Assessing the Development and Use of Virtual Personal Assistants

[1]Akhila Mohanan, [2]Anjana Sabu ,[3]Sandhra Susan Shinu ,[4]Swathi Surendran ,[5]Gargi ChandraBabu
[1,2,3,4]Student,[5]Assistant Professor
[1,2,3,4,5]Department of Computer Science and Engineering
[1,2,3,4,5]College Of Engineering,Kidangoor,Kottayam, Kerala, India
[1]akhilamohanan009@gmail.com , [2]anjanasabu25@gmail.com , [3]sandhrasusanshinu@gmail.com ,
[4]swatheendra70@gmail.com , [5]gargichandrababu@gmail.com

ABSTRACT: Intelligent personal assistants (IPAs) and virtual personal assistants (VPAs) have rapidly advanced in recent years, transforming how users interact with digital technology in both personal and professional domains. These systems, powered by artificial intelligence and natural language processing, are designed to streamline tasks, improve productivity, and enhance accessibility, enabling users to manage their daily schedules, answer questions, and perform routine functions through voice or text commands. Research has highlighted their increasing integration into various platforms, from mobile devices and smart speakers to specialized operating systems, which expand the functional scope of VPAs. Studies show that user satisfaction and interaction effectiveness are influenced by the degree of personalization, with advanced user modeling techniques enabling these systems to adapt responses based on individual preferences and usage patterns. However, alongside these advances, security and privacy issues are becoming more prominent, particularly with third-party applications and voice-controlled functions that can expose user data to unauthorized access or misuse. Addressing these vulnerabilities is essential to fostering trust and encouraging widespread adoption of these technologies. Additionally, the role of accessibility is crucial, as VPAs should accommodate diverse user needs, including those with varying levels of technical expertise, different languages, and specific accessibility requirements. Real-world implementations of AI-powered assistants, such as those tailored for specific operating systems, demonstrate the versatility of VPAs but also emphasize the need for continuous improvement in usability, responsiveness, and contextual awareness to deliver a seamless and natural user experience. Ultimately, the future of VPAs lies in overcoming these challenges by balancing advanced capabilities with stringent security protocols, ethical considerations, and inclusive design, paving the way for a generation of digital assistants that are more intelligent, secure, and user-centered.

INDEX TERMS: Virtual Personal Assistant (VPA),Artificial Intelligence (AI),Natural Language Processing (NLP),Machine Learning (ML), Automation,

1 Introduction

In recent years, Virtual Assistants (VAs) have become an integral part of everyday life, revolutionizing the way users interact with technology. These intelligent systems help individuals accomplish tasks by leveraging multiple communication mediums such as text, speech, images, and video. With the rapid advancements in Artificial Intelligence (AI) and Machine Learning (ML), VAs are increasingly capable of understanding and responding to user queries, making decisions, and performing tasks more efficiently. These assistants utilize various technologies like Natural Language Processing (NLP), Human-Computer Interaction (HCI), and different recognition systems (speech, image, video) to facilitate seamless human-machine interaction.

The role of Machine Learning is pivotal in the development of VAs, as it allows these systems to learn from past user behavior, predict future actions, and recommend personalized solutions. Through techniques like Clustering, Generalization, Artificial Neural Networks, and Support Vector Machines, VAs can adapt and improve their responses over time, reducing the burden of manual tasks and increasing overall productivity. This paper presents an overview of Virtual Assistants, highlighting their integration with Machine Learning methods, discussing their applications, and examining the technologies that make them effective tools in various domains, including education and professional settings.

Transfer learning is not a new concept and it isn't particularly related to deep learning. Making use of a methodology based on transfer learning principles makes a lot of difference from using traditional machine learning techniques. In traditional learning, tasks, datasets, and models are used to create specific isolated learning models based on them and no knowledge is retained which can be transferred from one model to another. Transfer learning enables you to reuse previously trained knowledge (features, weights, etc) for training newer models and even tackle situations where there is less data.

2 LITERATURE SURVEY

Veton Z. Këpuska et al. [9] introduce a framework for the Next Generation of Virtual Personal Assistants (VPAs), which is a cutting-edge system designed to facilitate conversations with humans using advanced Machine Learning models like Deep Neural Networks. These VPAs integrate multiple communication channels, such as speech, visuals, video, and gestures, for both input and output. They are employed in diverse applications, including customer interaction, service support, education and training, transaction facilitation, information retrieval, and taxi reservations. To enhance the system's accuracy and complete its development, the project requires collaboration and financial support from partnering organizations.

Katragadda et al. [9] describe a system that provides a unified platform for performing various tasks without manual intervention. The system relies on Support Vector Machines and Natural Language Processing techniques, featuring functionalities like Automatic Profile Management, Location-Based Reminders, and Calling Services. These features enable efficient task tracking and management, such as setting reminders based on time, location, or calls.

Ross Mead et al. [9] designed a cloud-based system, supported by Semio, that allows robots to interpret and respond to human communication through speech and body language. Semio equips developers with tools to create an operating system, an app ecosystem, and browser-based software for crafting, animating, and deploying robot applications using natural language. Developers can use the Semio SDK to build and deploy gesture- and speech-driven robot applications on the cloud, allowing non-experts to access and use them via natural interaction. The platform enables robots to automatically process speech and gestures, make decisions, and respond appropriately.

Tom Mitchell et al. [9] introduced Calendar Apprentice (CAP), a practical tool capable of autonomously learning a user's meeting preferences based on prior interactions. CAP allows users to review and modify the rules it learns, as shown in Fig. 1, ensuring the system adapts effectively. Users can also override outdated or inefficient rules as needed. By leveraging past experiences, CAP reduces the user's workload by efficiently managing tasks and delivering highly accurate predictions in most scenarios. Additionally, when a user's schedule is busy, the system adapts and predicts outcomes to provide suitable recommendations.

The Hidden Markov Model (HMM) is a statistical model used for systems where the actual states are hidden, but we can observe certain events that help infer these statesin the Fig.2. HMMs are highly applicable in fields like speech recognition, bioinformatics, and natural language processing, where data often comes in sequences. In HMMs, the hidden states represent the underlying structure of the system, such as phonemes in spoken language or parts of a gene, which are not directly visible. However, we can observe outputs—such as audio signals in speech recognition or DNA bases in gene sequencing—that provide indirect clues about these states. Each state has associated transition probabilities, representing the likelihood of moving from one state to another. Additionally, each hidden state has an emission probability, which links it to the probability of generating a specific observation. The initial state

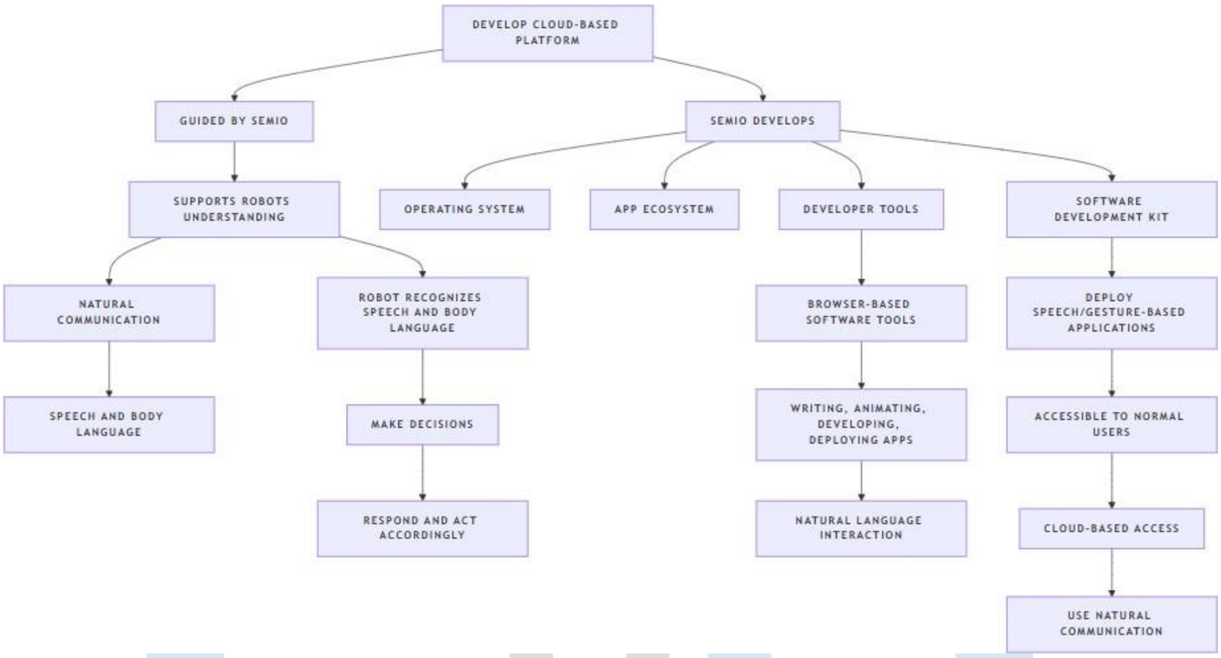

Figure 1: Semio Cloud Platform: Enabling Robots to Understand and Respond to Human Communication".

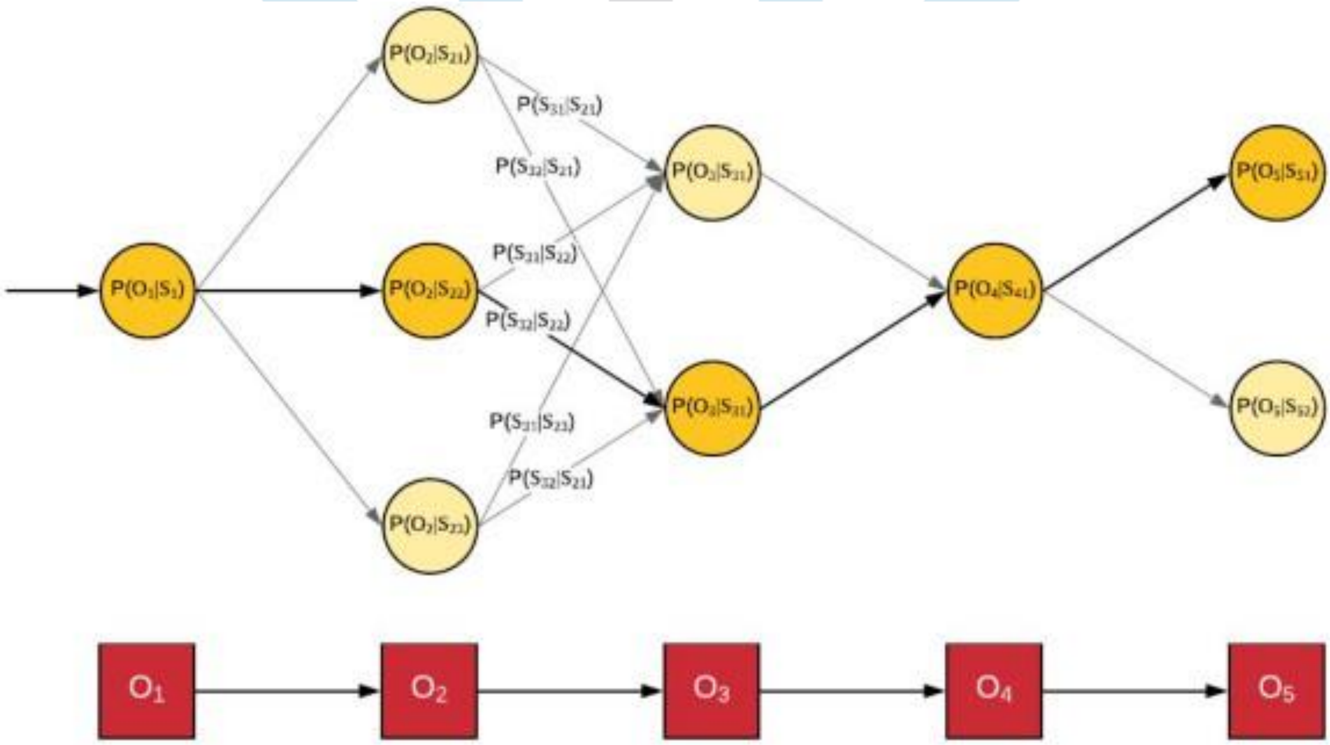

Figure 2: Hidden Markov Model . [9]

distribution sets the starting probabilities for each hidden state in the sequence.

Zakharova (Zakharova, 2018) et .al [3] has created a pro fessional development forecast model. She used ma chine learning algorithms to analyze students' papers. Pre-processing and statistical analysis of texts were carried out to highlight features characterizing the general vocabulary and the use of general scientific and professional terminology. These approaches al lows students to get immediate

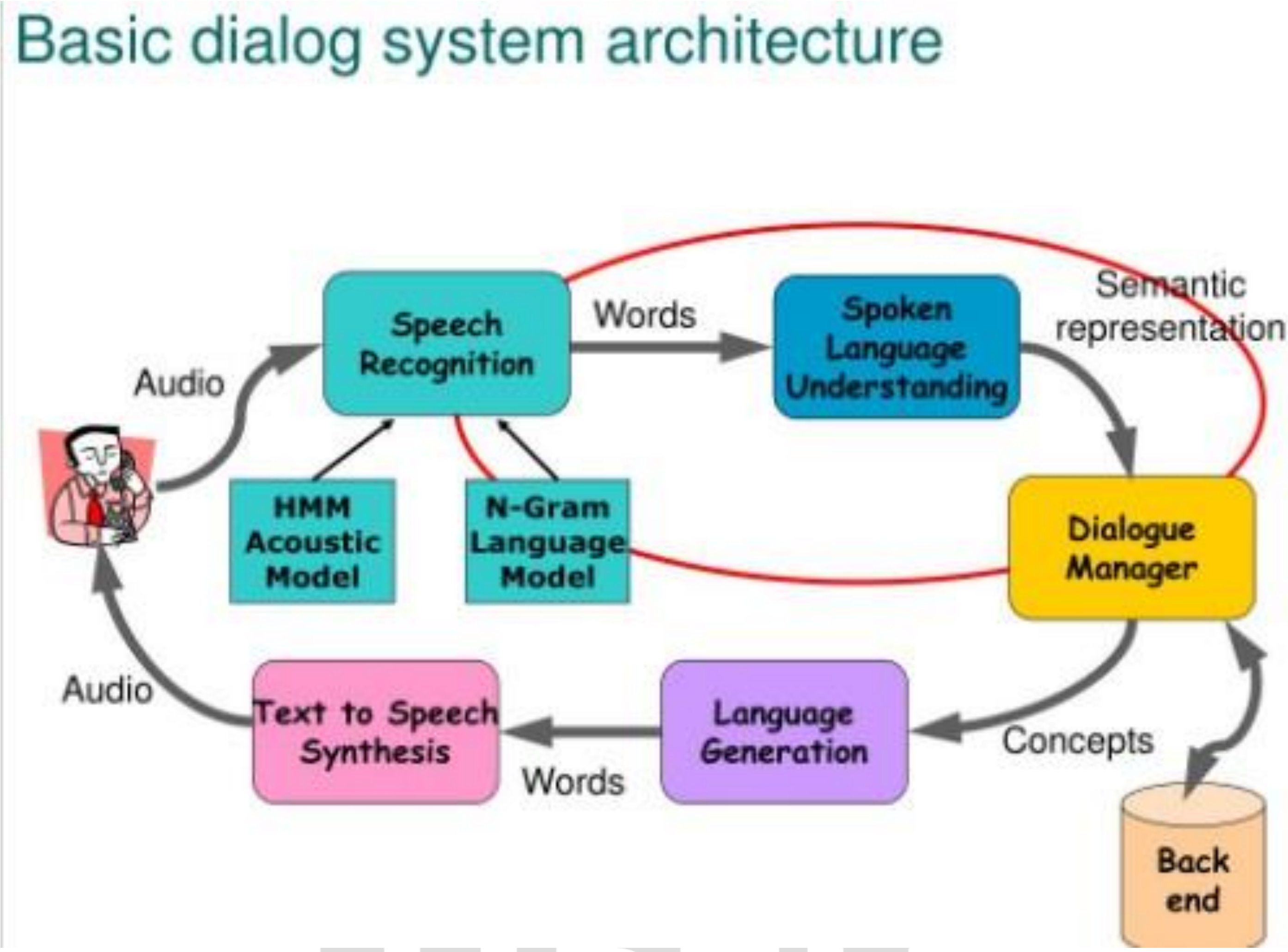

Figure 3: . Categories of support infrastructure [?]

feedback, contribut ing to better learning and academic success. In fact, this literature review further confirms that the effec tiveness of a virtual assistant depends on the quality of feedback. There are three kinds of feedback elements given by the virtual assistant: evaluation, disturbances and hints a basic dialog architecture in Fig.3. Disturbances are questions that are intended to test the confidence of the students about their work. Hints are documents (text, graphs, formulas, etc.) re lated to a particular question of the experimental pro tocol that help students find by themselves possible errors or problems (Geoffroy et al., 2002).

Kiseleva et al. [2]found that users tend to be more satisfied with using VPAs for simple tasks (such as device control) rather than more complex, multi-turn tasks (e.g. making travel arrangements etc.) where preserving the context is crucial. Despite being comprehensive, the study was limited to one type of virtual assistant (Cortana). Our survey adopts a broader perspective by probing users satisfaction with a variety of VPAs. Cowan et al. [2] focused on the experiences of infrequent VPA users and their reasons for not using VPAs on a regular basis. The feedback obtained from the focus groups indicates that privacy concerns over data usage, and lack of trust in the assistant's capability to perform the task are some of the main reasons for people not to use the technology on a regular basis. Although Cowan et al.'s study was limited only to infrequent users of Siri, it can be argued that data permanency, ownership of data and limited human-like interaction abilities are the factors that are relevant across different devices. Thus, it would be worth extending the scope of investigation to other assistants. Our study applies a broad approach towards evaluation of VPAs by considering the feedback of different user groups

In a similar study, conducted by Sorensen et al. [2], expecta tions of novice users were compared between two different chatbot systems (i.e. a human-like and a robot-like system). The results indicated that the system that asked questions, pro vided feedback and informed user about its current status was perceived as better meeting users performance expectations. While Sorensen's study dealt exclusively with text as the input method, in our work we focus on voice interaction. Moore et al. [2] analysed the opinions of members of the public expressed in two surveys on spoken language technology. One that compared the opinions of experts and non-experts, and another that evaluated the degree of usage of voice technology on a daily basis. The overall results suggest that the ordinary people (non-experts) are more optimistic about the future capabilities of voice technology. However, poor system accuracy and inadequate language understanding skills prevent regular usage. In current work, we ask our respondents about their perception of their VPA's current language capabilities, and expectations regarding future natural language and conversational capabilities. . Another study examining privacy and social acceptability was conducted by Efthymiou and Halvey [2] in the context of voice based smartwatch search. The findings indicated that voice search has low social acceptability when carried out in front of strangers, mostly due to privacy concerns. Our work builds up on previous studies by incorporating questions on respondents' con cerns over privacy and social acceptability of VPAs when used in private and public spaces. In this paper, we replicate, collate and extend the prior work on VPAs to get better understanding of current usage and fac tors that drive VPA adoption. Our focus is to determine how voice interaction and VPA usage vary between frequent and in frequent user groups. Park et al. [7] also used the TAM and confirmed that people considered platform-wise benefits more important than device-wise benefits. Consumers perceived smart speakers as an outlet connecting home automation. As these early studies share a common focus on users' motivation to purchase smart speakers, they can help us understand the mechanism of the adoption process. Apart from the adoption reasons, there is little information on the psychological factors and interaction related factors affecting users' satisfaction while interacting with VPAs in daily routines. There are studies exploring smart speaker users' interactions with smart speakers and VPAs. The Alexa users' daily usage behavior and interac tion pattern were observed more comprehensively using both history logs and in-home contextual interviews (Sciuto et al., 2018). Lopatovska and colleagues (2018) also analyzed user interaction using Alexa users' online daily diary responses. Meanwhile, in the conext of human-computer interaction, parasocial interaction could occur in situations where human VPA communication takes place, as discussed by Han and Yang (2018), who focused on the parasocial relationship the ory. Purington and her colleagues (2017) analyzed user reviews of the Amazon Echo and the Alexa, and found that people personified the device and agent. Another user review analysis revealed this personification phenomenon along with usage behavior patterns (Gao et al., 2018). A study discussed how the elderly perceived the voice assistant of Echo in terms of anthropomorphism, and its link to their loneliness (Pradhan et al., 2019). Few studies have reported the possible relationship between IPA use and loneliness. All of these qualitative approaches and quantitative approaches using review data revealed usage patterns in general and interesting phenomena such as personification in VPA usage experience. However, these concepts have not yet been examined statistically using responses from users. Therefore, the current study focused on the effects of interactions with VPAs.

Ramanathan G et al. [4] highlight the significant scaling challenges involved in developing a modeling system for a personal assistant in production. The first observation is that computing features such as categories, forums, and entities dynamically during query and click processing is prohibitively expensive. To address this, they pre-compute these com putationally expensive features for a few billion frequently occurring clicks and queries, storing them for runtime lookups. For less common (tail) queries and clicks, they fall back on simpler and cheaper features such as query words and click titles. Consider a batch implementation that partitions input data by user and parallelizes computation using MapReduce [8]. Running this system for over 100 million users, with approx imately 5 seconds of computation per user (as discussed in Section 4.3.4), would result in around 6,000 days of CPU time just for segmentation—excluding the additional load on the back-end for feature lookups. One potential optimization involves skipping users with no new data since the last run. However, since active users tend to generate significantly more data than inactive ones, this optimization yields only a marginal reduction in overall com putation time. Consequently, any system that relies on repeated batch processing will face substantial latencies between the arrival of new observations and updates to the user model.

Lucchese et al. [4] proposed techniques for identifying task based sessions : sets of possibly noncontiguous queries within a session that correspond to a user task. They recog nized that standard clustering algorithms are too expensive, and proposed a new clustering algorithm QCHTC, which avoids computing the full similarity graph by just consid ering a cluster to be represented by the rst and the last queries in the session. Li et al. [4] also consider in-session tasks. They use query words, query co-occurrence, and the temporal sequence of queries as their main signals. Their learning algorithm is quadratic, but they have a linear approximation which works well in practice. Hua et al. [10] showed the importance of semantic features for in-session task identi cation.

Wangetal et al. propose an algorithm for identifying cross session search tasks using a variety of signals. They built a training set by completely annotating 5 days of query logs for 1436 users, and learned a latent structural SVM classi er to compute the similarity between pairs of queries. Then they use a bestlink clustering algorithm to cluster queries each query is clustered with the most similar query that occurred before it, as long as their similarity exceeded a threshold

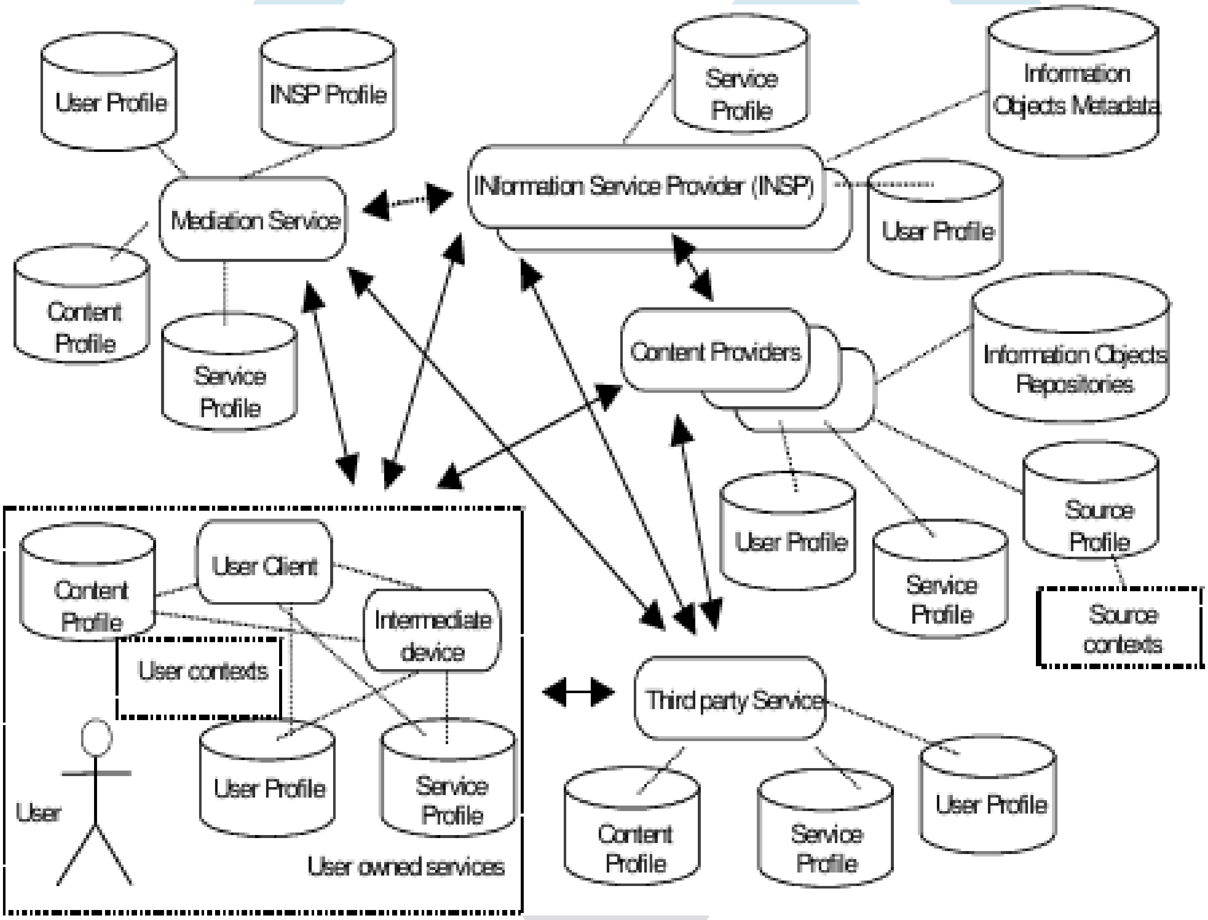

Figure 4: . Categories of support infrastructure [8]

imrie et al. [6] drew significant attention to the issues surrounding personal data and the fragile boundaries between private and public systems. Increasingly, questions arise about who has access to data generated by careless social media posts or the monitoring of supposedly private emails and telephone conversations. As data storage continues to shift toward ”the Cloud,” concerns about privacy and security have intensified. In this context, the concept of a Virtual Personal Assistant (VPA) becomes particularly appealing, as it refocuses supporting systems within the user's private control. The authors proposed a hypothesis that a chatbot, enhanced with natural language processing algorithms and localized data, could be trained to function as a VPA with user controlled data privacy. Such a system would require the ability to create metadata by linking the data it receives and producing contextual outputs during interactions that are meaningful and useful to the user (see Fig. 4)

Abishek Narayanan et al. [5] conducted an exploration of the ”Virtual Personal Assistant” (VPA), focusing on foun dational techniques in context-aware computing using Mel Frequency Cepstral Co-efficients (MFCC) and Natural Lan guage Processing (NLP). To enhance spoken word recognition accuracy, the system integrates neural network-based speech recognition with machine learning-driven lip movement detec tion. The proposed VPA system comprises four key components: voice input, detection, deduction, execution, and output. Fur thermore, Narayanan et al. introduced a ”Virtual Personal Assistant with Facial Recognition Login System” to enhance security through a two-phase login process. The first phase involves a traditional username and password login, followed by a second phase employing user-specific facial recognition logic for added security.

Rodriguez et al.[10] explored the idea that speech acoustic characteristics can serve as indicators of trust and familiarity among team members. The speech data used in this study was collected by the Air Force Research Laboratory. Individual voice recordings were divided into 1-second intervals and converted into RGB amplitude spectrogram images. These spectrograms were then classified using a fine-tuned CNN ResNet-18 model to identify either the level of trust or familiarity. Two trust-related constructs were assessed using survey data: TTP (Trustworthiness, Propensity to Trust) and RIS (Reliance Intentions Scale). Speech samples were classified into Low/High-trust categories for the RIS and TTP constructs, both before and after the mission, using single transfer learning (TL), achieving an average classification accuracy of 82 A separate single TL classification into Known/Unknown partners categories achieved an accuracy of 85 The proposed training and testing framework for the study is depicted in Fig. 10.

Guha et al. et al. [1], explain that Virtual Personal Assistants (VPAs) rely on advanced techniques in machine learning and user modeling, allowing them to adapt and respond to individual user behaviors and preferences over time. These developments enable VPAs to deliver highly personalized services, such as customized recommendations, reminders, and even proactive assistance. While VPAs were initially designed for straightforward tasks like setting alarms or checking the weather, their applications have expanded significantly into do mains such as smart homes, healthcare, and customer service. According to Gubareva and Lopes (2020), VPAs have found a valuable role in education, where they support stu dents by facilitating learning, giving feedback, and handling administrative duties. In the healthcare sector, VPAs help manage appointments, remind users about medications, and even provide initial health-related suggestions. However, Jang (2020) highlights that achieving high levels of user satisfaction is still a hurdle. Even with improvements in voice recognition and natural language processing (NLP), VPAs often encounter challenges in interpreting complex or ambiguous commands, which can lead to user dissatisfaction. Privacy concerns are another significant issue. Since VPAs require access to personal information to function effectively, users often worry about data security and misuse. Imrie and Bednar (2013) point out that although advancements in encryption and data handling practices aim to address these concerns, many users remain wary of the extent to which their data is collected and utilized.

to Dubiel et al. [11], Virtual Personal Assistants (VPAs) leverage advanced technologies like Natural Language Pro cessing (NLP), Speech Recognition, and Machine Learn ing to comprehend user commands and adapt to individual preferences. These innovations have evolved VPAs from basic command-response systems into intelligent, context-aware tools capable of learning from interactions. Gubareva and Lopes (2020) emphasize the increasing role of VPAs in educa tion, where they enhance learning by delivering personalized assistance, organizing schedules, and providing resources for both students and educators. Similarly, Katragadda (2023) discusses how machine learning-powered VPAs and chatbots are transforming customer service by automating repetitive tasks and increasing operational efficiency. Despite these advancements, challenges remain. Jang (2020) observes that VPAs still struggle with understanding com plex, multi-turn conversations and queries requiring contex tual interpretation. Privacy concerns also persist, as many users remain cautious about the extent of data collection by these systems, a point highlighted by Imrie and Bednar (2013). As VPAs advance, future innovations are anticipated to focus on overcoming these challenges by offering more secure, personalized, and seamless experiences across different platforms. However, addressing issues like privacy, contextual understanding, and support for multiple languages will be crucial to ensuring their broader adoption and success in the years ahead.

Manasa Sri Vardhan Kottamasu and C. K. Gomathy et al.[10] along with other researchers, have outlined the creation of a Python-based voice-enabled assistant designed for personal computer use. This system leverages APIs to perform various tasks, including performing calculations and retrieving news quotes from websites. Users can interact with the assistant by issuing requests and obtaining specific results. This voice as sistant allows users to ask questions in a conversational format and offers features such as opening applications, reading news, taking notes, and conducting Google searches. Additionally, the system utilizes artificial intelligence and natural language processing techniques to provide a variety of services, includ ing sending emails, performing web searches, setting alarms, delivering weather updates, and playing YouTube videos. The tool operates using three learning modes: supervised learning, unsupervised learning, and reinforcement learning. The inte gration of machine learning and deep learning is essential to carry out the wide range of tasks effectively.

3 COMPARATIVE STUDY

Virtual personal assistants (VPAs) have evolved significantly over the years, with advancements in user interaction, application areas, and underlying technologies. Studies such as Jang (2020) and Dubiel et al. (2018) focus on user satisfaction and interaction, noting that ease of use and accuracy in understanding user commands are key factors. Jang's research emphasizes the role of smart speakers, where voice clarity and the assistant's ability to understand diverse accents are essential for user experience, a limitation also highlighted by Dubiel et al. (2018). In terms of applications, Gubareva and Lopes (2020) explore the use of VPAs in educational settings, showing how these assistants help with scheduling and personalized learning. This contrasts with Katragadda (2023), who examines the use of machine learning-driven chatbots in automating customer support, where VPAs are focused on task efficiency rather than complex interactions. Technologically, Guha et al. (2015) delve into user modeling, demonstrating the importance of creating personalized systems that adapt over time, a key difference from earlier, simpler VPAs discussed by Imrie and Bednar (2013), which were limited to basic functionalities like scheduling. Despite these advancements, challenges remain, particularly in adapting VPAs to dynamic user needs, as highlighted by both Gubareva and Lopes (2020) and Katragadda (2023). These findings reflect the rapid progression from early VPAs with limited capabilities to more sophisticated, AI-driven assistants that can learn from and respond to user behavior in more personalized ways. However, achieving consistent user satisfaction and understanding complex contexts continues to be a challenge. This comparative study underscores the diverse applications of VPAs across domains and the ongoing need for improvements in natural language processing, machine learning, and user modeling.

4 CONCLUSION

In conclusion, this comparative study of recent literature on intelligent personal assistants (IPAs) and virtual personal assistants (VPAs) reveals distinct trends and challenges in the field, as summarized in the above Table1 . Each study offers valuable insights into different facets of VPA functionality, user experience, and technical challenges. For instance, several studies (e.g., Papers 1, 2, and 4) underscore the importance of advanced user modeling and personalization techniques, which can enhance user satisfaction and interaction by tailoring responses to individual preferences.

Additionally, security concerns are prevalent, with Papers 4 and 10 particularly emphasizing vulnerabilities introduced by third-party integrations, underscoring the need for enhanced data protection measures. In terms of application, Papers 3 and 8 explore the growing role of VPAs in educational and customer support environments, suggesting that these tools can increase accessibility and efficiency across diverse fields. Moreover, studies such as Papers 5 and 9 illustrate technical implementations and practical applications of VPAs, showcasing the feasibility of deploying VPAs in specific operating systems and highlighting the ongoing need for refinement in usability and functionality.

Overall, this comparative analysis points to a promising future for VPAs, provided that advancements in security, personalization, and cross-functional applications continue to address the evolving demands and risks in user interaction and digital assistance.

	Deep Learning	Support Vector Machine	Decision Tree
[1]		✓	
[2]		✓	
[3]			✓
[4]		✓	
[5]			✓
[6]	✓		
[7]	✓		
[8]	✓		

Figure 5: Technologies used

References

[1] Allan de Barcelos Silva, Marcio Miguel Gomes, Cristiano André da Costa, Rodrigo da Rosa Righi, Jorge Luis Victoria Barbosa, Gustavo Pessin, Geert De Doncker, and Gustavo Federizzi. Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications*, 147:113193, 2020.

[2] Mateusz Dubiel, Martin Halvey, and Leif Azzopardi. A survey investigating usage of virtual personal assistants. *arXiv preprint arXiv:1807.04606*, 2018.

[3] Regina Gubareva and Rui Pedro Lopes. Virtual assistants for learning: A systematic literature review. *CSEDU (1)*, pages 97–103, 2020.

[4] Ramanathan Guha, Vineet Gupta, Vivek Raghunathan, and Ramakrishnan Srikant. User modeling for a personal assistant. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 275–284, 2015.

[5] SVS Gunasekara, B Fernando, et al. Gypsy: Ai-powered virtual assistant for windows os. 2024.

[6] Peter Imrie and Peter Bednar. Virtual personal assistant. In *ItAIS 2013. The 10th Conference of the Italian Chapter of AIS*. Università Commerciale Luigi Bocconi in Milan, Italy,, 2013.

[7] Yeibeech Jang. Exploring user interaction and satisfaction with virtual personal assistant usage through smart speakers. *Archives of Design Research*, 33(3):127–135, 2020.

[8] Abdul Rehman Javed, Muhammad Usman Sarwar, Saif ur Rehman, Habib Ullah Khan, Yasser D Al-Otaibi, and Waleed S Alnumay. Pp-spa: privacy preserved smartphone-based personal assistant to improve routine life functioning of cognitive impaired individuals. *Neural Processing Letters*, pages 1–18, 2023.

[9] VAMSI KATRAGADDA. Automating customer support: A study on the efficacy of machine learning-driven chatbots and virtual assistants. *IRE Journals*, 7(1):600–601, 2023.

[10] SI Parihar, Neha Tarale, Ruchika Kumbhare, Chaitanya Kumbhalkar, and Kartik Daine. Voice assistant using python and ai.

[11] Nan Zhang, Xianghang Mi, Xuan Feng, XiaoFeng Wang, Yuan Tian, and Feng Qian. Dangerous skills: Understanding and mitigating security risks of voice-controlled third-party functions on virtual personal assistant systems. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1381–1396. IEEE, 2019.