AMERICAN INTERNATIONAL UNIVERSITY-BANGLADESH
Faculty of Science and Technology

Assignment Cover Sheet

Assignment Title:	**Misinterpretation, Privacy and Security Challenges in Artificial Virtual Assistant**			
Assignment Type:	Research Proposal		Date of Submission:	19 December 2023
Course Title:	RESEARCH METHODOLOGY			
Course Code:	CSC 4197		Section:	B
Semester:	Fall	2023-24	Course Teacher:	DR.MD.ABDULLAH-AL-JUBAIR

Declaration and Statement of Authorship:

1. I/we hold a copy of this Assignment/Case-Study, which can be produced if the original is lost/damaged.
2. This Assignment/Case-Study is my/our original work and no part of it has been copied from any other student's work or from any other source except where due acknowledgement is made.
3. No part of this Assignment/Case-Study has been written for me/us by any other person except where such collaboration has been authorized by the concerned teacher and is clearly acknowledged in the assignment.
4. I/we have not previously submitted or currently submitting this work for any other course/unit.
5. This work may be reproduced, communicated, compared and archived for the purpose of detecting plagiarism.
6. I/we give permission for a copy of my/our marked work to be retained by the Faculty for review and comparison, including review by external examiners.
7. I/we understand that Plagiarism is the presentation of the work, idea or creation of another person as though it is your own. It is a formofcheatingandisaveryseriousacademicoffencethatmayleadtoexpulsionfromtheUniversity. Plagiarized material can be drawn from, and presented in, written, graphic and visual form, including electronic data, and oral presentations. Plagiarism occurs when the origin of them arterial used is not appropriately cited.
8. I/we also understand that enabling plagiarism is the act of assisting or allowing another person to plagiarize or to copy my/our work.

* *Student(s) must complete all details except the faculty use part.*
** Please submit all assignments to your course teacher or the office of the concerned teacher.

Group Name/No.:

No	Name	ID	Program	Signature
1	NOSHIN FARZANA	21-44647-1	BSc [CSE]	
2	AVIJIT SAHA ANTO	21-44630-1	BSc [CSE]	
3	SADIKUL MOBASSHIR	21-44415-1	BSc [CSE]	
4	ISRAK HOSSAIN PANTHO	21-44401-1	BSc [CSE]	
5	KHONDAKER FAISAL IBN AZIZ	21-44398-1	BSc [CSE]	

Faculty use only		
FACULTY COMMENTS	**Marks Obtained**	
	Total Marks	

Research Proposal

MISINTERPRETATION, PRIVACY AND SECURITY CHALLENGES IN ARTIFICIAL VIRTUAL ASSISTANT

INTRODUCTION

Artificial Virtual Assistants have emerged as innovative solutions designed to improve human-machine interactions in the rapidly changing technological landscape. It is a software agent that can perform a variety of tasks or services for a user, when given commands or questions, including verbal ones. These Artificial Virtual Assistants are incredibly versatile. They can help with a lot of different things, like managing daily reminders, scheduling appointments, answering questions, and even doing data analysis. By integrating into smart speakers, smartphones, or other gadgets, they are becoming essential components of our everyday lives. They make tasks easier and provide helpful assistance in both the personal and professional domains. They can make technology more accessible to people with disabilities, such as those with visual impairments, by providing voice-activated interfaces and assistive capabilities. Amazon Alexa, Apple Siri, Google Assistant, and Microsoft Cortana are examples of well-known AI virtual assistants. Amazon Alexa is used in Amazon Echo devices, which can control smart home devices, answer questions, and perform various tasks through voice commands. Apple's built-in personal virtual assistant, Siri, can be voice-activated and used on devices running iOS, iPad, mac. It uses voice recognition technology to perform tasks. Google Assistant is integrated into android devices and available on iOS, offers information retrieval, task automation, and voice-controlled functions. Microsoft Cortana is Microsoft's virtual assistant, which is used in Windows and other Microsoft products. It uses the Bing search engine to perform tasks like setting reminders, sending emails, and answering questions. Using artificial virtual assistants can offer convenience and functionality, but they also come with several potential problems and challenges. As AI virtual assistants have become more integrated into daily life, there is a growing need to address ethical concerns related to privacy, security, and transparency as well as improving the user experience. Understanding these problems is crucial to enhancing the performance and usability of these virtual assistants and ensuring that they meet the evolving needs and expectations of users.

BACKGROUND

The roots of Artificial Virtual Assistants can be traced back to the intersection of Artificial Intelligence and Human-Computer interaction. Although the idea of creating intelligent entities to help humans has been around since the early days of computer science, but the major advancements in the field have only recently resulted in the creation of sophisticated Virtual Assistants.

In the early stages, basic rule-based systems were employed to simulate conversation, enabling simple interactions. However, the true breakthrough came with the advent of Natural Language Processing (NLP) and Machine Learning (ML). These technologies enabled virtual assistants to not only understand and interpret human language but also to adapt and improve based on user interactions. As computing power increased and data became more abundant, the capabilities of Artificial Virtual Assistants expanded exponentially. The integration of cloud computing allowed these assistants to access vast amounts of information in real-time, providing users with dynamic and up-to-date responses. The rise of smartphones and smart devices further accelerated the adoption of virtual assistants, making them inescapable in various aspects of daily life. Major technology companies have played a significant role in shaping the landscape of Artificial Virtual Assistants, introducing flagship virtual assistants that set new standards for functionality and user experience. Voice-activated assistants like Siri, Google Assistant, and Alexa have become household names, demonstrating the potential of artificial intelligence to streamline tasks and enhance convenience. They are designed to provide users with a seamless and personalized experience across a variety of tasks and applications.

As the field continues to evolve, the combination of AI, machine learning, and other emerging technologies holds the promise of creating even more intelligent, context-aware, and emotionally intelligent virtual assistants. The background of Artificial Virtual Assistants reflects a journey from basic rule-based systems to sophisticated, learning entities that are becoming integral parts of the digital ecosystem, anticipating, and meeting the diverse needs of users across the globe.

PROBLEM STATEMENT

The widespread adoption of AI virtual assistants, such as Google Assistant, Siri, Alexa, and Cortona has brought about numerous benefits in terms of convenience and automation. In an increasingly digitized world, they have become integral parts of our daily lives, assisting us with tasks ranging from answering questions to controlling smart home devices. It is estimated that by 2026 more than 150 million voice assistants will be used **[1]**. However, despite their advantages, users encounter a range of challenges and issues while interacting with these AI virtual assistants. AI virtual assistants face several challenges, including natural language understanding limitations that can lead to misunderstandings and misinterpretations, privacy concerns regarding the storage and use of user data, the potential for bias in decision-making due to biased training data, difficulty in handling complex or nuanced tasks that require human intuition, the risk of job displacement in industries where virtual assistants are increasingly used, and the need for ongoing updates and improvements to keep up with evolving user expectations and technology advancements. Additionally, ensuring the ethical and responsible deployment of AI virtual assistants remains a critical challenge, requiring careful consideration of their impact on society and individuals. This paper explores the misinterpretation, privacy, and security challenges of artificial virtual assistants. The voice recognition mechanism can typically distinguish sounds from others, but the mechanism is far from flawless. Most of the time it doesn't seem perfect. These problems can decrease their functionality, usability, and overall user experience. The purpose of this study is to identify, categorize, and address the reasons for misinterpretation, privacy, and security issues that users encounter when using AI virtual assistants with the goal of improving their effectiveness and user satisfaction. Addressing these issues in AI virtual assistants is an ongoing challenge, continuous research, and necessary to improve their capabilities.

Therefore, there is a need to gain knowledge about transparency and user education about the AI virtual assistant’s limitations to help users better understand and interact with these systems, reducing hampering and improving overall user experience.

RESEARCH OBJECTIVE

General Objective: The general objective of this paper is to improve the AI virtual assistant's efficiency and ensure user satisfaction.

Specific Objective: To investigate the causes of misinterpretation, privacy, and security issues that users face when using AI virtual assistants, overcome misinterpretation, privacy, and security problems by coming up with new ideas, recommend future scope for working on these problems.

CONTRIBUTION OF THE STUDY

This study contributes to AI domain by finding the reasons and possible solutions for misinterpretation, privacy, and security problems. The role of misinterpretation, privacy and security solutions in Artificial Intelligence is crucial for ensuring responsible and ethical use of AI technologies. Implementing privacy and security measures is essential to protect AI systems. This includes secure model training and deployment, encryption of data and model parameters, and continuous monitoring for potential attacks. These are crucial to mitigate privacy and security risks. Moreover, the solutions of misinterpretation play a significant role in improving the robustness and interpretability of AI models.

RELATED WORK

According to a report, Google has admitted listening to private recordings of customer conversations via Google Assistant when the leak of 1,000 private conversations in Dutch has come to light [2]. A Belgian news site claimed that third-party contractors working for Google could also access these multiple sensitive user conversations, which were recorded unintentionally. Users who have Google Assistant on their phones or smart speakers must typically say "Ok, Google" to initiate a conversation with the AI-powered virtual assistant. But VRT NWS claimed that the recordings were made even though some Google users did not even say the wake word,

"Ok Google" [3]. Moreover, Google also acknowledged the fact that it does not always remove recorded data. The company stated that users can completely disable the storage of audio data to their Google accounts or select to have data deleted automatically after three to eighteen months. The company retains the transcripts until a user "manually deletes the information."

In another report from ZDNet stated that Apple identified this in its newest iOS 15 updated version, acknowledging the fact that the AI-powered virtual assistant captured people's interactions, even if they opted out of it [4]. This bug of the latest version automatically enabled the settings of Siri, recorded, stored, and reviewed the conversations with Siri. Apple has since erased the recordings. When Apple found the bug, the organization switched off the setting for "some" Siri clients with the arrival of iOS 15.2. A similar update likewise fixed the bug. Additionally, with the second beta of iOS 15.4, clients will be inquired as to whether they need to opt-in or opt-out after the update has been introduced on their iPhone.

A demonstration of an experiment result found out that questions aimed at Siri and Google's voice search are routed to their respective corporations, where they are coupled with unique device IDs that aren't associated with specific users **[5]**. Therefore, this study found that our requests, like almost anything else on the Internet, will leave a trail of breadcrumbs.

Alexa specifically tunes in for its wake word “Alexa”, but it cannot selectively filter out voices. This implies that anybody in the room is possibly likely to its tuning in and recording powers. Any kid or house visitor might be drawing-in purposely or not with Alexa. Recording a youngster without parental consent is unlawful in certain states, like video recording regulations. In June 2019 Amazon was hit with two claims over Alexa recording kids' voices without the consent of the guardians [6]. It is now paying a fine of 25 million dollars. Amazon has agreed to change the way it handles children's data stored by Alexa. Under the agreement, the company will now delete all data, including audio recordings, upon request. In one unusual occurrence, Alexa sent a confidential discussion of somebody with their partner to a colleague unintentionally **[7]**. On the other hand, in another example, Alexa advised a 10-year-old to connect a coin to a half-embedded plug. Alexa had gathered the response from a web pattern (called the "penny challenge") and utilized information from that to recommend it **[8]**.

For privacy considerations, these speakers listen for their respective wake words, such as "Alexa" or "Hey Siri," before transferring the audio stream to the cloud for further processing. Similar phrases or noises, such as "cocaine noodles," instead of "OK Google" can be used to fool the virtual assistant.

Luca Hernández Acosta stated that, by default the user of a virtual assistant has full privileges, including viewing and deleting. In a multi-user scenario, secondary users consequently have no means to safeguard their privacy from the primary user or even the manufacturer and third-party developers. Only the main user administering the account has access to the stored data **[12]**. Google Assistant is the only platform which came up with the solution of it by enabling secondary end users to protect their own privacy. In this case, primary users will be denied access to previous interactions with others. To the best of our knowledge, Google Assistant does not save additional recordings for users whose voices are not recognized by the system. The primary user cannot see any recordings. In addition, Alexa and Google used authentication mechanisms such as voice match or voice profiles to secure privacy. However, this authentication process can be bypassed with recording and playback user voice **[12]**.

Louisa Olafuyi observed that, when her husband requests a song, Alexa frequently misunderstands him. But if he alters it to sound more British, Alexa eventually responds correctly, but not always **[9]**. A lab technician in Vancouver in her breezy West Coast accent, asked Alexa to tell her about the weather in Berlin (70 degrees), the world's most poisonous animal (a geography cone snail), and the square root of 128, which it offered to the ninth decimal place. But when Andrea Moncada, a college student and fellow Vancouver resident raised in Colombia, says the same thing in her light Spanish accent, Alexa only shrugs. She requests that it add a few numbers, and Alexa apologizes **[10]**. When a speaker with a British accent read one of the headlines — "Trump bulldozed Fox News host, proving once again why he likes phone interviews". Alexa came up with a more imaginative story: "Trump bull diced a Fox News heist, demonstrating yet again why he enjoys pain and beads" **[10]**.

A survey claimed that Alexa and Google Assistant are 30% less likely to understand a non-native English speaker **[11]**. They may also appear very different, inattentive, and unresponsive. People

with Southern accents, for example, were 3% less likely to get accurate responses from a Google Home device than those with Western accents. Furthermore, Alexa understood Midwest accents 2% less than those from the East Coast. The problem is that these devices are designed to learn some accents more effectively than others. Amazon said in a statement. "As more people talk to Alexa and with different accents, Alexa's understanding improves" **[10]**.

RESEARCH METHODOLOGY IN FLOWCHART

In this paper the proposed research methodology is Qualitative method. Among the three methodologies (Qualitative, Quantitative and Mixed) the most suitable one for our research topic is the Qualitative methodology. Because qualitative research is used to investigate, explore, and understand people's perceptions, ideas or feelings and often used to gain a deeper understanding of the complexity of a situation to draw a rich picture of what's going on. Therefore, qualitative data can be used to develop hypotheses and theories. The purpose of this method is to investigate the reasons for accent, misinterpretation, and privacy problems and propose solutions to these problems. For that, we need to explore and understand the complexity of privacy and misinterpretation problems and generate new ideas/theories based on the hypothesis to overcome them.

Proposed Solutions: Here, we have provided some solutions that might partially solve the problems.

Privacy and Security Issue:

- **Sensor:** Voice biometrics are vulnerable to spoofing. Attackers can attempt to spoof voice authentication systems by playing a pre-recorded voice sample in front of AI virtual assistant which is collected from a genuine user and can easily steal private information. In addition, fake calls or phishing messages can be used to trick the user's social contacts. This first approach needs an additional device(sensor) near the user’s mouth that monitors vibrations caused by speaking on the user's end. In this case, the user’s voice recording will be saved into the device prior to detecting the original user in future. To authenticate the user, vibration, and

movements of various articulators, such as the upper lip, lower lip, or jaw, in order to pronounce the phonemes will be combined with the user's command. It's a lively detection system. But carrying additional hardware, however, may be inconvenient for the users and may reveal more about them.

- **Volume and Gaze Control:** Another approach that is based on volume and gaze control, when a virtual assistant should start to listen rather than turning the device off or even unplugging it. This approach additionally makes use of a built-in camera in a virtual assistant to analyze the speaker's look direction and identify the original user. This would enable better control over when recordings are taken and could be implemented in the AI virtual assistants. As a result, less data is transferred to cloud services in case of accidental activations, ensuring that data is only processed when the original user gives command by gazing at the virtual assistant. Additionally, anonymization and encryption of stored data is needed to make it more secure. This method's drawback is that it needs a camera that would also gather data of the users.

Misinterpretation and Misunderstanding Issue:

- **Training:** Accents and dialects can significantly impact the accuracy of voice assistants. Users with non-standard accents or dialects may experience more misinterpretations. An AI virtual assistant requires different ways of speaking to learn different accents. The solution is: Different trainers with different accents should train the virtual assistant. Companies should use resources to train and test the systems with new languages and accents, including creating games that encourage more intercultural communication in different dialects.
- **Noise Cancellation:** Another approach is noise-cancelation technique that should be implemented for suppressing unwanted background sounds effectively. Virtual assistants can be triggered accidentally by similar-sounding words or phrases, leading to unintended actions or voice recordings. Similar phrases can be misunderstood by virtual assistants because of different noise and frequency. Frequencies of noises sometimes detected as wakeup word. Spectral subtraction or filtering algorithms can reduce frequencies or waves that are likely to be noise.
- **Robust Contextual Memory:** Human language is inherently ambiguous, and the same sentence can have different meanings depending on context. AI virtual assistants often struggle

to disambiguate and understand context. Understanding context is crucial for accurate communication. AI virtual assistants may not always grasp the context of a conversation, leading to misunderstandings. Implementing a more robust contextual memory system can help AI virtual assistants remember and refer back to previous parts of a conversation, reducing misunderstandings. It must be ensured that the AI model is trained in a wide variety of data sources and contexts to better adapt to different conversation styles and domains.

Flowchart: The flowchart of the solutions is given below-

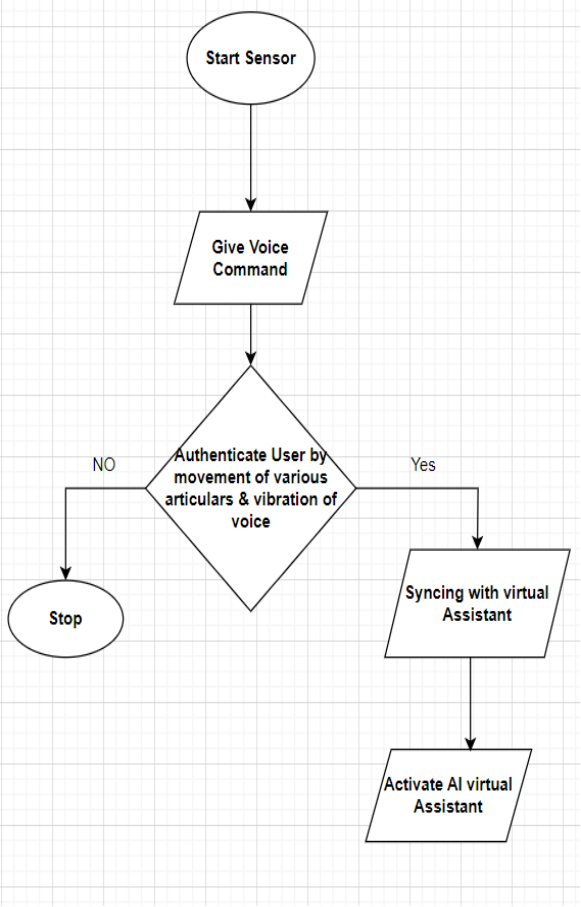

Fig: Sensor Device

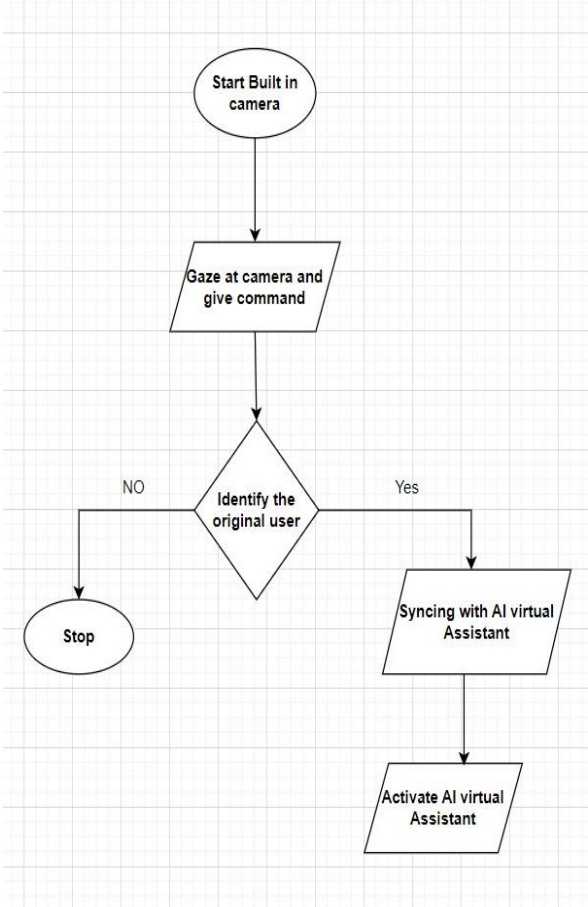

Fig: Volume and Gaze Control

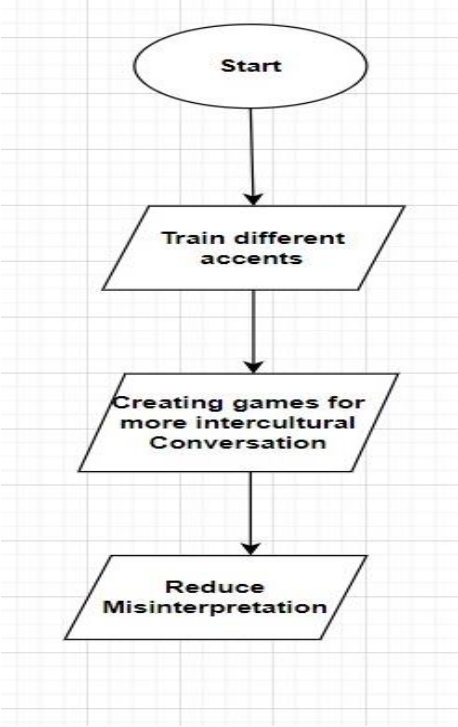

Fig: Training

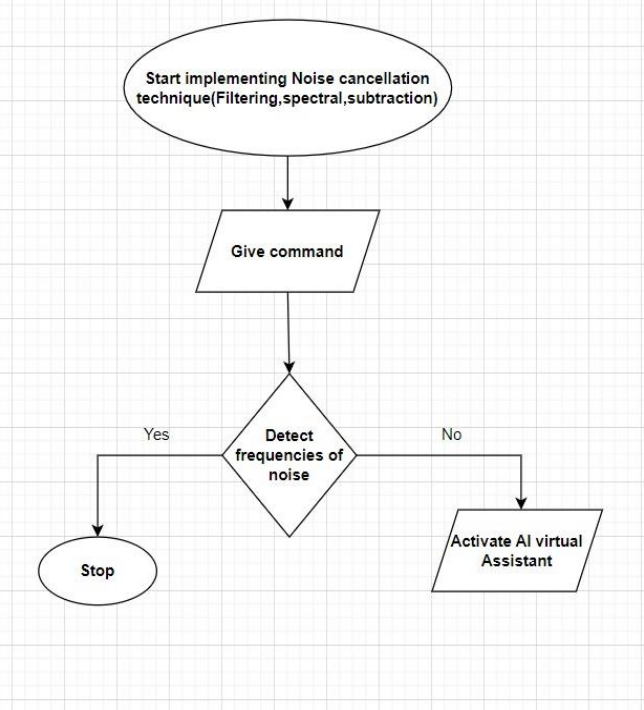

Fig: Noise Cancellation

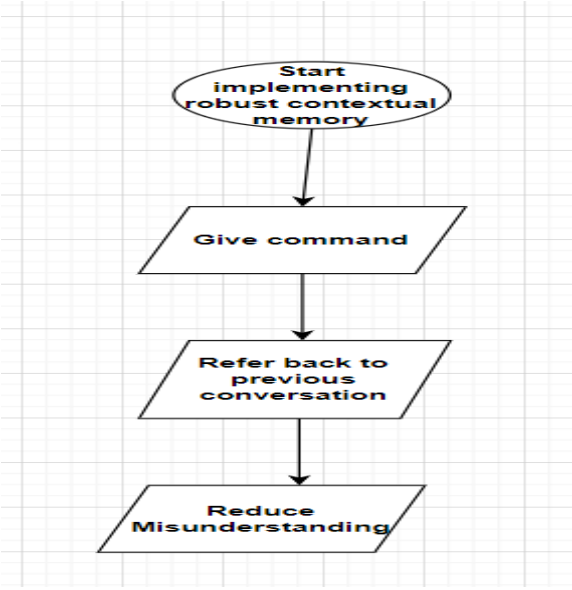

Fig: Robust Contextual Memory

SYSTEM DEVELOPMENT METHODOLOGY

After studying several SDLC models, it is decided that Agile (SCRUM) model is appropriate for the research of Misinterpretation, Privacy and Security challenges because this model is used to improve software quality and responsiveness to customer requirements. As AI Virtual Assistants need continuous change in their features at any time, SCRUM is best. Here's how Agile (SCRUM) model can be adapted to tackle these challenges -

Agile (SCRUM) has 3 phases-

1. **Pre-game**

- **Identify and Prioritize Issues (Product Backlog):** Creating a product backlog that includes currently known requirements related to misinterpretation, privacy, and security concerns in AI Virtual Assistants and prioritizing the backlog items based on the severity and impact of the issues. The product backlog list will be constantly updated with new and more detailed items.
- **Architecture:** Selecting prioritized backlog items and breaking down each selected item into smaller tasks to make them manageable during the sprint.

2. **Development/ Game phase**

- **Iterative Development:** Implementing iterative development cycles, typically 2-4 weeks (Sprints), to address specific aspects of misinterpretation, privacy, and security issues. Moreover, regularly reviewing and adjusting the approach based on feedback.
  - **Privacy and Security:** Implementing Sensor Device and Built-in Camera to adopt privacy, ensuring that privacy considerations are integrated into the development process from the start. Regularly reviewing and updating privacy and security policies and practices based on changing regulations and user expectations.
  - **Misinterpretation:** Implementing Filtering Algorithm, Robust Contextual Memory and training of AI Virtual Assistants to solve misinterpretation issues.

3. **Post-game**: This phase entered when requirements are completed. Testing of Sensor Device, Built-in Camera, Filtering Algorithm, Robust Contextual Memory will be done here. After this AI Virtual Assistant is going to be ready for release.

During the processing SCRUM has several meetings which will help the team to complete AI Virtual Assistants deliverables quickly and efficiently. Meetings are conducted to discuss progress such as what worked well, what could be improved, how to address ongoing misinterpretation,

privacy, and security challenges and emphasize communication and collaboration among team members working on different aspects of AI Virtual Assistant issues. Feedback gathered from users and stakeholders continuously improves the AI Virtual Assistant system. For fast-moving development SCRUM works well. Moreover, SCRUM ensures effective use of time and money. So, that's why Agile (SCRUM) has been chosen.

SCHEDULE & BUDGET

Schedule

Research Planning	1 Month
Related Work (Literature Review)	1 Month
Methodology Design	2 Months
Data Collection and Analysis	2-3 Months
Report and Paper Writing	1 Month
Review and Revision	1 Month
Finalization and Submission	1 Month
Total	**9-10 Months**

Budget

Personnel: • Researcher(s) • Data Analyst(s)	10k
Equipment and Software: • Computers/Laptops • Statistical analysis tools • Project management software	1-1.5 lacs
Data Collection: • Travel • ISP Bill	20-30k
Publication: • Publication fees • Printing and distribution costs	50-60k
Miscellaneous: • Fund for unforeseen expenses	10k
Total	**2.5 – 3 lacs**

DATA COLLECTION METHOD

As Qualitative method is used in the study, so data collection method is Literature Review (Related Work). It is a survey of scholarly articles, books and other sources. Articles which are related to misinterpretation, privacy and security issues, were collected from different journals, conference papers and books. Then, in-depth exploration of a particular case and complex problem was done to understand user's perspective and to find the solution as well as their thoughts are written in own words.

SIGNIFICANCE OF THE STUDY

This study represents a multifaceted approach to address the critical issues concerning AI virtual assistants, like privacy and security as well as accent and misinterpretation problems. The proposed solutions, such as implementing Sensor Devices and Built-in Cameras, offer potential benefits for user identification and interaction but also raise important privacy concerns that must be carefully managed. This study also reveals the emphasis on training AI virtual assistants, Noise Cancellation, and developing Robust Contextual Memory stands as a promising strategy to enhance their functionality and reduce instances of misinterpretation, making them more accessible and reliable for users. By doing so, it will be possible to work towards enhancing the overall user experience while maintaining the privacy and safety of virtual assistant users.

REFRENCES

[1] B. Thormundsson. "Users of AI Virtual Assistant." *Statista*. www.statista.com. (Accessed: October 3, 2023).

[2] R. Roy. "Google admits to listening in on private conversations via Assistant." *Business Today*, June 30, 2020. [Online]. Available: https://www.businesstoday.in. (Accessed: October 3, 2023).

[3] T. Verheyden, D. Baert, L. Hee & R. Heuvel. "AI Virtual Assistant Data." *VRT News*, April 25, 2020. [Online]. https://www.vrt.be/vrtnws/en/. (Accessed: October 3, 2023).

[4] J. Cipriani. "Apple's Siri recorded a conversation." *Journal of Artificial Intelligence*, vol. 118, no. 20, Jul. 2020. Accessed: October 3, 2023. [Online]. Available: https://www.zdnet.com/article/ios-15-4-update-why-youre-asked-to-help-improve-siri-after-updating/.

[5] K. Waddell. "Privacy problems with digital assistants." *The Atlantic*. www.theatlantic.com. (Accessed: October 3, 2023).

[6] G. Clauser. "Alexa never stops listening to you. Should you worry?" *Nytimes,* August 8, 2019. [Online]. https://www.nytimes.com. (Accessed: October 3, 2023).

[7] T. Warren. "Alexa got confused." *The Verge*, vol. 110, May. 2019. (Accessed: October 3, 2023).

[8] M. Yang. "Alexa told a kid to touch a penny to a live plug socket." *The Guardian*. www.theguardian.com. (Accessed: October 3, 2023).

[9] L. Olafuyi. "Alexa's struggle with understanding accent." *Medium*, October 24, 2019. (Accessed: October 3, 2023).

[10] D. Harwell. "The accent gaps." *Washington Post*, July 19, 2019. (Accessed: October 3, 2023).

[11] M. Graye. "AI Virtual Assistant User Data." *Venturebeat*. https://www.venturebeat.com. (Accessed: October 3, 2023).

[12] H. Acosta, L. & Reinhardt, D. "A survey on privacy issues and solutions for Voice-controlled Digital Assistants." *Goettingen*, February. 2019. (Accessed: October 3, 2023).