Voice Assistants: The Digital Age's Emergence of Virtual Friends

Atharva Singh*, Komal Goel, Haider Abbas, Mohd Izhar, Faisal Rais

ADGIPS, FC-26, Panduk Shila Marg, Zero Pusta Rd, Shastri Park, Shahdara, New Delhi, Delhi 110053

* Corresponding author

doi: https://doi.org/10.21467/proceedings.178.19

ABSTRACT

The main goal of the Desktop voice assistant is to provide people with an easy and quick way to get their questions answered to manage their time more effectively and increase productivity and interaction or to build a Desktop assistant that can perform tasks and provide information to users through voice commands. It can play music, manage schedules, perform home automation, and personal organisation using our voice. The desktop assistant will provide a user-friendly interface for various tasks as there is a large scope of voice technology. Voice assistants have transformed human-computer interaction by enabling intuitive, hands-free gadget control. The creation and operation of a desktop voice assistant intended to improve user experience, accessibility, and productivity are presented in this study. Utilising developments in artificial intelligence (AI), speech recognition, and natural language processing (NLP), the system decodes user input to carry out operations such system automation, application control, and information retrieval. The suggested solution incorporates custom frameworks for smooth desktop integration and APIs such as OpenAI's GPT for contextual understanding. In addition, this article examines issues with privacy, accuracy, and real-time performance and suggests solutions. The system's effectiveness and user happiness are demonstrated by experimental evaluations, indicating that it may be widely used in both personal and professional contexts.

Keywords: Voice assistant, voice commands, Speech Recognition, AI, Desktop Assistant

1 Introduction

All of us **l**ive in exponential technological growth and development, where everything is becoming more automated and sophisticated. A new field called voice assistants has emerged, offering a reliable and easy way to communicate with computer systems. One kind of voice enabled artificial intelligence (AI) is voice assistants (VA). Artificial Intelligence (AI) is the degree of intelligence that digital interfaces can exhibit, or the capacity of algorithms to imitate intelligent human behaviour. Even so, AI denotes to processes like learning and problem solving that are connected to the human mind. Voice assistants are now included in a wide range of devices, including computers, Bluetooth headsets, and smartphones. Voice assistants are becoming a necessary component of our daily existence. They support us in all areas, such as education (some apps come with their assistants that help students learn new things) and healthcare (they help with scheduling consultations and appointments). Voice assistants are generally very significant in our lives and will only get bigger in the future. Although voice assistants are still a relatively new market, voice-enabled technology is becoming more and more widely used. Between 2021 and 2023, the voice assistant market is predicted to expand by an average of 28% annually. Moreover, estimates suggest that by 2023, there will be over 8.4 billion voice assistants worldwide (including those with built-in software and Bluetooth speakers) more than the number of people on the earth. With over one-third of Americans using voice assistants, there were 115.2 million users in 2019 and 135.6 million by the end of 2022. Although use is growing across all age groups, millennials are the largest consumers.

Proceedings DOI: 10.21467/proceedings.178; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-8-1

2 History of Voice Assistant

It's amazing to see how much the history of voice assistants has changed over time. Here's a quick rundown Early Examples: The idea of voice-controlled assistants originated in science fiction, and the computer on Star Trek's USS Enterprise and HAL 9000 from "2001: A Space Odyssey" are two examples of early examples in popular culture. The first useful voice recognition systems appeared in the 1980s and 1990s. IBM's "IBM Shoebox," which debuted in 1986 and was capable of understanding and reacting to about sixteen spoken words, is one famous example. 2000s: More sophisticated voice recognition technologies emerged in the 2000s, although they were primarily confined to particular sectors like call centres and the automobile industry. The capacity of products like Dragon Naturally Speaking to convert spoken 2010s Voice assistants saw significant progress thanks to the introduction of smartphones and the growth of artificial intelligence. When Apple's Siri originally came out in 2011, it was the first voice assistant that was utilised widely. 2012 saw the release of Google Assistant (previously Google Now), while 2014 saw the release of Amazon Alexa along with the Amazon Echo. Extension and Integration: Voice assistants have been incorporated into a growing number of platforms and devices, such as smart TVs, smart speakers, smartphones, and even automobiles. They went above and beyond basic duties like answering inquiries and setting reminders to include managing smart home appliances, making tailored recommendations, and more.

Enhancements in Artificial Intelligence and Natural Language Processing: Progress in NLP, AI, and machine learning has made it possible for voice assistants to comprehend and react to natural language inquiries more effectively .Confidentiality and Ethical Concerns: As voice assistants proliferated, worries about data security and privacy also increased. Concerns about data exploitation, private conversation recordings, and inadvertent activations spurred talks about privacy protections and openness. Persistent Innovation: The accuracy, usefulness, and integration of voice assistants with other technologies are constantly being enhanced. They are getting more and more individualised, adjusting to the tastes and routines of users to deliver more customised experiences. The history of voice assistants, from early trials to global adoption and continuous improvement spurred by advancements in AI and consumer demand for more user-friendly and convenient interfaces, generally mirrors the broader developments in technology. words into text have led to their rise in popularity.

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

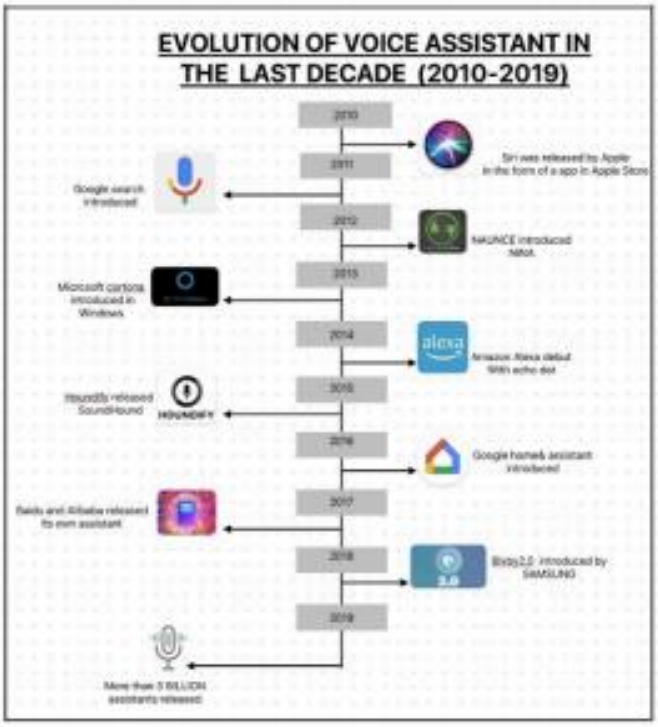

Figure 1: Evolution of Voice Assistant

3 Evolution of Voice Assistant

Voice assistants have gone through a remarkable transformation, evolving from basic command-based systems to complex AI-powered instruments. Only numbers and simple commands could be recognised by early speech recognition systems, like IBM's Shoebox in the 1960s. Improvements were made in the 1990s with programs like Dragon Dictate, which offered dictation features but necessitated a lot of training. When Apple released Siri, the first widely used voice assistant built into a smartphone, in 2011, it was a significant advancement since it used natural language processing (NLP) to comprehend conversational requests. This was followed by Google Assistant, Amazon Alexa, and Microsoft Cortana as shown in Figure 1 which added further integrations with smart home devices and the cloud.These days, AI models such as GPT and Whisper have improved voice assistants' conversational skills, allowing them to understand sentiment, context, and complicated commands. Voice assistants are now essential tools in industrial, professional, and personal settings thanks to these developments as well as the incorporation of machine learning and the Internet of Things.[3].

1) SIRI: An initiative created by the SRI International Artificial Intelligence Centre gave rise to Siri. Nuance Communications supplied their speech recognition engine, which operates by utilising cutting-edge machine learning techniques. Around 2005, its original voice actors—American, British, and Australian—recorded their voices, not realising that the recordings would eventually be used. The Siri iOS app was released in February 2010. Two months after purchasing it, Apple included it in the October 4, 2011, release of the iPhone 4s, removing the independent application from the iOS App Store. Since then, in addition to the more recent iPhone models, Siri has been integrated into several Apple hardware products, including the iPad, iPod Touch, Mac, Air Pods,

Apple TV, and Home Pod [4].

2) ALEXA: The creation of Alexa began on November 6, 2016. Inspired by the computer speech and conversational systems aboard the U.S.S. Enterprise in Star Trek, it was unveiled by Amazon concurrently with Echo. With Amazon Echo, we can link a wide range of gadgets.

3) GOOGLE ASSISTANT: Google Assistant is a virtual assistant that can be accessed on mobile and home automation devices. It is developed by Google. In October 2016, it debuted on Google Pixel smartphones. Later, it spread to additional Android and iOS devices, wearables, smart speakers, and other gadgets [14].

4) CORTANA: Microsoft developed Cortana, which was first included in Windows Phone 8.1 in April 2014. The AI character from the Halo video game series inspired the name of the system. Like other virtual assistants, it can carry out tasks and comprehend and carry on discussions. Cortana was said to have around 800 million users as of 2018.

4 System Structure and Technical Requirement

The spoken words must be captured and converted to text. Next, it must be determined whether it can use its artificial intelligence algorithms to extract the purpose. Verifying that the Virtual Assistant can reply by the intent inferred in the previous stage is the next logical step. Certain answers must carry out system commands, while others must change certain variables on Internet of Things (IoT) devices or obtain information from third-party Application Programming Interfaces (APIs) (such as weather and other applications). The software's response will depend on how confident it is in the intended outcome [5]. We chose Python as a programming language above other languages to fulfil our voice assistant's need for voice commands. Python has a clear and concise syntax that is very easy to understand and learn, as well as rich ecosystems of libraries that are specifically designed for developing AI & ML models. Python is an extremely flexible and adaptable language that can be used for a wide range of tasks. Apart from all this python can be easily integrated with databases and web frameworks making it a convenient option for deploying models. Here for the overall development of the assistant, we have used various Python libraries like pyttsx3(TTS library for converting text to speech), speech recognition, WEB browser, bs4, date time, google trans, etc. Because of Flutter's efficiency and versatility with other languages, we decided to choose it over other frameworks to create the user experience, which takes the shape of a personalised desktop assistant.

Flutter makes it incredibly flexible and efficient for developers to build a single code that can be used to create both iOS and Android devices. This helps developers to have a single codebase for each platform. Additionally, Flutter offers a variety of customisable features and widgets that contribute to a rich user interface.

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

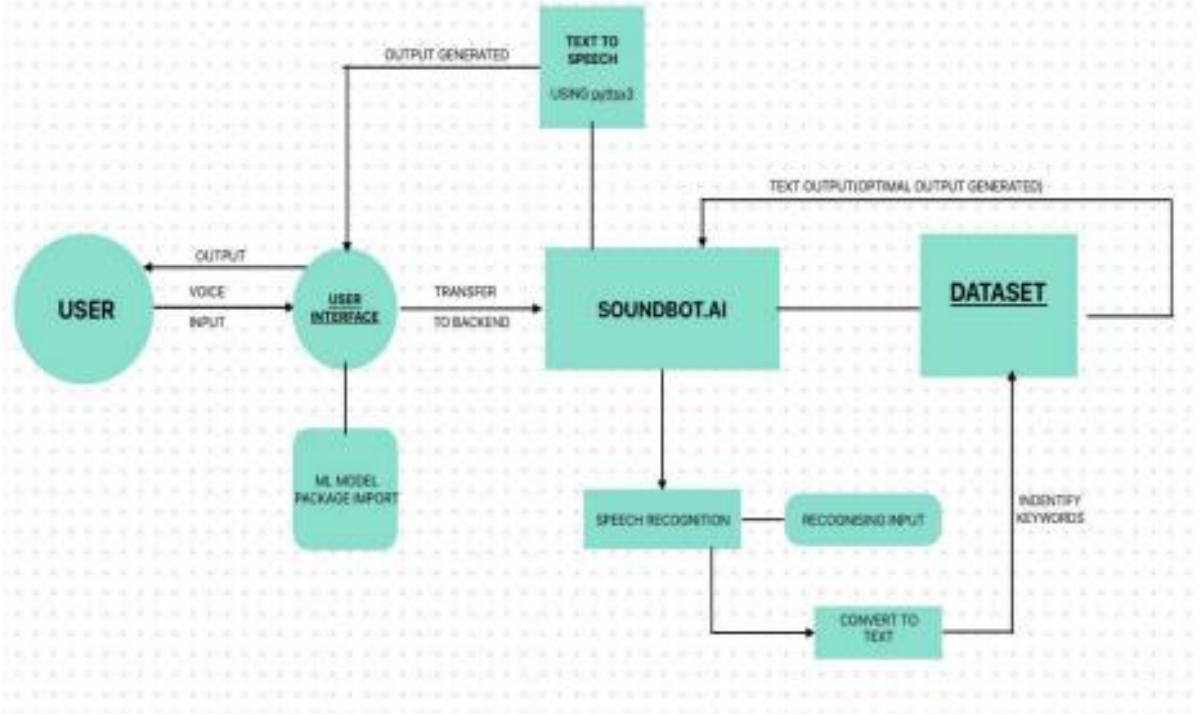

Figure 2: Workflow of the voice assistant

5 Workflow of Voice Assistant

This interaction model processes speech queries and produces speech output by using commercial NLP algorithms (Amazon) as an endpoint, which is an internet accessible skill hosting service. Based on the interaction model, the skill receives requests and responds with actions, such as giving audible feedback or interacting with the device shadow, which is a digital depiction of the state of the device at that moment [6].

As a desktop assistant, our voice assistant has been incorporated into the user interface. The model can be activated by saying the keyword "WAKE UP." This will cause the model to begin operating and collect user voice input, which it will then analyse and transform into text using the Python speech recognition module.This sends the text to the model, which analyses it and then completes the designated task or function. If a question has been posed, the output will solely use the Python Text To voice (Pyttsx3) package to produce voice as shown in Figure 2 .To enhance the model's adaptability, many queries have been produced and numerous functions have been developed as subclasses and connected to the primary model.

6 Result

Thus, by allowing blind people to get information simply by using their voice to provide requests, this voice assistant is making their lives easier. Additionally, our model guarantees accuracy. The images below display a few of our model's outcomes. Here the voice assistant is performing and responding to the commands and providing user with a desired output like opening wikipedia as shown in figure 3, checking the temperature as shown in figure 4, opening word in figure 5

Figure 3: Opening Wikipedia

Figure 4: Checking the temperature of the day

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

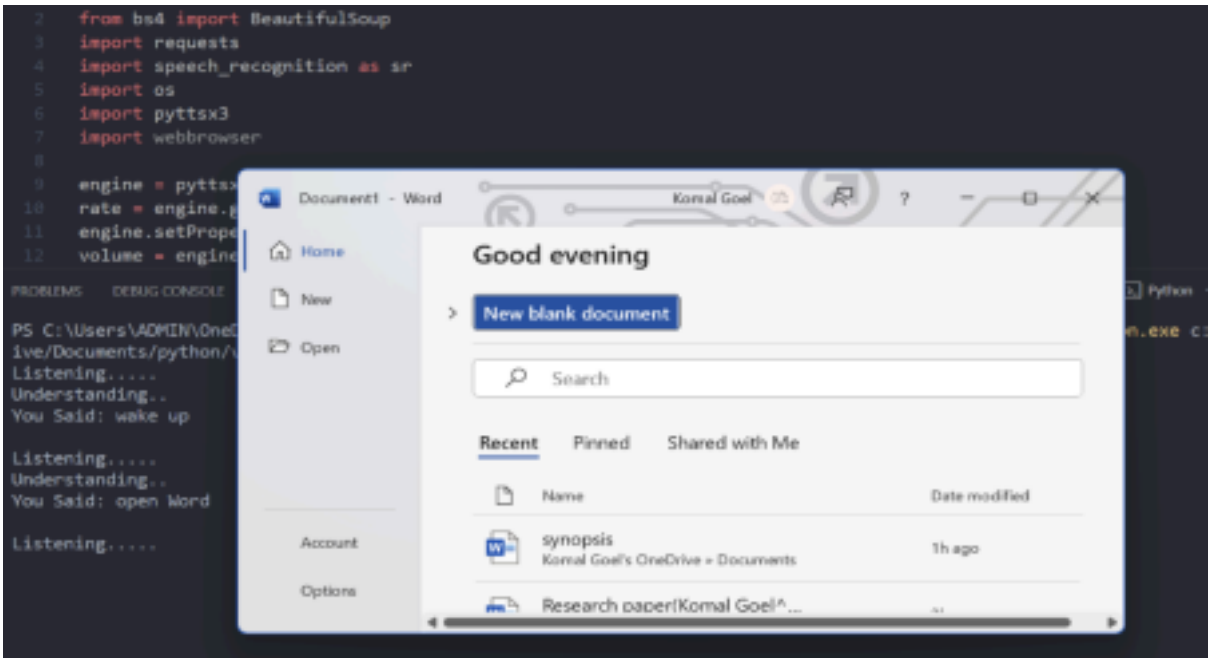

Figure 5: Opening Word

7 Applications of voice assistant

SMART HOME AUTOMATION: Voice assistants used for smart home control have completely changed how we interact with our living areas. Natural language commands enable homeowners to easily handle a variety of smart gadgets through the integration of platforms such as Google Assistant, Apple's Siri, Amazon Alexa, and Google Assistant. Without requiring physical interaction with each device, voice assistants offer seamless control for tasks like locking doors, preparing coffee, adjusting lighting, and setting thermostats. This degree of convenience improves daily activities and encourages home security and energy efficiency. Voice assistants significantly simplify household administration by providing the ability to personalise requests and automate operations. Voice assistants used for smart home control are expected to become increasingly more streamlined and user-friendly as technology develops [9].

HEALTHCARE: In the healthcare sector, voice assistants are helpful tools that enhance patient care, speed up administrative tasks, and boost overall facility productivity. Patients can utilise voice assistants for a variety of purposes, including scheduling appointments, remembering to take their prescriptions, accessing medical information, and even receiving basic first aid instructions. By dictating clinical notes, retrieving patient records, and encouraging staff contact, voice assistants can also support healthcare providers. Voice assistants are useful technologies in the healthcare industry that improve patient care, expedite administrative processes, and increase overall facility productivity. Voice assistants can be used by patients for a variety of tasks, including as making appointments, setting up reminders for prescriptions, retrieving medical records, and even getting basic first aid instructions. Voice assistants can also help healthcare providers by obtaining patient records, dictating clinical notes, and easing staff communication [10].

EDUCATION: Voice assistants are increasingly being used in educational settings because of their innovative ways to enhance learning and the educational process. Teachers, administrators, and students can all benefit from these digital companions in a range of educational contexts. Students gain from voice assistants because they make material easily accessible, assist with homework, and promote interactive learning through games and assessments. Additionally, they can serve as language instructors by providing feedback on pronunciation and assisting with practice. Additionally, voice assistants allow students to access educational resources hands-free, allowing them to concentrate on other activities while listening to lectures, audiobooks, and podcasts. With voice assistants, teachers may automate administrative tasks like scheduling, reminding students of appointments, and making announcements. They can use voice assistants to provide students tailored feedback in addition to developing interactive lessons and distributing instructional materials. Voice assistants can also facilitate class discussions, surveys, and tests to promote students' involvement and active participation. [11]

ENTERTAINMENT: These days, voice assistants are an essential component of the change in entertainment. To increase accessibility and user engagement, they are utilised across a range of platforms, including smart TVs and game consoles. Voice commands allow players to communicate to control characters or do in-game tasks, which enhances immersive gaming experiences. Streaming services employ voice assistants to facilitate hands-free navigation, which lets customers control smart home appliances, look up content, and adjust playback without ever having to touch a button.Voice assistants also enable audience participation during live events and performances, allowing for real-time surveys, trivia, and even crowdsourced storytelling. In general, voice assistants have brought in a new era of entertainment, where our interactions with media and experiences are redefined through seamless integration and simple interactions [12].

NAVIGATION AND TRAVEL: Voice assistants are essential for navigation and travel, providing users with support and convenience at every turn. Voice assistants offer seamless assistance for all tasks, including travel planning, unfamiliar route navigation, and attraction discovery in the vicinity.

When driving, walking, or taking public transportation, users can utilise voice commands to obtain turn-by-turn directions, the quickest routes, and real-time traffic information. Voice assistants also make it simpler for consumers to explore new places by directing them to neighbouring eateries, lodging options, landmarks, and other items of interest. Additionally, passengers can use voice assistants in airports and rail stations to find gate information, check flight or train timetables, and even purchase tickets [13].

8 Privacy and Trust Issues in Voice Assistant

Indeed, there are serious privacy and trust issues with voice assistants. Here are a few crucial points: Data Collection: To increase accuracy and offer individualised experiences, voice assistants frequently gather enormous volumes of data. Although users may not be aware of the extent of data collection or the uses it is put to Security: There's always a chance that someone not authorised could obtain private data that voice assistants have recorded. This could contain private communications, financial information, or other information. Eavesdropping: Even in situations where a wake word is not used, there have been worries expressed regarding voice assistants listening in on private discussions. Despite tech companies' claims that they only record after the wake word, reports have surfaced of unintentional activations and recordings [7].

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

Storage and Retention of Data: Users may not know where or how long their data is stored. When data retention laws are unclear, questions about the long-term privacy implications surface. Access by Third Parties: Integrations with third-party services may expose user data to further risks. It's possible that users have little control over how these third parties share or use their data. User Profiling: Voice assistant companies are able to create thorough user profiles based on user interactions, preferences, and behaviour. While this can enhance personalised experiences, it also raises concerns about user privacy and targeted advertising.Trustworthiness of Providers: Users must have confidence that voice assistant providers will properly handle their data.Users may not understand the full ramifications of using voice assistants or provide their informed permission to data collection and usage practices. To provide informed consent, clear information about privacy policies and data practices is required. To overcome these obstacles, a combination of government oversight, user education, and technology protections is required. For voice assistant technology to be successful in the long run, businesses must prioritise the privacy and trust of their customers [8].

9 Conclusion and Future Work

Voice assistants are now a crucial part of contemporary technology since they make daily tasks easier and enhance user experiences. An OS-transparent voice assistant was created for this assistant using the Python programming language. This AI-powered assistant offers seamless accessibility and saves customers a significant amount of time.Two of the primary advantages of voice assistants are their adaptability and simplicity of use. They are very practical instruments for daily living because they may be used at any moment. These personal desktop assistants, which provide a more hands-free, user-friendly interface that makes tasks like scheduling, reminding, and even managing smart devices easier, have the potential to fundamentally alter how the assistant interacts with technology. Voice assistants are also evolving rapidly. They are now more adaptive because they can speak multiple languages and even identify emotions. These advancements will undoubtedly boost user engagement and offer more customised digital experiences. However, there are still other challenges to overcome. One of the key issues is context awareness, or ensuring the assistant can understand and respond suitably to increasingly complex, nuanced requests. Because voice assistants must accurately recognise and interpret speech from users with a range of linguistic backgrounds, accent identification remains a difficulty.Despite these challenges, they present opportunities for innovation. As technology advances, voice assistants will probably become increasingly more sophisticated in order to meet consumer requests and streamline and simplify digital interactions.

10. Declarations

10.1 Competing Interests

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10.2 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Atharva Singh, Komal Goel, Haider Abbas, Mohd Izhar, Faisal Rais (2025). Voice Assistants: The Digital Age's Emergence of Virtual Friends. *AIJR Proceedings*, 159-168. https://doi.org/10.21467/proceedings.178.19

References

[1] A. Poushneh, "Humanising voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviours," *Journal of Retailing and Consumer Services*, vol. 58, p. 102283, 2021.

[2] C. Flavián, K. Akdim, and L. V. Casaló, "Effects of voice assistant recommendations on consumer behaviour," *Psychology & Marketing*, vol. 40, no. 2, pp. 328–346, 2023.

[3] M. B. Hoy, "Alexa, Siri, Cortana, and more: An introduction to voice assistants," *Medical Reference Services Quarterly*, vol. 37, no. 1, pp. 81–88, 2018.

[4] F. Shahid, *Siri or Google Assistant? The impact of voice assistant anthropomorphism on consumer usage intention*. Macquarie University, 2022.

[5] A. J. Abougarair, M. K. Aburakhis, and M. Zaroug, "Design and implementation of smart voice assistant and recognising academic words," *International Robotics & Automation Journal*, vol. 8, no. 1, pp. 27–32, 2022.

[6] J. Austerjost et al., "Introducing a virtual assistant to the lab: A voice user interface for the intuitive control of laboratory instruments," *SLAS Technology: Translating Life Sciences Innovation*, vol. 23, no. 5, pp. 476–482, 2018.

[7] T. Bolton et al., "On the security and privacy challenges of virtual assistants," *Sensors*, vol. 21, no. 7, p. 2312, 2021.

[8] D. Pal et al., "To trust or not-trust: Privacy issues with voice assistants," *IT Professional*, vol. 22, no. 5, pp. 46–53, 2020.

[9] S. Venkatraman, A. Overmars, and M. Thong, "Smart home automation—Use cases of a secure and integrated voice-control system," *Systems*, vol. 9, no. 4, p. 77, 2021.

[10] M. Eggert and M. A. Stanke, "Adoption of integrated voice assistants in health care-requirements and design guidelines," *Wirtschaftsinformatik (Zentrale Tracks)*, 2020.

[11] T. Tsourakas, G. Terzopoulos, and S. Goumas, "Educational use of voice assistants and smart speakers," *Journal of Engineering Science & Technology Review*, vol. 14, no. 4, 2021.

[12] M. Gordon and C. Breazeal, "Designing a virtual assistant for in-car child entertainment," in *Proc. 14th Int. Conf. Interaction Design and Children*, 2015.

[13] J. Silviya Nancy et al., "Context-aware self-learning voice assistant for smart navigation with contextual LSTM," in *Adv. Informatics for Computing Res.: 3rd Int. Conf. ICAICR 2019, Shimla, India, Revised Selected Papers, Part I*, vol. 3. Singapore: Springer, 2019.

[14] R. Mehrotra et al., "Hey Cortana! Exploring the use cases of a desktop-based digital assistant," in *Proc. SIGIR 1st Int. Workshop Conversational Approaches to Information Retrieval (CAIR '17)*, vol. 4, 2017.

[15] A. H. Michaely et al., "Keyword spotting for Google Assistant using contextual speech recognition," in *Proc. 2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2017.

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922