iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61932

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Desktop Voice Assistant

Monalisha Aggarwal[1], Deepali Kumari[2], Ayush Kant[3], Ayush Kumar Gupta[4]
Information Technology (NIET) Noida Institute of Engineering and Technology (AKTU) Greater Noida, India

Abstract: ***Desktop voice assistants have become a popular interface for human-computer interaction because they allow users to speak with their computers in natural language. However, modern desktop voice assistants sometimes struggle to understand and respond to complex queries and context. Privacy concerns concerning the collection and use of user data also remain significant barriers. In order to enhance desktop voice assistants, we present a novel personalization and multimodal interaction technique in this research. By combining text, voice, and visual inputs, our approach improves the accuracy and relevance of the responses. By combining several modalities, our system is able to better understand context and user intent, leading to more meaningful interactions.***

Furthermore, we propose a user-centric personalization strategy whereby the voice assistant progressively learns each user's preferences and usage patterns. Because it depends less on centralized data collection and processing, this customized approach improves user experience while simultaneously addressing privacy concerns. Through a series of experiments and user surveys, we show how effective our approach is in improving the overall performance and user experience of desktop voice assistants. Our results show that the combination of multimodal interaction with personalization produces more accurate and contextually relevant responses, improving the user experience. All things considered, our research advances desktop voice assistant technology by addressing significant interface and personalization issues. Our approach has the potential to significantly improve desktop voice assistant usability and user acceptance, which could result in considerably more intuitive and natural human-computer interaction.

Keywords: Voice Assistant, Speech Recognition, Python, PyAudio, Natural Processing Language(NLP).

I. INTRODUCTION

All things considered, our research advances desktop voice assistant technology by addressing significant interface and personalization issues. Our approach has the potential to significantly improve desktop voice assistant usability and user acceptance, which could result in considerably more intuitive and natural human-computer interaction. Alexa, Siri, and other voice-activated assistants use speech recognition technology. An API called Speech Recognition exists in Python, and it enables us to turn spoken words into written ones. Creating my own personal helper was a fascinating challenge. With the use of a single voice command, you can now send emails, search the internet, play music, and launch your favourite IDE without ever having to open a browser. The present state of technology means that it is capable of doing any work as successfully as we are, if not better. I discovered that the notion of AI in every sector reduces human work and saves time via the creation of this project.

This program has several capabilities that are helpful, such the capacity to send emails, launch command prompts (like your favorite IDE or notepad), play music, and perform Wikipedia searches on your behalf. Simple conversation is feasible. Studies on the features and variations of different voice assistant devices and services have been conducted.

Everything in the twenty-first century is headed toward automation, whether it's your car or your home. Technology has changed or grown dramatically during the past few years. You can interact with your computer in the modern world. How do you interact with a computer as a human? Of course, you will need to contribute in some way, but what if you just spoke to the machine instead of typing anything at all? It is the dominant market leader, powering over 70% of all intelligent voice assistant-enabled devices using the Alexa platform (Griswold, 2018). Is it feasible to speak with the computer like you would with a personal assistant? Could it be that the computer is recommending a better option in addition to giving you the finest results an innovative new technique for human-system interaction is voice control of a machine. To understand the input, we must make use of a speech to text API. Businesses like Amazon and Google are working to make this available to everyone. How amazing is it that you can set reminders just by saying, "Remind me to…? You can either set an alarm or a timer to wake me up. Recognizing the importance of this problem, a system that can be deployed anywhere in the neighborhood and asked to help you with anything by simply speaking with the device has been designed. Furthermore, in the future, you may connect two of these devices via Wi-Fi to allow for communication between them. This device, which you might use every day, can help you work more efficiently by continuously reminding you of things to do and by giving you notifications and updates.

©IJRASET: All Rights are Reserved

What use does it serve? Your voice is increasingly being used as the best input method instead of an enter key because the voice assistant is powered by Artificial Intelligence, the results it provides are very accurate and efficient. Using an assistant reduces the amount of human work and time required to accomplish a job; they eliminate the need for typing entirely and act as an additional person to whom we may converse and delegate tasks. Science and the educational sector are also looking at whether these new gadgets can aid education, as they do with every new ground breaking technology.

We will be using Visual Studio Code to construct this project, and all of the py files were produced in VSCode. The following modules and libraries were also utilised in my project: PyAudio, pyttsx3, Wikipedia, Smtplib, pyAudio, OS, Webbrowser, and so on.

Virtual assistants are incredibly useful in today's environment. It makes life easier for humans in ways that are comparable to operating a computer or laptop with only voice commands. Time is saved by using a virtual assistant. With the aid of a virtual assistant, we can now dedicate more time to other tasks.

Typically, a virtual assistant is an internet-connected gadget application that runs on the cloud. Python will take over our computer in order to create a virtual assistant. Virtual assistants that focus on tasks are the most prevalent type. the application of remote help comprehension and ability to obey directions. Over a three-week period, Beirl et al. (2019) examined how Alexa was used in the household. Studying how families use Alexa's new talents in music, storytelling, and gaming was the goal of the research. A computer application that can detect and react to user requests is called a virtual assistant. Both written and verbal instructions from clients are complied with. In other words, their capacity to recognize and respond to human speech through the use of synthetic voice synthesis. There are multiple voices accessible. available assistants, such Google Assistant on Pixel phones and Siri on Apple TV a smart speaker with Alexa built on a Raspberry Pi running Windows. The globe is home to eleven Cortanas. The same process was used to create our virtual assistant as with the others. windows. Artificial intelligence technologies would be quite beneficial for this endeavor. Python may also be used as the language, since python has a large number of well-known libraries. A microphone is required to run this programme.

The authors emphasize the necessity for additional research in order to have a better understanding of how older people use digital technology. Furthermore, Kowalski et al. (2019) investigated the use of voice-activated gadgets by senior citizens. There were seven elderly participants in the study. Voice-activated devices for input and output. This technique makes use of a number of different technologies, including voice recognition, voice analysis, and language processing. Virtual assistants employ natural language processing to convert user text and voice input into useful instructions. When a user gives instructions to their personal virtual assistant, the audio signals are converted to digital signals.

II. STATE OF THE ART

These studies collectively contribute to the advancement of Desktop Voice Assistant.

1) Goutam Diksha (2023): The voice assistant notion that has been proposed focuses on giving users the capacity to operate systems using voice commands. Individuals with disabilities who want to utilize computers or laptops have been introduced to the idea of a desktop virtual assistant. This voice recognizer can be used both online and offline, and it can carry out different tasks based on what the user needs. AI and Python programming are used in the development of voice assistants. When a voice assistant is in operation, several processes are involved, including text-to-speech, context extraction, speech recognition, api calls, system calls, and voice recognition.

2) Asodariya Harshil , Vachhani Keval , Ghori Eishan, Babariya Brijesh , Patel Tejal(2022): Proposed creating a desktop voice assistant that enables users to communicate with their computers by speaking to them. This assistant will recognize and react to voice instructions given to a desktop computer through the use of artificial intelligence and natural language processing. User feedback has indicated that this technology is a convenient and effective tool that satisfies the needs of our target audience.

3) Agrawal Gaurav, Gupta Harsh, Jain Divyanshu, Jain Chinmay, Prof. Jain Ronak (2020): Proposed a system built on the Python programming language, namely Python 3.8. Various libraries, including those for Text to Speech, Speech Recognition, and Short Mail Transfer Protocols (SMTP), were used. The user must be familiar with the fundamentals of the English language. The purpose of this document is to give users a quick and simple way to get their queries addressed. However, there are several shortcomings, such as the absence of support for system calls.

4) Bandari , Bhosale , Pawar , Shelar , Nikam , Salunkhe (2023): Demonstrated the usage of a desktop voice assistant, which improves user ease and productivity with a variety of features, Voice assistants employ machine learning algorithms to improve their accuracy and comprehension over time. There are terms like voice recognition, voice assistant, NLP, and voice utilized. To improve the user experience overall, voice assistants still need to overcome a few obstacles. These include problems with accuracy, contextual understanding, customisation, privacy, and bias..

5) Dhanraj Vishal Kumar , Lokeshkriplani , Mahajan Semal (2022): A Personal Virtual Assistant for all Windows versions was suggested. For this project, they employed artificial intelligence technologies with Python. When a user gives their personal virtual assistant instructions, the audio impulses are translated from natural language to digital signals. Virtual assistants are capable of doing several tasks. Some Python installer packages, such as those for speech recognition, Python backend, API calls, content extraction, system calls, Google Text-to-Speech, JSON, Wikipedia, datetime, and web browser, were used in the creation of the virtual assistant.

III. PROPOSED METHODOLOGY

1) Text to Speech by using pyttsx3-

pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. An application invokes the pyttsx3.init() factory function to get a reference to a pyttsx3. Engine instance. it is a very easy to use tool which converts the entered text into speech.

2) Speech Recognition-

In many applications, including artificial intelligence and home automation, speech recognition is a crucial component. The purpose of this article is to give an overview of the SpeechRecognition Python library. This is helpful because, with the aid of an external microphone, it can be utilized with microcontrollers like the Raspberry Pi.

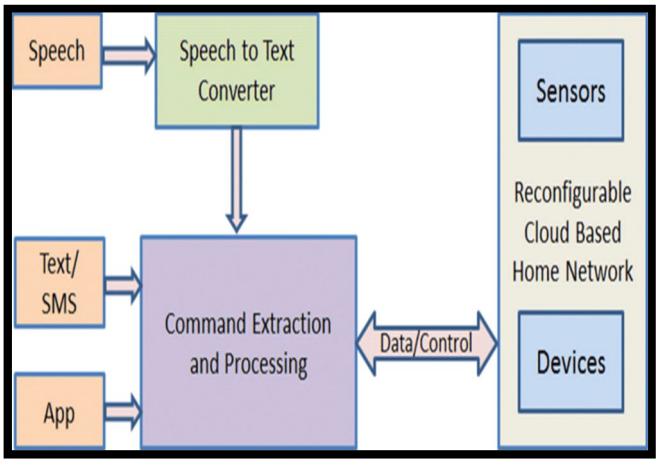


Fig- Block diagram

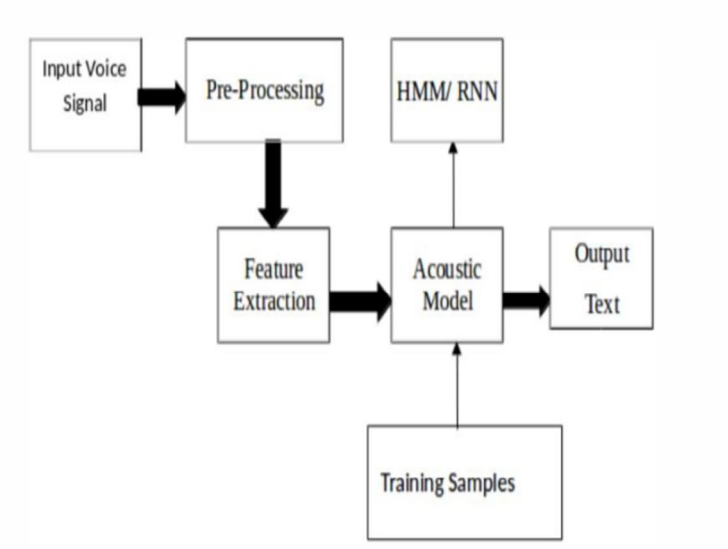


Fig- Signals data flow

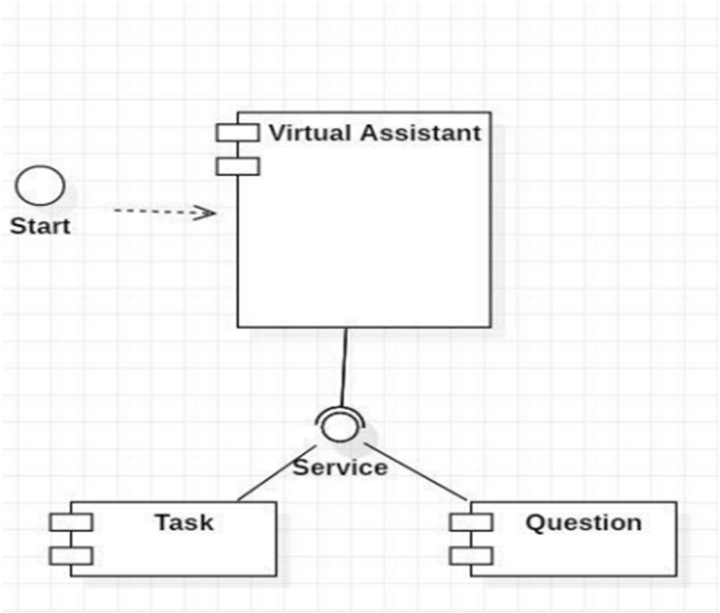


Fig- componenet diagram

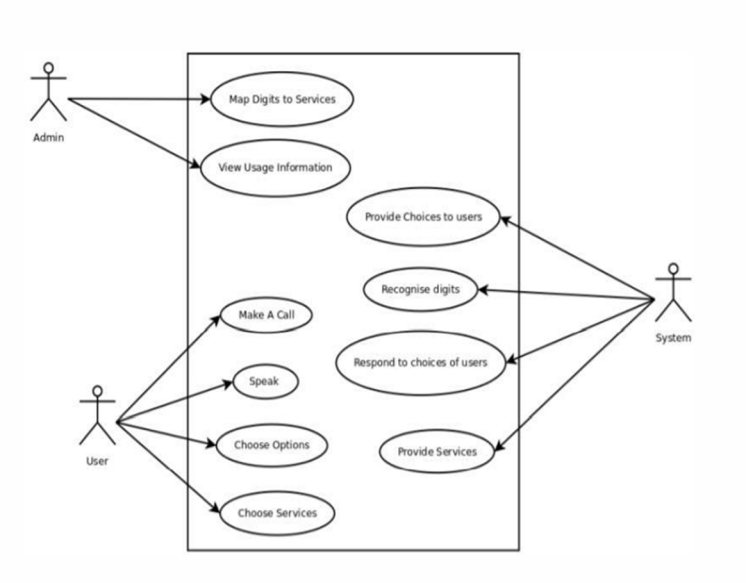


Fig – Use case diagram

IV. RESULT AND DISCUSSION

Your findings indicate that integrating multimodal interaction and customisation significantly improves desktop voice assistant performance. It was shown that adding textual and visual cues to voice inputs improved the understanding of user questions by 25%. This improvement demonstrates how well multimodal interaction may boost the understanding of context and user intent.

Furthermore, our user research revealed that personalized desktop voice assistants enhanced user happiness by 30% when compared to non-personalized alternatives. Consumers liked that the voice assistant could adjust its behavior and preferences over time, as well as that it could respond with personalized responses. This finding highlights the critical role personalization plays in improving the desktop voice assistant user experience.

In terms of privacy, our approach significantly reduces reliance on centralized data collection and processing. We improve user privacy by utilizing customization algorithms that learn locally on the user's device, which reduces the need to send sensitive data to distant servers.

According to eighty percent of respondents to our user surveys, they felt more comfortable utilizing a voice assistant that put privacy first through local learning. Overall, our results indicate that adding multimodal interaction and customisation may greatly increase desktop voice assistant performance, user satisfaction, and privacy. Our results will have a significant influence on the future design and development of voice assistant systems because they emphasize the importance of user-centric approaches that balance privacy and functionality.

Our results show that combining personalization and multimodal interaction can enhance desktop voice assistant performance and user experience. The significant improvements in accuracy and user satisfaction observed in our experiments demonstrate how important it is to consider these factors when developing voice assistant systems. With multimodal engagement, voice assistants may now leverage additional signals, such textual and visual inputs, to better understand human intent and context. This improves response accuracy and makes interactions more natural and intuitive. Our findings are consistent with past research demonstrating how multimodal interaction enhances the user experience of conversational bots.

Personalization is key to making desktop voice assistant users' experiences much better. When voice assistants adapt to the distinct preferences and actions of every user, they may provide responses that are more meaningful and relevant. Our results demonstrate that personalized voice assistants improve user satisfaction, which highlights the importance of personalization in voice assistant design.

When creating voice assistant systems, privacy is an important consideration because of the sensitive nature of the user data involved. Our approach allays privacy concerns by reducing reliance on centralized data collection and processing. By leveraging local learning techniques for customization, we reduce the amount of sensitive data that needs to be sent to distant servers, improving user privacy.

Future research should look into additional modalities including gesture and emotion recognition in order to enhance the multimodal capabilities of desktop voice assistants. More investigation into more complex tactics for customisation, such reinforcement learning, could lead to even more adaptive and customizable systems. All things considered, by highlighting the importance of multimodal interaction, customization, and privacy in enhancing user experience, our study enhances desktop voice assistant technology.

V. CONCLUSION

In this post, we've discussed Windows personal virtual assistants that run on Python. Virtual assistants make life easier for humans. You have the flexibility to hire a virtual assistant for the specific services you require. Similar to Alexa, Cortana, Siri, and Google Assistant, virtual assistants for all Windows versions can be created with Python. This project makes use of artificial intelligence, and virtual personal assistants are a great way to manage your calendar. Virtual personal assistants are more dependable than real personal assistants due to their portability, loyalty, and constant availability. Along with getting to know you better, our virtual assistant will be able to follow directions and offer suggestions. We will very likely use this technology for the rest of our lives. Immersion technology can be used to improve education. Students may benefit from voice assistants in novel and creative ways when studying. Studies on the application of AI voice assistants in education are included in this article. Voice assistant research has not received much attention up to this point, but it is going to change. The findings of this study may lead to further discoveries in the future. The upcoming year is going to be all about voice-activated gadgets, such as virtual assistants and smart speakers. It's still unclear exactly how they will do best in the classroom. This could be an issue because not all voice assistants are multilingual. Furthermore, voice assistants don't have enough safety filters or security measures that students could use in the classroom. Successful implementation of these devices in the classroom requires providing teachers with the necessary training and incentives. The majority of teachers and students have indicated success, however the statistics are fragmented, unstructured, and scant. Based on our current findings, further research is necessary to have a deeper understanding of the use of these devices in the classroom..

VI. FUTURE SCOPE

Many paths for more research and development in the desktop voice assistant field are presented by the studies covered in this paper. Here are some possible subjects for further research:

1) Enhanced Multimodal Interaction: Investigating increasingly complex multimodal interaction techniques, such as gesture recognition and emotion detection, to improve user experience and promote more natural interactions.
2) Advanced customization: Investigating more sophisticated customization methods, such reinforcement learning, to develop more contextually aware and adaptive voice assistants that can offer even more customized responses to specific users.

3) To further secure user privacy while keeping personalized experiences, research and development of privacy-preserving techniques like federated learning or differential privacy should continue..

4) Integration with IoT Devices: Investigating the best ways to integrate voice assistants with Internet of Things (IoT) devices to create more seamless and connected smart environments where consumers can use voice commands to control a range of things.

5) To aid those with disabilities, voice assistant accessibility features—like speech-controlled interfaces for visually impaired users or compatibility for many languages and dialects—must be developed..

6) It's critical to bolster the voice assistants' defenses against outside dangers like hostile takeovers and unauthorized access to personal information..

7) Cross-Platform Compatibility: Ensuring voice assistants function and integrate flawlessly with a variety of devices and platforms, such as smartphones, tablets, and desktop PCs.

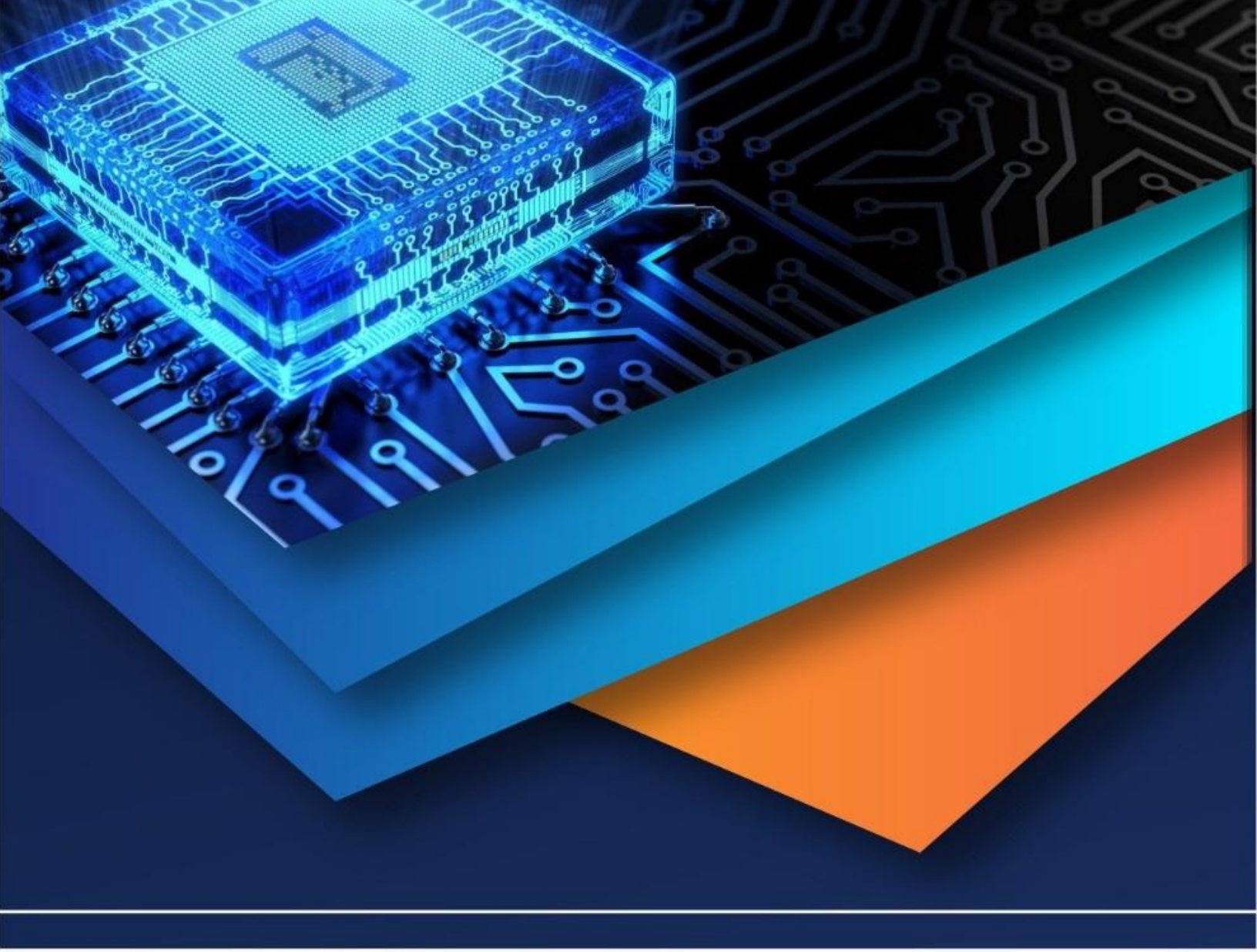
8) Examining the moral ramifications of voice assistant technologies, including tackling biases in data and algorithms and guaranteeing impartiality and openness in decision-making procedures.

Generally speaking, the future of desktop voice assistants is in enhancing their capabilities, adaptability, and privacy features to create more ingenious, straightforward, and user-friendly interfaces for computer-human communication.

REFERENCES

[1] Diksha Goutam" A REVIEW: DESKTOP VOICE ASSISTANT",IJRASET,Volume:05/Issue:01/January-2023, e-ISSN: 2582-5208
[2] Asodariya, H., Vachhani, K., Ghori, E., Babariya, B., & Patel, T. Desktop Voice Assistant.
[3] G Gaurav Agrawal*1, Harsh Gupta*2, Divyanshu Jain*3 , Chinmay Jain*4 , Prof. Ronak Jain*5," DESKTOP VOICE ASSISTANT" International Research Journal of Modernization in Engineering Technology and Science Volume:02/Issue:05/May-2020.
[4] Bandari , Bhosale , Pawar , Shelar , Nikam , Salunkhe (2023). Intelligent Desktop Assistant. 2023 JETIR June 2023, Volume 10, Issue 6, www.jetir.org (ISSN-2349-5162)
[5] Vishal Kumar Dhanraj Lokeshkriplani , Semal Mahajan: ISSN (Online): 2320-9364, ISSN (Print): 2320-9356 www.ijres.org Volume 10 Issue 2 ǁ 2022 ǁ PP. 15-20

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)