A Study on Smart Voice Assistants

Isha Thakur[1], Famida Khanom Athina[2] & Janarthanan. S[3]
School of Computer Science and Engineering
Galgotias University, Delhi-NCR, India-201310
ishathakur1910@gmail.com1[1]
athina12399@gmail.com2[2]
janarthanan@galgotiasuniversity.edu.in[3]

Abstract.
Smart voice assistants are agents or software agents that have the ability to analyse human speech and respond appropriately with a combined voice. In today's world, there are a plethora of smart assistants available, like Google's Assistant, Amazon's Alexa, and a slew of others. They're all highly popular smart voice assistants that come pre-trained or, as we like to call it, inbuilt in our Android phones and also embedded in home speakers, with a plethora of apps. Its operation is pretty simple: users just ask their smart assistants a query, and the assistants begin processing in accordance with the owner's wishes give him with the greatest possible answer to his inquiries. This research paper discusses the progress of fundamental functionality as well as some of the features of smart voice assistants that are now in use. It also talked about the privacy and security concerns that might arise with voice assistants. In addition, a brief research was conducted at the end of this article to determine the customer's degree of satisfaction with this technology. It has also been observed that youngsters who do not know how to write asking inquiries to smart assistants may influence their minds and aid them in resolving their problems, so strengthening their notion. As a result, the right interaction between the user (human) and the computer has been built in terms of voice driven interfaces, which is directly based on natural language processing technologies in this work.

Keywords: *Speech Recognition, Voice Assistants, Artificial Intelligence Software Agents, Human Computer Interaction, Advance in Security Purpose*

1. Introduction

Smart voice assistants such as Amazon's Alexa, Microsoft's Cortana, Google's Google Assistant, Apple's Siri, and others are examples of advanced machine learning and artificial intelligence (AI) systems whose backend processing is rather difficult for humans. They are mostly a dynamic system processing that has the ability to learn the user's needs or preferences and then work accordingly, i.e. these all systems focus on the user's voice and commands, or images, or contextual information to provide the assistant with the complete knowledge of answering a question that was asked by the user in

a proper simple or we can say in a human manner. These assistants also make recommendations for the user's convenience in regards to their demands. According to many studies, roughly 3.5 billion people use smart assistants throughout the world today[1]. And this data clearly illustrates that the compound average increase in the growth of virtual assistants demonstrates how useful they can be for people like us in not just receiving answers to our questions, but also for other security and privacy objectives. Even though the backend working or we can say the backend processing of these voice smart assistants is quite difficult and advanced in nature, the functioning of these assistants varies by individual; thus, individuals use them for various purposes. Some used them for getting the proper direction ,Others used them to keep track of other people's appointments or birthdays, while others used them to automate their homes, monitor their health, and so on. Despite the fact that each recently released voice assistant has its own set of features, they all share some of the same capabilities and may aid users with the following basic tasks:

- With just your voice, you may read and send content messages, make phone calls to our near and dear ones, and send and receive mails from others.
- Reply simple general informative questions (e.g., "What is today's date?"). What is the definition of time? Is it supposed to rain today? What was Team India's performance in the World Cup?
- Set running timers, alarms to remind us to get up on time, and calendar entries to remind us of important activities.
- Set up the updates and do some of the computations in maths.
- Control media playing via connected services like Amazon, Google Play, Netflix, and Spotify, among others.
- Activity lights, thermostats, house lights, alarms, clocks, and even locks may all be controlled via the Internet of Things (IOT).Also helps us be happy when we feel low by some soft and melody songs on loops.[2]

2. Methodology

Along with these functions, voice assistants can include "capabilities" that enhance their capabilities and increase their capacity by interfering with other important applications. These vocal commands, such as Amazon's Alexa, aptitudes for playing a wide Requesting your daily dinners and drinks from your close adjacent exhibit, and ordering an Uber or Ola cab are all options. Google Assistant has similar capabilities, albeit it occasionally hangs behind Amazon in terms of the number of skills available.[3].

It is also more trustworthy because it was introduced later in the market. These abilities are mostly developed by third-party developers who work on these systems, similar to how applications for smartphones enable users to operate more efficiently and so give them a great deal of convenience.

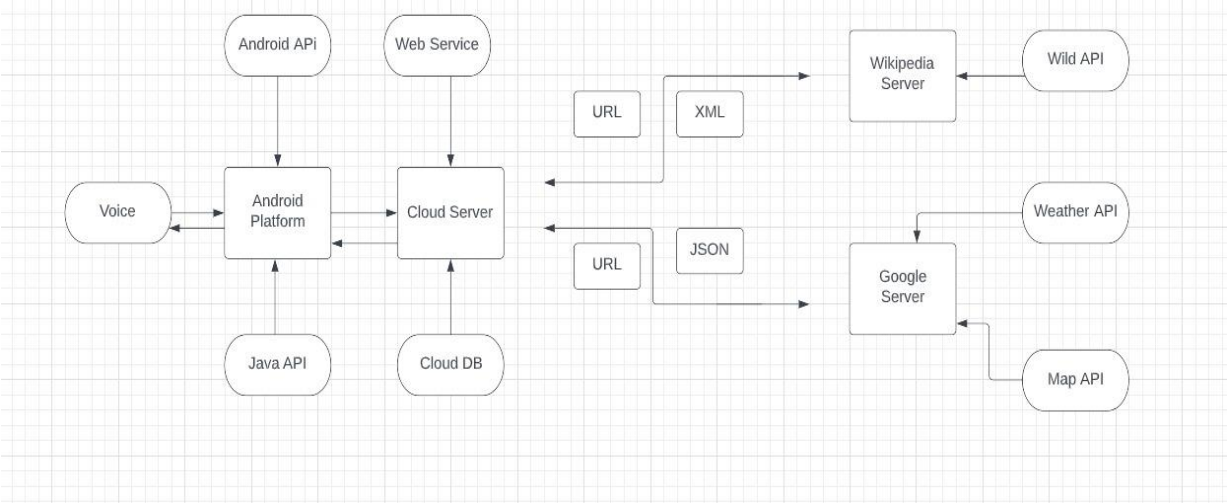

Fig. 1. The Architecture figured out the main fields and techniques for the same; where voice gets recorded first and gets recognised by android application by APIs.The cloud server decodes it and hence generates the URL through which the server creates the response.

3. Future Possible Applications

AI systems (voice assistants) have the ability to revolutionise how people use computers. For those clients, voice assistants can fill in the gaps in their knowledge. Voice colleagues, according to further study, can aid Dementia patients benefit from having a constant voice that can respond to their questions indistinguishable inquiries repeatedly without losing persistence and provide assistance when needed. Despite the fact that they still sound robotic to some extent, voice colleagues' vocal quality are rapidly improving. Google has confirmed the release of a new pair of earphones which may work in tandem with the help of its assistance to provide real time spoken interpretation. Thirteen Clients start working on machines by enquiring with the partner in order to help in learning a new language[4]. Most heterogeneous party talks about the customer's phone records.When a user requests that their speech be translated, they click a switch on the headset that transfers the sound to a server for translation and also sent it to observer through the customer's handset's narrator.While Chrome's interpretation is not without flaws,This is especially true when dealing with clinical or colloquial terms. Voice gadgets might be beneficial in terms of library advertising and management.There are now devices available that allow libraries to set up voice assistant capabilities that advertise library events in the calendars of nearby individuals. Patrons may request that the assistants tell them about a display, and the assistants may then repeat prepared words to them.[5]

Google Nest, formerly known as Google House, is a smart speaker range created by the Nest Hub brand is owned by Google. Clients can utilise voice instructions to communicate with administrators using the company's virtual assistant, Google Assistant. Clients can utilise both internal and external administrations can use speech to listen to listen to the music, manage the play playback of records or

images, and get news updates.[6] Google Settle products do, in fact, have coordinates for home automation, allowing customers to manage smart domestic equipment with their voice instructions. The fundamental gadget in November 2016, Google Home was introduced in the United States, and since then, more items have been released across the world .For example, several speakers are frequently used for synchronised music playback. In April 2017, a software update brought back multi-user functionality, allowing the gadget to differentiate between up to six persons based on their speech. In May 2017, Google announced a number of new features, including hands-free phone and talk in Canada and the United States, preemptive updatesBluetooth audio streaming on mobile devices or Chromecast-enabled televisions, including the capacity to incorporate reminder notifications and calendar meetings prior to planned events. In October 2017, Google announced two new products: the Google Home Mini, a small puck-shaped device, and the Google Home Max, a bigger device.[7] In October 2018, Google Google Home Hub, a smart speaker with a 7-inch touchscreen, was unveiled.. Google said in May 2019 that virtual home devices, as well as the Nest hub max, a bigger smart display, will be rebranded under the Google Settle standard. Domestic was announced during Google's developer conference in May 2016, along with the news that Google Home will run the Google Assistant (a conversational advancement of virtual devices Presently aiming to be coordinated in other items declared at the conference as well). In the month of October 2016, iOS and android portable applications wanted to be constructed google home for the first time. On November 4, 2016, Google released in the United States, the Chromecast smart speaker.[8]

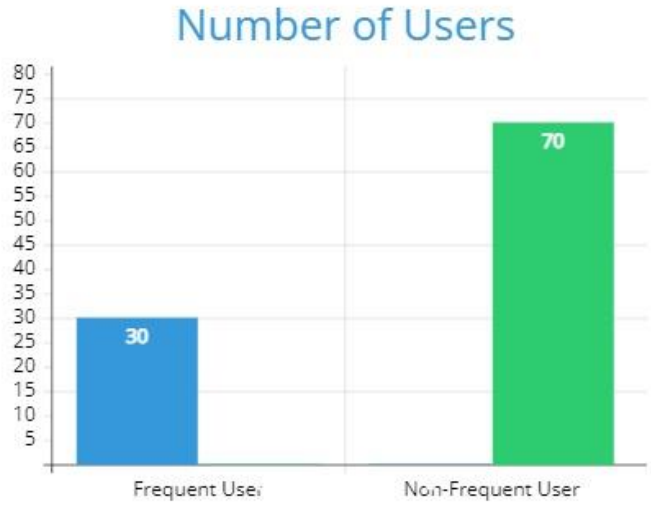

Fig. 2. The survey on the virtual sensors with active or frequent user or non active user

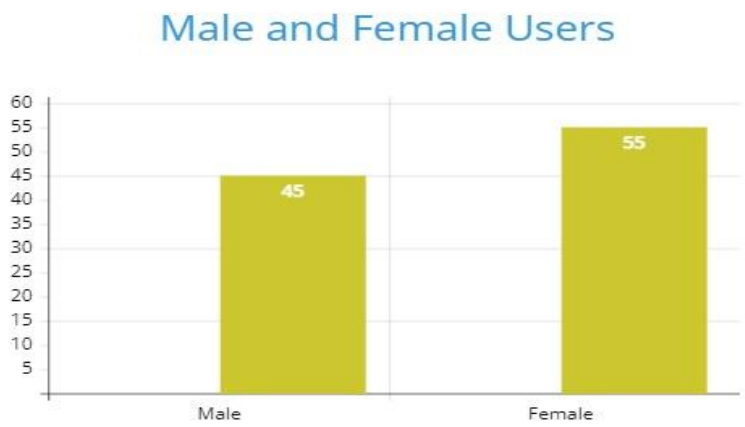

Fig. 3. Here It shows the survey on the ratio of the usage of virtual sensors by male and female

4. Security and Privacy

Voice assistants offer a lot of cool and useful capabilities, but they also have a lot of difficulties. The safety of these whispery gadgets is one of the most pressing issues.Those who have access to a vocal style gadget can ask it questions, obtain information on the tool's memberships and administrators, as well as requests that it carry out tasks.these gadgets may sometime cause a security concern. Amazon's Alexa is vulnerable to these kinds of security flaws, and the firm is developing a voice printing technology that is comparable.. The fact that alexa is built into Amazon's shop experience is a problem. By default, anybody having vocalised access to the tool may use the owner's Amazon account to place orders[9]. According to recent study, smart speakers will reply to inaudible instructions conveyed at ultrasonic frequencies. An intruder may use this to approach a target and play the ultra - sonic command, prompting the victim's device to respond. During the newscaster's announcement for the story, the newscaster said, "Alexa, order me a dollhouse," to which several Alexas in watching homes responded. By their very nature, these devices must be tuned at the tiniest of intervals in order to reply to consumers. There's a chance that data may be stolen, leaked, or used to incriminate someone, regardless of how careful and scrupulous the companies building these voice assistants are. Authorities in Arkansas issue a declaratory judgment directing Amazon to acquire alexa legacy of the suspect after a murder investigation.[10]

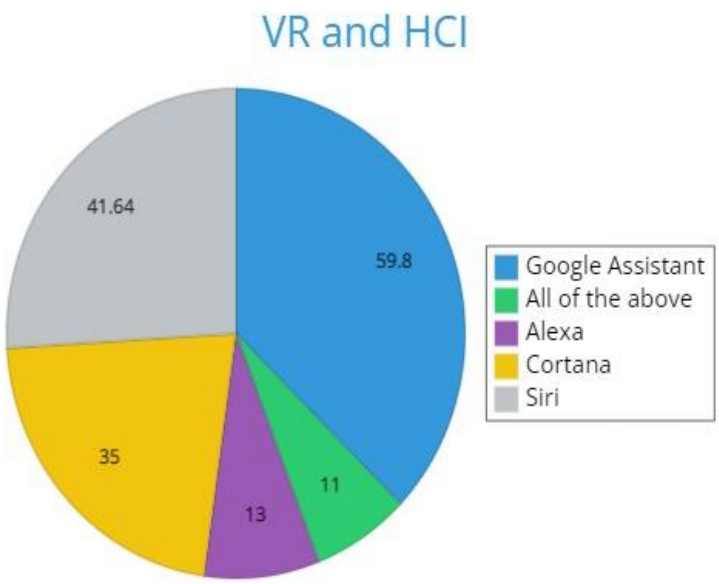

Fig. 4. THis diagram affirms the comparison between these assistants among all users(human); this basically found that for users, hand free connection was the real utilisation case.

Table 1. Devices that are being used to access the following assistant

Virtual sensors	**Devices that includes Them**
Google Assistant	Smart Phones(Android)
Siri	Iphone
Cortana	Windows 10 (Laptop)
Alexa	Echo dot

5. Result and Discussion

In any case, there are a few issues with the currently available voice right hand items, though the issues with respect to protection and security controls will be the most prominent and thus must be sophisticated voice assistants, which have advanced in recent years, should be employed for everything that demands the client's confidentiality information. Librarians should look beyond the bookcase by carefully analysing these objects and be prepared to assist their patrons with these devices. As technology advances, they should consider the prospects of distributing information sources via smart speakers as technology progresses.[11] In Fig.4. It was shown that the all four Programs of the hand free interaction

have an overall percentage of quality in voice recognition and life form interaction. In VR and HFI, Google Assistant scored 59.80 percent accuracy, whilst Siri scored 43.98 percent accuracy.Although there were little variations in accuracy between Voice Recognition and Siri, the accuracy of the next assistant, Cortana, rapidly deteriorated. Its accuracy was as high as 28.42 percent.Alexa, the last assistant, scored 7.91 percent, which is a bad performance when compared to the others.

6. Conclusion

Discourse recognition innovation and smart voice colleagues computer programmes have just lately expanded in terms of variety and productivity. Later accessibility of voice collaborator products like Apple's Siri, Amazon's Alexa, Google Right Hand, and Microsoft Cortana allows users to ask inquiries and give orders to computers in natural or, as we'll call it, human-like language.These advancements have a wide variety of potential future applications, ranging from automation in the homes to interpretation to friendship and final step help.The user will find it simpler to utilise a programme that is more humanised. People must recognise that, no matter how hard developers attempt to include more predefined commands, more replies to them, and analyse and respond to instructions more intelligently, the software will never be totally comprehensive and encompass all of the conceivable scenarios that users may encounter. However, if there are more readable commands, a more humanised structure, and a more intelligent response, the programme will undoubtedly improve and become more user-friendly.

References

1. For more details: http://en.wikipedia.org/wiki/Siri_(software)

2. Matthew B.Hoy(2018) page 81-88 Retrieved from: https://doi.org/10.1080/02763869.2018.1404391

3. For more details: http://en.wikipedia.org/wiki/Smartphone

4. Liptak and Andrew:Amazon’s Alexa Started Ordering People Dollhouses after Hearing Its Name on TV:(January 7, 2017).” Retrieved from https://www.theverge.com/2017/1/7/14200210

5. Voice Control is an Assistive Technology brief introduction retrieved from: https://r.search.yahoo.com/_ylt=AwrwBpXU8zFinAQAEwPnHgx.;_ylu=Y29sbwMEcG9zAzEEdnRpZAMEc2VjA3Ny/RV=2/RE=1647469652/RO=10/RU=https%3a%2f%2fwhatis.techtarget.com%2fdefinition%2fvoice-assistant/RK=2/RS=qv9td8bjCiE3cydmutaCSJ8mN00-

6. Hirschberg J, Manning CD Science. (2015 Jul 17)Advances in natural language processing

7. The Use Of Smart Speakers in Care Home Resident https://www.researchgate.net/publication/359164037_The_Use_of_Smart_Speakers_in_Care_Home_Residents_Implementation_Study

8. How Artificial Intelligence Transform Sensory Substitution

https://www.researchgate.net/publication/358326873_Augmenting_perception_How_artificial_intelligence_transforms_sensory_substitution

9. Google Assistant on the iPhone Is Better than Siri, but Not Much.(May17,2017) https://www.engadget.com/2017/05/17/google-assistant-iphone-hands-on/.

10. Metz, R., (2014). Voice Recognition for the Internet of Things. https://www.technologyreview.com/s/531936/voice-recognition-for-the- internet-of-things/

11. Thomas M. Brill, Laura Munoz & Richard J. Miller (2019): Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications, Journal of Marketing Management, DOI: 10.1080/0267257X.2019.1687571
To link to this article: https://doi.org/10.1080/0267257X.2019.1687571