ISEC)*, March 2021, pp. 165–172. https://ieeexplore.ieee.org/document/9764005/ doi: 10.1109/ISEC52395.2021.9764005.

[15] R. Aldhafiri, G. Powell, E. Smith, and C. Perera, "Voice-Enabled Privacy Assistant Towards Facilitating Successful Ageing in Smart Homes," in *2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, March 2023, pp. 343–345. https://ieeexplore.ieee.org/document/10150259/ doi: 10.1109/PerComWorkshops56833.2023.10150259.

[16] S. Dallmer-Zerbe and J. Haase, "Adapting Smart Home Voice Assistants to Users' Privacy Needs using a Raspberry-Pi based and Self-Adapting System," in *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*, June 2021, pp. 1–6. https://ieeexplore.ieee.org/document/9576469/ doi: 10.1109/ISIE45552.2021.9576469.

[17] S. Y. Torres-Maldonado et al., "Design and Implementation of a Voice Assistant to be used in an IoT Home Automation Environment," in *2023 27th International Conference on Circuits, Systems, Communications and Computers (CSCC)*, July 2023, pp. 1–8. https://ieeexplore.ieee.org/document/10361668/ doi: 10.1109/CSCC58962.2023.00027.

[18] R. P, K. Suresh, M. Boobalan, et al., "IoT Based Voice Assistant using Raspberry Pi and Natural Language Processing," in *2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, December 2022, pp. 1–4. https://ieeexplore.ieee.org/document/10046890/ doi: 10.1109/ICPECTS56089.2022.10046890.