WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Surveyon Voice AssistantWith Secure Framework Using AI AND ML

Geetha Rani E[1],Sunil Gowda S[2], Shashank S[3], Tejas S[4],and Suman A[5]

[1]Assistant Professor, Department of Computer Science and Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.
[2,3,4,5] Undergraduate Scholar, Department of Computer Science and Engineering, MVJ College of Engineering, Bengaluru, Karnataka, India.
[1]geetharani@mvjce.edu.in , [2]sunilgowdas11234@gmail.com , [3]shashank191220@gmail.com , [4]tejasdanny007@gmail.com, [5]suman2901a@gmail.com.

ABSTRACT:

In order to increase the precision and effectiveness of voice assistants while maintaining high levels of security, researchers have also looked at the use of machine learning and natural language processing techniques. These methods entail listening for irregularities in speech patterns that could point to illegal access or fraudulent activity. In general, creating a safe framework for voice assistants is essential to guaranteeing the security of sensitive data and upholding user confidence. Future studies in this area may build on the current work's solid basis, and continuous work will keep voice assistant security getting better. The start of voice assistants has completely changed how we use technology. Security issues have, however, surfaced with the growing usage of voice assistants in delicate jobs like banking and healthcare. Researchers have been working hard to create safe frameworks for voice assistants in order to solve this problem.Earlier research in this field concentrated on a number of security-related topics, including data privacy, encryption, authentication, and permission. To make sure that only authorized users can use the voice assistant, several researchers have suggested employing multi-factor authentication. Others have created ways to store and transmit voice data safely.

Keywords: Voice Assistant, Smart personal assistants, Privacy and Security

Introduction:

Speech recognition is used in smart cars to enhance the driving experience and give the driver hands-free access to the car's features, allowing them to concentrate on the road. One of the biggest worries about cars that may be started using voice commands has the advantage that children access the keys or keys usually in the car would possible to be able to start the vehicle without difficulty by voicing the assisted wake word. It is used to assist with various tasks, such as starting cars.[1]PVAs are beneficial, and their deployment density is rising quickly. For instance, 81% of individuals in the U.S. own a smartphone, and 21% of people own at least one smart speaker. Users are thus highly likely to be within range of at least one PVA constantly. Users could not be aware of them, have no control over how they behave or be unable to deactivate them. People are concerned about privacy because the devices can listen in on and comprehend speech.Questions like, "How can I control which PVAs are listening to my conversation?" might be of interest to users. How can I listen to recorded conversations?and how can I convey my need for privacy?[2]The most well-known applications that make use of VUIs or VAs, such as popular assistantssuch as Siri from Apple and voice enabled navigating apps, such as Google Maps and Apple Maps. Some apps require several strong permissions relating to user privacy to operate correctly, including access to the contacts, geolocation, the calendar, microphone, and storage. They can then use the internal speakers to speak back to users when they ask questions. These VUI responses frequently contain sensitive personal data. Stealthy IMU is a cutting-edge and useful threat that employs motion sensors with zero permission to obtain private information that is protected by permission from the VUI answers. These permissions are strongly connected to user privacy and are specifically provided to the VUI by the user. [3] Promoting the ideas that support informed agreement in the systems that control informationassemblage and authorization inside voice assistant-driven systems is, thus, a major driving force behind our study.[4] The SPA links Internet-based distant services and in-home integrated devices. The user can, for instance, check the conditions of weather and traffic, hearing to podcasts, conduct audio and video chats, shop online, or operate other devices like smart lights and similarly temperature sensors. SPA differs from conventional voice interaction systems in how it handles the user commands that are gathered. The best response to the user's question is given. SPA "understands" the user's voice instructions and derives the message's intent by utilising recent developments in natural language processing. To activate a function in a conventional system, the user must issue a precise command with correct pronunciation and a rigorous structure.A typical voice assistant listens for keywords like "Alexa," "Siri," or "Google" and continuously samples the audio coming from its microphone. The main method used to find keywords is called "Keyword Spotting" (KWS) contactregulator to avoice assistant which is on. The VAs streams the upcoming recording of audio to be used for voice command analysis when it recognizes the wake

keyword.[5] The use of home IPAs, also known as multipurpose speakers, in daily life is demonstrated in this commercial. The IPAs used at homes were immediately before IPAs on mobile devices. Voice-activated virtual assistants like Siri from Apple, Google Now/Google Assistant, Cortana from Microsoft, and Bixby from Samsung are just a few examples (Kinsella and Mutchler, 2019). Because privacy concerns vary, we make a distinction between Issues with platforms, security, and surveillance are discussed in the literature section. Finally, the results part tries to provide a holistic knowledge of IPAs used at homes and privacy concerns, thus highlighting the role of affordability[7].Due to the high level of personalization offered by VAs, this technology helps both consumers and businesses. (Brill et al., 2019). Because of these advantages, people may start to view veterans favorably. (Moriuchi, 2019). However, for some people, virtual-digital assistants are a causeof privacy terms. (Ghosh & Eastin, 2020; PWC, 2018). According to research, users' attitudes toward the VA may change due to privacy concerns since VAs have to actively keep listening to their surroundingsto recognize the commands necessary to carry out their duties. (Easwara Moorthy & Vu, 2015) despite the significance of gaining this information, research into the various effects of privacy skepticism on the uptake of new techs is still lacking. There isn'tmuch research looking at privacy.[8]The gathering and processing of personal data have virtually permeated every aspect of modern living online and is largely controlled by a few large technology firms. This is especially true for smartphones, which are carried everywhere and have a variety of always-on sensors, and where just two firms control both the operating systems and methods for distributing apps .The lack of openly accessible analysis tools, combined with iOS app encryption and the murky legality of their decryption, have contributed to the limited scope of iOS privacy research.[9]Despite the popularity of SPAs, there is growing worry about the hazards users may face from third-party capabilities. As bad actors may create potentially destructive applications that can compromise the users' privacy and security, skills increase the attack surface of SPA. Recent research has examined several third-party skill-related issues, such as publishing potentially harmful skills, collecting unjustified data, listening in on conversations covertly, and performing squatting attacks (invoking an action with a name like another skill but is conveyed differently).[10]The assistant must continuously record conversations to be proactive and provide contextual suggestions. This is a clear privacy risk and will only increase the myriad of worries consumers currently have about smartspeakers**.** Permissions, likethose utilized by mobile devices to restrict an app's access to personal resources like location or cameras, are one approach to limiting what assistants can hear. In reality, permissions are already used by existing voice assistants: When "skills" (third-party add-ons) are installed, Alexa, for instance, displays them when accessing users' names, addresses, or emails.[11]Without charge, permission is granted to make alphanumerical or physical duplicates of each or part of this work for private or educational use, provided that copies are created and distributed without a view to profit or commercial gain, bear this notice, and include the complete citation. on the opening page. It is required to respect any copyrights for parts of this work that belong to somebody other than the author(s). Credit-assisted abstraction is acceptable. Current voice-based VA consent implementation has a number of serious issues and disregards the fundamentals of informed consent.**[**12**]** As more individuals adopt them into their daily lives, voice assistants (VAs) are becoming more and more common. In addition to being integrated into smartphones, smart speakers are expected to have strong sales.to continue growing in the future. For instance, Tenzer (2021) estimates that 205 million smart speakers will be sold globally in 2025. Understanding voice assistants: a growing trendis crucial because for numerous individuals, it might or would be their primary opportunity to interact with speech-based technology, which will affect how they perceive it and how they will use it.[13][14]Together with voice assistants like Amazon Alexei and Google Assistant, Luis (voice user interfaces) for particular agents arebecoming typical in various uses like banking, contact centres, and medical services. These services frequently rely on speaker recognition to validate the person before employing speech recognition to comprehend a spoken language.We can think of an opponent in the security space who wants to deceive the target model. Aggressive attacks are sometimes called evasion attacks.In the privacy area, the adversary seeks to get either the model itself or sensitive training data attempts to violate data privacy.**[15]**[16]Alongside this, enhancements in the areas of Natural Language Processing(NLP) and speech synthesis and processing have significantly changed the world of voice assistants. (VA). Large technology companies like Google, Amazon, and Apple(Google Assistant, Alexei, Sirs, etc.) have suggested their own voice assistants, so long as a perfectly adaptable interface for the user. Other open-source projects and other large technology companies have also proposed voice assistants. Additionally, a lot of the voice assistants available today let programmers design functions (also known as "skills" or "actions") that can communicatewith the user as well as the surroundings' things. This user crossing point has a lot of latent for older people due to its ease of use and short time skill.[17]According to studies, if constraints. The potential benefit of smart homes may be maximized after challenges like interoperability cross compatibility of IoT devices, security as well as privacy concerns, and the utility of IoT data gaps are resolved. Also, there is a lack of information regarding how to effectively deploy artificial intelligence for voice-activatedcohesive smart home mechanization using the capabilities of various IoT devices. Although studies show that if limitations such system compatibility, security and privacy concerns, and the usability of IoT data gaps are resolved, the potential value of smart homes may be fulfilled, users presently concentrate on health and fitness. Also, there is a dearth of understanding of how to effectively apply machine intelligence for a speech-based IoT voice assistant.[18]The inadequate support for authentication in existing VAs is one of the main

security issues. Apart from simple customized wake words like "Alexa" or "Hello, Google," there is little effort for authentication in VAs. Although VAs may distinguish between people based on their speech profiles, this method has been demonstrated to be subject to direct replay assaults. Additional characteristics include limiting sensitive procedures like voice-based internet orders by employing PIN codes that can only be accessed by voice. Once more, the PIN code must be repeated aloud and is vulnerable to passive listening.**[19]**The use of intelligent voice assistants (IVAs) is expanding. IVAs can help users with a wide range of tasks and can recognize and comprehend voice-based user requestsIVAs are primarily used as digital assistants at the moment.IVAs will eventually take the role of laptops, tablets, and mobile phones in online shopping, according to Gartner. Consumers are expected to interact with businesses through their IVA and purchase goods and services from physical locations, the internet, and mobile apps in the future. Customers will therefore use their IVA to look up items and brands, compare them to alternatives, and post product reviews.**[20]**Voice assistants have developed into a popular technology product in recent years. The developments in machine learning (ML), artificial intelligence (AI), and Modern voice assistants, such as Alexa from Amazon,Siri from Apple, GoogleAssistant, Cortana from Microsoft, or Bixby from Samsung, can now be trained to mimic people and mimic the bio-neural thinking of human conversations thanks to natural language processing (NLP). The capacity of voice assistants to enable human to machine communications in a usual and instinctive manner, similar to daily human interpersonal dialogues, contributes to their appeal.**[21]**Governments and businesses have shown a growing interest in creating smart homes over the past ten years. Several internet-connected home appliances, like smart locks, smart speakers, and smart metres, provide a variety of services to enhance quality of life. Put, virtual assistants (VAs), also known as "smart speakers," are software programs that can understand user commands as a question or speech, carry out tasks, and respond with voice from the speaker. Examples of these programs include Alexa from Amazon, Siri from Apple, Google Assistant, Cortana from Microsoft, or Bixby from Samsung. These programs compatible with the dedicated hardware found in personal computers, cell phones, and tablets. The user can have natural conversations with the VA using this method of interaction.**[21]**Intruders have the ability to remotely enter the medical device's control unit and take control of it, putting patients' lives in danger. A passive network observer being able to deduce private patient information from network traffic would also constitute a serious threat to the patient's right to privacy, especially if the deduced information might later be misused. It is clear that the absence of a proper understanding of MIoT security among end users and other key stakeholders may worsen existing flaws and encourage attackers to further exploit MIoT technology, putting patients' lives in jeopardy in most circumstances. Moreover, data leakage or information loss would be the largest worries or hazards for healthcare in the event of any cyberattack.**[23]**The computing world has become incredibly big and complicated as expectations go beyond just connecting individuals. We are going to enter a new era in which everything will be interconnected. As a result of the quickening pace of technological development, an increasing number of people and companies are starting to provide services to clients through the use of intelligent devices including smartphones, home appliances, automobiles, wearable embedded devices, sensors, and actuators. Massive WSNs(Wireless Sensor Networks) and vast network-linked devices, together known as the Internet of Things (IoT), carry out the underpinning work. The IoT has gainedthe interests of tech developers in recent years and is often referred to as the Internet's future.**[24]**Voice and text chats make up the majority of the interaction. It is intended to mimic human interaction patterns, enabling human communication with machines. 1 Joseph Weizenbaum developed the first chatbots in 1966 when he created an application for computers that showed the possibility of natural language interaction between humans and computer.A chatbot is made in such a wayas to functionwithout a human operator's aid. It makes an effort to comprehend the questions and offer relevant solutions. If the conversation ever goes beyond the scope of the system's current understanding, it is either redirected or passed to human operators. Yet, contemporary chatbots also make use of machine learning algorithms to absorb the interaction and improve future answers.**[25]**Our goal isto implement these ideas to allow consumers to perceive the value of any suggested model, with common-use scenarios as due to the adoption of several third-party smart services, which might invade outside of the house and pose security and privacy risks, today's smart homes show gaps in privacy rules. To adopt smart home automation with a voice-based command and control strategy, we suggest a secure and integrated model connecting IoT devices and appliances. This system exploits the unified connection of IoT platforms to recognize network connected gadgets instantly over a secure network. **[26]**Voice-activated verbalized frameworks and content-predicted personal assistants can lock in crucial conversations and, at the same time, interact with guests in ways that blur the distinction between human cognition and calculated by a computer. Today, artificial intelligence has been shown to significantly reduce the need for human assistance when it comes to answering questions with simultaneous quick replication both inside and outside, at their sophisticated fingertips, and resolving conundrums that frequently arise a guest remain. A few traveling hotel chains have already implemented AI-enabled housing to respond to visitor queries via text or voice, reducing calls to the human concierge work area by 30% or more.**[27]**Traditional symmetric cryptographic systems, like Diffie-Hellman and Rivest-Shamir-Adleman, are difficult to apply on smart devices because of their little memory, processing capacity, andbandwidth. (RSA). As a result, this initiative offers a secure key establishment method that also enables mutual authentication between each entity, allowing participation in the home network only after that. Many different smart home protection

systems have been put forward in recent years. Additionally, many of these methods only consider eavesdropping attackers.[28]The innovation that enables users to ask for information via spoken commands is called voice search. (Li et al. 2009, p. 769). This research demonstrates how businesses respond to this modern disruption by using an example of the innovation diffusion of these services in Swiss tourism. Processes for voice recognition and natural language processing are becoming more sophisticated as they can depend on more powerful computers and algorithms. These services are useful to consumers and are made available through products like Apple's Siri, Google Home, or Amazon's Alexa thanks to the processes' steadily improving quality. A speech assistant is independent of physical devices like the Amazon Echo Dot. [29]One of the most prominent instances of the paradigms is the vocal assistant and associated smart home apps, which permit users to regulate lights and set reminders using voice commands and natural language at home. Nevertheless, this strategy has unforeseen privacy and security consequences: All data are continually transferred across the local and public networks with regard to a highly popular use case, a smart home. Using the Amazon Echo Plus hub and the Alexa Voice Assistant, the suggested technique illustrated how it is feasible to locate a working assault on such a system utilising relatively affordable specialised hardware and standard tools.[30]voice assistant devices, which take commands from their users.Siri, Google Assistant, and Amazon Alexa are a few speech assistant devices as examples. Siri, Google Now, and Cortana are a few voice assistants available on mobile platforms. For users to give verbal instructions or ask questions to the device, voice assistant devices use the IoT in the form of a speaker or a mobile device. The instructions or queries are transmitted to a computer via Wi-Fi, which then sends them to "the cloud".Rapid growth in the sharing requires the transfer of multimedia data via open networks like the Internet. dependable and strong security measures to maintain data secrecy and avoidillegal access to the transmitted data. Data protection is one of many possibilities. To make data unintelligible, undetectable, or impenetrable during transmission, encryption techniques alter data (such as text, images, audio, etc.). Data encryption now plays a major role in several applications because differentWith the end objective of strengthening the security and confidentiality of sensitive data, encryption techniques have been created, enhancing its secrecy and security.[32]Voice assistants are becoming more and more popular. Every platform now tries to have a voice-controlled assistant, and every big company has its assistant, like Alexa, Bixby or google assistant. With the rise of those assistants, users are increasingly accustomed to do basic tasks by voice, and voice recognition frameworks are released frequently. When interacting with others, employing the idea of privacy limitations and guidelines for third-party leaks gives a methodical approach to how people choose to expose or keep their private information hidden. extended this theory to human-agent interactions. In addition, three key aspects of users' perceptions of privacy in multimedia communication, particularly in video conferencing systems, have been found.[33]Voice assistants enable users to efficiently complete their everyday tasks thanks to new technological trends. Most virtual assistants use artificial intelligence and offer the users individualized support in the form of schedule management, smart environment control, navigation, appointment setting, wake-up calls, and many other functions. Now, many apps from various fields come, and Numerous well-known businesses have unveiled their own virtual assistants, includingSiri rom Apple, the Google Assistant, also Bixbyfrom Samsung, and Alexa from Amazon. These personal assistants offer ancollaborative user console (text, speech, or both) with the funtionality to comprehend requests, handlechallenging tasks, and produce the right answer using machine learning.[34] In the future, intelligent personal assistants (IPAs) will be used extensively to improve learning. The examination of the voice assistant taxonomy using IPAs in voice user interfaces (VUIs) implies that these assistants can reply to natural language stimuli instantly and intuitively, enabling the user to create speech interaction with the computer system. Many of them also allow for the creation of applications for their use and growth, such as the Google Home and Amazon Echo assistants. The functionality provided by IPAs consists of instructions for using the platform or the internet, advice for reading and writing texts and giving feedback on the outcomes of tests, such as quizzes. IPAs are also used in Moodle components, such as the "Lesson," which is one example. It can be used as a tool to create learning situations and scripts while keeping in mind the IPA architecture.[35]In order to examine the security issue we discovered, we also search the Google Assistant Store and the Alexa Skill Shop. Our findings indicate that about2,300 out of 2,400 Google Assistant Actions and 26,800 out of 32,900 Amazon Alexa Skills are possibly vulnerable. Further sampling-based verification reveals that 29% of these Google Assistant vApps and 71% of these Amazon Alexa vApps are truly susceptible. [36]This fine was imposed because the privacy policy was inadequate and did not provide consumers with enough information, not because it was not provided at all. Researchers have discovered that there are various contradictions between the privacy policies of mobile apps (like Android apps) rules. These inconsistencies may be the result of negligent planning by good developers or deliberate deception on the part of unscrupulous developers.[37]The existing system has the drawback that it can only hold limited commands and that only predetermined spoken directions are possible. As a consequence, the client cannot understand all of the facts clearly. These systems only use voice commands to carry out the restricted assignment. [38]The user's speech samples are currently only accessible to voice assistant service providers. Contradictory claims make it unclear, though, if application developers will have access to user voice samples. For instance, it has been stated that Amazon Echo and Google Home both intend to provide application developers with access to the raw voice command audio. As a result, there are more opportunities than ever to listen to consumers' unprocessed voice samples. [39] Voice biometric-

based authentication approaches that implicitly examine users' speech using trained (known) voice biometric features have been presented as a more practical authentication technique. Users do not need to physically authenticate themselves as a result. Nevertheless, when there are background sounds present, vocal biometric based authentication systems only achieve about 80–90% accuracy. This is because they are frequently employed with threshold settings that lower false rejection rates, which compromises security.**[40]**

2.Literature survey:

Over the past ten years, PVA usage has significantly expanded; more than 21% of Americans now own a smart speaker, and 80% of people in the country now have smartphones. Personal voice assistants are gadgets made to hear user input voice and recognize skill to carry out the necessary action in accordance with the command given. PVAs frequently include hardware—such as speakers and microphones—along with software to make it easier to listen, record, analyse, and take action. Typically, a PVA searches continually for a wake word**.[1]**In MSS since the sensor that simply records vibrationartefacts alsodoes so at a far reducing the sampling rate than microphones, MSS is a severely understudied issue. Since the sensor only records vibration artefacts and does so at a considerably lower sampling rate than microphones, recognising general speech from the MSS is a very understudied subject. But, StealthyIMU can get around the obstacle thanks to two crucial observations. First off, there aren't many machine-rendered voices available. Secondly, the sound properties of VUIs resulted to be more predetermined than those of real human voice**.[2]** A number of ethical issues with voice assistants have been previously recognised. These issues largely concentrate on social relationships and privacy. There is widespread agreement that people's perceptions in what waythe voice assistants work, including their understanding of confidentiality restrictions, are hazy and/or inaccurate. People's faith in outside safeguards, especially privacy laws, is one justification cited for using VAs despite having reservations. Speaking as a mode of interaction can be enjoyable and is associated with trust, according to research on anthropomorphism in voice assistants desire in more resourceful helpers about how datadistribution and permissions might be needed to be modified, particularly when dealing with sensitive topics like finances or health. **[3]**As SPAs become more commonplace of our daily lives, they could be specifically watched by bad actors. Also, the specific user-device interaction characteristics broaden the attack surface as they allow malevolentusers to interact with the SPA, access personal data, and potentially conduct transactions on the user's behalf; attacks in the Access Control category are of special importance.**[4]**A set of predefined keywords must be found in an audio stream by the KWS task. The VA's microphone(s) typically record the unauthorized audio stream. Finally, the VA conducts KWS classification and audio pre-processing. The KWS's functionality is essential to the VA's user experience. The responsiveness and utility of the device depend on a near-optimal true-positive rate. A KWS mistake, on the other hand, jeopardizes both the user privacy and the integrity of the Virtual assistant. A mis-activation occurs when an unauthorized command activates the VA.**[6]**The concept of privacy is nebulous, abstract, and fraught with moral dilemmas. The idea of privacy is often challenged because of its openness and internal complexity. Due to several definitions of privacy related to many facets of human existence, a holistic approach to privacy that conform to Koops et al.'s typology of privacy is usedin order to deal with its complex nature. Several issues in recent years have sparked a public discussion about IPAs in homes "listening in" and taking action without user consent. The news that a Google Home Mini was recording and relaying information to Google servers shocked many round-the-clock as a result of a broken touch button, which has since been disabled on all Google Mini devices.**[7]**Although consumers are utilizing virtual assistants (VAs) more frequently despite the privacy hazards, there are currently few research that indicatesthe variables affecting perceptions of VAs and their post-adoption behaviour (e.g., Ashfaq, et al., 2021; Kowalczuk, 2018; Moriuchi, 2019; Pal et al., 2020). For instance, Ashfaq et al. (2021) examined how attitudes and intent to continue using smart speakers are influenced by functional, hedonic, economic, and social values. They discovered that, except from this associationattitudes and social values, no further connections are relevant in explaining attitudes about the use of VAs**.[8]**Dynamic analysis monitors an app's behavior while it is running and gathers proof that private information is leaving the device. Early research centered on OS instrumentation, or changing iOS or Android. Recent research has shifted to analyzing network traffic because mobile operating systems are becoming more complex. Static analysis analyses apps without executing them. Apps are typically decompiled, and the resulting computer code is examined. The primary advantage of static analysis is its speedy app analysis, which enables it to grow to millions of apps.**[9]**Ina similar vein, 399 additional abilities were added between 2020 and 2021 that request permissions, and Amazon anticipates that developers will declare their collection in the privacy policy. It's interesting to note that according to the data, more abilities with full traceability have been added than those with broken or incomplete traceability. Particularly, new talents added with full traceability in 2020 include 518 (52%) and 256 (64%) capabilities. Device Address and Location Services are collected by this skill. Nevertheless, the skill displays faulty traceability because the links to the privacy statement take users to a dead page. Although things are generally getting better, privacy concerns persist with many recently submitted skills. As a result, we suggest that the screening procedure could yet be enhanced.**[10]**Architecture The studies presented in this paper could be used to inform

any of the numerous helper architectures that runtime permission requests may be a part of. Therefore, to clarify why the study permits cross-check to certain. Finally, we now outline a specific architecture that serves as the foundation for the application of permission in our study.It's a good idea to get a second opinion if you have any doubts about the accuracy of the information provideddesign decisions and presumptions. We see that in a broad design space, these possibilities only represent one conceivable set.**[11]** We summarise here a number of significant problems with the voice forward consent process that are suggested by an analysis of the pertinent literature and regulations. the addition of time constraints to the permissions procedure; the difficulty of communicating the necessary quantity of information via voice; the absence of separationthe Alexa Operating System (OS), voice from third-party skills, 2 Even if they do it in diverse ways, they all go against the accepted rules of informed consent.**[12]** We carried out a study in which families with kids who had never owned a voice assistant interacted with in order to test our hypotheses and research topics, 5-week period with an Amazon Alexa. This section details the precise characteristics of the sample of data we used as well as the metricswhich were derived from both empirical and conversational data. Ten families with a total of twelve kids were given an Alexa Echo Dot for five weeks. From the middle of January and the end of February 2020, the study was carried out in Germany. Regrettably, no log file information was found for one family.As a result, nine complete datasets were used for the calculations.Recruitment was done through leaflet distribution, local Facebook groups, and personal connections**.[13]**Several security applications depend on speaker verification. This is done in order to confirm someone's identity based on voice features. Two steps are involved in modern ASV systems: offline training and runtime verification. The ASV system builds a speaker model during Leveraging features for offline training. In the runtime verification step, the received voice is then compared to the trained speaker model, and the verification score is compared, to a predetermined threshold.**[15][16]**According to the "Ageing Europe" report from the European Commission, by 2030, 24% of the population in Europe will be 65 or older, up from 20% now. Also, the majority ofSeniors desire to age in place and continue living in their homes These statistics have prompted an increasing number of researchers to focus on initiatives that would assist older persons in ageing better at home and make information more accessible.and communications technologies (ICT) more accessible to them on one side, numerous research have concentrated on enhancing the home environment of older persons through the use of home automation systems that facilitate daily activities and incorporate ambient intelligence ideas like adaptation to the user's surroundings and Internet use.**[17]**According to a survey of current studies' literature, the effective consumer adoption of voice technology and the newest wireless technologies capable of connecting Internet of Things devices seamlessly is on the rise.Its usage is most frequently found in intelligent home automation, where it may serve as a personal assistant. A voice control system for smart homes that uses cloud data storage has recently been examined. The Alexa Voice Service and Amazon Echo were the foundation of the system, which operated on the cloud. It has been discovered that Alexa is "always on" and captures all voice interactions in the house even when they are not intended. Notwithstanding the fact that this data collection complies with the conditions.**[18]**To correctly extract signal features, a hand gesture's beginning, and end must be detected. This means that we must identify signal components that correspond to hand gestures. We discovered empirically that the signal phase is deterministically non-zero in the presence of hand movements as illustrated, whereas it is near to zero in the absence of any hand movement. As a result, the coefficient of variation for any signal value is significantly lower when there are no hand motions than when there are. So, to identify the beginning and end of a gesture, we choose a threshold based on the coefficient of variation.**[19]** IVAs are artificial intelligence-based digital assistants (AI). They are incorporated into smart speakers like Google Home and Amazon's Echo as well as cellphones like Apple's Siri (Chattaraman et al., 2019). These IVAs are capable of understanding voice-based user requests and responding in a manner consistent with human communication. IVA-capable smart speakers are becoming more widely used. With 30% of Americans using the Internet, the US is the second-largest market for IVAs behind China (Lee, 2018; and Peres, 2020). In 2020, 19% of Internet users in the Netherlands, the research's focal market, will own a smart speaker with IVA functionality. This figure was just 6% in 2018 after increasing by more than three times in the previous two years.**[20]**Literature survey research focuses on two different facets of voice assistant: (i) technical facets that aim to enhance the efficiency of speech identification, add emotion to speech, or enhance privacy and security facets of voice assistants; and (ii) user behaviour and acceptance facets that are mainly focused on user interactions with voice assistants as well as elements that affect the taking and approval of these intelligent devices. As the objective of the current study is more directly related to latter, we provide a current state of literatures relevant to the usage and acceptability of voice-assistants element.**[21]**A thorough study of the literature regarding privacy and safetyissues associated with virtual assistant revealed significant gaps in the state of the art of research. The issues of user concerns, the danger of malicious attack, and enhancing authentication have all been the subject of research. These studies, however, do not take a broad perspective on how these themes might interact, which could result in a disconnect between these domains. Many studies focused on user behaviour and identified privacy and security concerns; however, they did not address how these concerns might be addressed, with the exception of, which made a few recommendations for developments to the controlling, confidentiality default settings, and audio log capabilities, as well as for the addition of security layers to speech recognition, for privacy and security design. and

other technologies.[22]Attacks at the application layer typically target sensitive user data and aim to gain unauthorised access to it, which ultimately compromises user privacy. In order to make the services and apps vulnerableprovided by the application layer, attackers typically take use of software and hardware flaws (such as buffer overflow, code injection) on the device. Malware in the form of worms, viruses, and trojans frequently poses a threat to applications and services in addition to these attacks.[23]Several methods have been put out among the writings that address security concerns raised by the burgeoning Fog computing. The majority of these research publications either discussed issues with fog security or only paid attention to one particular area of fog security. By merging the views from numerous of these research studies, we have provided a succinct summary of Fog security in this article. Brief examination of several security issues in relation to fogsecurity attempted to identify various problem domains that corresponded to the solutions of the Fog computing environment. The aforementioned possible security and trust challenges were examined, analysed, and researched by Zhang et al., along with current accessible remedies.[24]Review of the literature outlining crucial steps in chatbot design methods with regard to chatbot security. Also, the paper discusses security threats and weaknesses and provides detailed instructions on how to create a secure chatbot platform. In conclusion, using chatbots for communication does not provide any security risks things haven't been properly addressed and dealt with yet. Since chatbots generate a significant quantity of user data, the major problems are more with how that data is used and kept on the backend side.[25]According to a review of the literature on recent studies, the effective consumer adoption of voice technology and the newest wireless technologies capable of connecting Internet of Things devices is on the rise. It is most frequently used in smart home automation, where it may operate independently without helper. A voice assistant system for smart homes that uses cloud database was recently examined. The Alexa Voice Service and Amazon Echo were the foundation of the system, which operated in the cloud.[26]A quick and safe method for setting up session keys in voice assistant settings is to create a session key and authenticate the IoT devices, it uses a brief authentication code. The minimal proceduremade certain that it could be used to gadgets with limited resources while maintaining the message security and freshness. It was intended to withstand well-known assaults including replay attacks, denial-of-service attacks, and known key attacks. The application of an Identity-Based Cryptography-based sign encryption system that authenticates without the need for a third party and offers secrecy and integrity. The connection is secured against a variety of potential assaults. This method implements elliptic-curve cryptography (ECC) and bilinear pairing.[27]Processes of idea diffusion and adoption among involved people and groups, as well as the adoption of new objects and content. Innovations have an impact on the creation of value and, as fresh concepts and methods, they present fresh tools and methods for its cocreation. Accordingly, these innovations suggest that businesses must modify their business models to adjust to environmental changes. They can continue to co-create value with consumers and, as a result, remain in existence. Thus, the theories of value cocreation, BMI, and innovation diffusion are discussed together in this part. The critical incident, in which a business rejects an innovation on behalf of customers, is the subject of theoretical focus.[28]It is now twice as common to have Alexa or Google Assistant in a household in the United States as it was in 2018. One element influencing the rise in utilization is the rivalry between Amazon, Google, and Apple.Also, well-known hardware manufacturers are experimenting with and implementing smart home technology, demonstrating the growing popularity and demand for this type of technology. The "newness" and "innovativeness" of the technology also piques the interest of potential consumers. However, expensive smart home technology may deter some prospective buyers.Before speaking a command, voice assistants watch for users to say a "wake word" before turning on the gadget. Yet it has been shown that Amazon Echoes can record and listen to conversations even when their owners haven't shouted the wake word. Echoes can occasionally start listening in on conversations when they mistake a word for the wake word. Users can attempt to avoid this by practising voice recognition so their voice assistant device can better recognise their voice. Users can also approve different permissions allowing voice assistants to access specific features, but occasionally voice assistants don't ask for permission, which is a potentially dangerous invasion of privacy.[29]Home Digital Voice Assistants (HDVA) are physical objects that function as agents and can carry out duties in response to voice commands. They are a special case of hardware-software systems. These devices can be used by regular people to carry out a variety of everyday tasks, such as asking their assistant a question, adding an event to their calendar, sending an email, playing music, or looking up events online, purchasing a plane ticket, and so forth. Today's HDVA technology can accurately understand human speech and react with synthesized voices. Due to their simplicity of use and the advancement of their technology, which can now handle numerous commands and perform a variety of duties, these systems have become very popular in recent years.[30]The method is based on popular question formats that people use to ask voice assistants for information or to get general responses to queries. Secondthe challenges associated with assessing the accuracy of information provided by voice assistants because these systems rely on speech synthesis and transcription algorithms, which have inherent flaws, to listen to, process, and react to user inquiries. To achieve this, we use a mix of experimental and crowd-based evaluation techniques to firmly establish the limits of our evaluation framework. This helps to demonstrate a significant technological gap in the ability of voice assistants to provide timely and reliable news information about important societal and political issues.[31]An overview of linked research found that chaotic systems could be effectively used to encrypt a variety of data. techniques of cryptography based on chaos that are

common might work well for text data, but they fall short when it comes to voice data protection. This is primarily caused by voice signals' high levels of redundancy and bulk data capability. Additionally, some confusing image encryption techniques have safety flaws, such as resistance to chosen-plaintext attacks, sensitivity to chaotic secret keys, the inability to properly restore the first pixel in the cypher image, and additional limitations on the inverse rectangular transform system's parameter selection. Contrary to general encryption methods, careful strategies only encrypt the most important data to render the entire multimedia content not possible to penetrate. Our cryptographic method is shown in the essential records from any of the very last bits-stream or the intermediate steps are typically the data section that requires encryption. For the same level of deterioration, encrypting this small amount of crucial résumé data requires less computational resources than encrypting a large amount of irrelevant data. **[32]**Design influence peoples' perception of privacy. To measure the experience of privacy in various real-life scenarios, a conversational speech corpus was provided in earlier work. While the compilation of the corpus was the focus of the work, an initial analysis revealed that various environments have an impact on how people perceive privacy. Based on these results, from the previous study we used a densely populated public listening test to allow detailed analysis on a sufficient sample size and to analysed characteristics that people use to gauge how much personal information they are willing to divulge in a given situation. Our goal was to examine how people perceive privacy in relation to background noise; listening test design and experimental setting had an impact on people's perceptions of privacy. **[33]** Virtual helpers help people with daily tasks like managing calendars, scheduling meetings, and giving wake-up calls.They make their daily lives manageable by offering clients a conversational service around-the-clock. With this new trend, numerous popular companies have introduced their own virtual assistants to handle the daily tasks of their clients. Authentication using voice effectively validates the user and guards against masquerading attacks. In addition, according to benchmark statistics from the User Experience Questionnaire (UEQ), the user experience performs admirably in all areas. **[34]** Since users must log in before using them, IPAs also include computer security tools. Their acceptance in university learning settings that use blended learning is also starting to spread, expanding their practical applications. IPAs also produce high levels of student satisfaction because they allow students to obtain instruction at their convenience and to receive individualized feedback. The use of this technology also gives academic leaders and instructors another way to communicate with students and learn about their main concerns. IPAs can also be used to educate students about administrative matters, and they are very helpful for learners who need both visual and auditory cues. These people value the flexibility of information searches.**[35]**NLU uses two steps, NLP transformation and intent classification, to obtain precise intents. The NLP component adheres to industry standards, including Named Entity Recognition (NER), Coreference Resolution (COREF), and tokenization (word segmentation). This collects the syntactic data for the provided textual material. NLU then compares the syntax data with an intent classification tree that has already been generated in order to understand the semantic meaning. Humans frequently make speech errors, which have been extensively researched as speech errors in the psycholinguistic. **[36]**Preliminary user study using the Amazon Mechanical Turk crowdsourcing platform found 48% of the participants said they were aware of the privacy regulations of the voice-apps they use when asked whether they were. Yet, 73% of respondents said they "rarely" read the developer's privacy statement when asked how frequently they do. 11% said they read it 50% of the time. 66% of the participants stated they never read the VA's privacy policy, while 34% indicated they do so using the smartphone companion app or a web browser. 47% of respondents were unaware of the data that the skill is gathering from them, while another 21% were unsure. **[37]**Making it a stand-alone personal assistant that can be used independently will enable it to accomplish its aim while overcoming the drawbacks of the current system exclusively through the voice of the client. **[38]**The body of literature reveals a dearth of studies examining existing problems with—and suggestions for—PVAs in the context of smart home control. Yet, it might reveal essential guidelines for creating smart homes.**[39]**The most popular method is speech biometric based authentication, which trains and categorises users using raw waveforms, intricate spectra characteristics, and log-mel features.Users do not have to complete extra tasks or recall extra data in order to verify themselves.It is a continuous solution that may be used to confirm each command issued by users.**[40]**

3.Related Work:

Machine learning algorithms are unable to process raw audio representations when working with audio data. The models instead use the audio features that were retrieved from the audio files. Pitch, loudness, and timbre are examples of features that can be used to categorise audio aspects that can be heard by humans. These characteristics are known as perceptual traits. Physical features are other features that cannot be expressed except by mathematical and statistical representation**.[1]**PowerSpy suggests using the shift in power usage brought on by cellular phone modems to pinpoint the user's position. Despite the fact that these assaults are connected to the StealthyIMU location authorization attack, presuppose either that the hacker is usually aware of the hosts starting position or that the victim is following a select few well-known routes. In contrast, StealthyIMU could reconstruct the user's GPS route by combining numerous navigable voices, extract the location using a single navigation voice, and even reveal sensitive spots.**[2]**To rule out situations if

either the smartphone is stationary or moving suddenly, MSS's first stage uses an empirical standard threshold to detect the vibration of interest. The signals are resampled to 500 Hz in the second stage, after which the magnitude values are normalised. Given that only a tiny portion of MSS survive the first stage, resampling here results in a manageable processing burden. The segments without substantial high-frequency components are then removed using a discrete Fourier transform (DFT).Finally, we must divide the signal buffer connected to a full voice sequence.[2]The results from first cycle: treaty and thoughtful the role of permissions for Points of Pact, The Nature of permissioned act as a Legal Need and RightFinestTraining, Joint Controllerships, disagreements and open questions, Adding Nuance to Consent Decisions, What to Include in Consent Dialogues, Opportunities Generated Through Speech—Round Two Results.[3]Personal assistants are also referred to as "virtual assistants" at times. Personal assistants are supported by many different families of electronic devices, including speakers, cell phones, Desktops, and wearables. They serve the purpose of gathering user utterances and sending them to the cloud of the SPA service provider. The assistant's "intelligence" is contained in this component. The requests are evaluated at this stage in order to determine the user's purpose. Third-party clouds and auxiliary clouds Users may interact with outside components not provided by the SPA provider, such as third-party skills or smart devices made by other companies.SPAs are connected to thecompanies cloud that provide these external parts in order to access them.[4]Google, Apple, Amazon, and Facebook paused the default enrollment following public outcry regarding the manual transcription of voice instructions and are now providing users with the opt-in and opt-out options .A configurable keyword sensitivity setting has recently been made available by Google to provide consumers more flexibility over the utility and privacy tradeoff However, these efforts fall short of users' expectations for hands-free interactions and do not address the privacy issues posed by VA miss activations.[6]The selected groups give details about the beliefs and experiences that underlie a variety of privacy concerns and show how these are connected to certain affordances. It's interesting to see that most respondents prioritise immediate and physical hazards while viewing hypothetical and long-term concerns from a practical standpoint. On the latter, the focus group findings reveal similarities between our Dutch respondents and prior studies. According to US-based study, practical attitudes based on how difficult it is for platforms to keep and analyse recordings from all devices constantly deter concerns about recordability and listening in (Lau et al., 2018). [7] The five-statement A Choi et al. study's scale was utilised to measure privacy cynicism (2018). Ratten's study served as the source for the five assertions on perceived utility and usability. (2015). Five statements taken from the research of Moriuchi et al. served as the basis for measuring attitudes towards VAs. Evidence suggests that this association also exists in this situation, where customers' privacy concerns might be minimised even when they are real thanks to their privacy scepticism. Because of this, privacy cynicism serves as a precursor to increased levels of trust for Vas.[8]We outline the process we used to gather and examine decrypted network traffic for potential PII exposure. In Section 3.4, we finally go into depth about how we dealt with tracking behaviours that were discovered by network analysis and code analysis. Next, using these terms to search the app stores, we were able to find a lot of apps and gather pertinent meta data (including title, release date, and time of last update). Due to the fact that these apps must adhere to the General Data Protection Regulation, we only included apps that were accessible through the UK region of both app stores in our research.[9] SPA stream of work like assesses the attack surface of SPA by looking into the sensitivity of the voice commands the SPA skills accept. We do not just look at how sensitive a voice command is; we also look into whether such a command is intended to collect personal information from the user. It introduces SkillVet, an automated tool that uses machine learning and natural language processing strategies to pinpoint skill traceability. The issue of gathering permissions while utilising poor privacy procedures is also motivated by our earlier research. By offering insight on how traceability has evolved through time and what causes these changes, this study expands on our earlier research. The traceability of privacy has also been examined in other research in domains including online networks or nodes, social media collectors, and mobile phone applications.[10]The Ambient Spotlight performed an automatic file search for records of meetings.Carrascal et al's research looked at how to extract crucial information from phone call transcripts. Real-time conversation-relevant online search results were ambiently shown by IdeaWall. The hazards they pose and the privacy concerns even current technologies arouse are the driving forces behind privacycontrols for proactive voice assistants. Since permissions are intended to protect especially sensitive resources.[11]The positioning of VFC the Alexa'srotation approach and the communication itself give the consent process a feeling of urgency that has never previously been observed. In alternative, Alexa's timer runs out after eight seconds , next reminds the user once more, teaching people to react within this time frame to prevent the device from following up. When VFC shifts the permitting of consent to the voice boundary clients consent since there is no way to revoke it via speech to the sharing of their data without being shown or advised of how to do so (or even that they can do so). The Alexa OS is in charge of voice-forward consent delivery. No discernible distinction can be made between speech produced by the Alexa OS.**[**12**]** disillusionment, lowered expectations, and unfavourable outcomes resulted from the VA's pre-adoption in the investigation.Recognition of the assistant's role can last all the way up to a lack of interest that leads to low use for basic operations or, in extreme situations, device abandonment. The discrepancy between users' expectations and the voice assistant's real capabilities is one of the potential causes for the development of this phenomenon. People with more technology exposure have further

realistic expectations of the assistant, according to Luger and Sellen (2016) and Cho et al. (2019). Still, users with less experience tend to base their expectations on known traits of Lack of information offered by the system itself prevents human-human connection, which ultimately leads to disillusionment at the absence of human-like conversational skills.[13][14]Several voice inputanalyzing techniques, voice conversion techniques include speech analysis, spectrum conversion, prosody conversion, speaker characterization, and vocoding. Speech analysis, mapping, and reconstruction modules are typically included in voice conversion pipelines.The analysis-mapping-reconstruction pipeline is now implemented differently thanks to deep learning techniques. A novel method of obtaining the intermediate representation is made possible by the deep learning idea of embedding. Two examples include speaker embedding for speaker identity and latent code for linguistic content.[15][16]An end-user interface with verbal capabilities is referred to as a voice assistant (VA), also known as smart home system. By utilizing several artificial intelligence techniques, such as speech recognition, NLP, conversation systems, and speech analysis, a voice assistants can converse with and comprehend users. Although there is two-way interaction between the user and the personal assistant, the user initiates the dialogue by introducing their assistant to do something. The personal assistant employs artificial intelligence algorithms to comprehend the data it receives and, when appropriate, carry out the action. Usually, the action denotes a reaction from the personal assistant. In the survey conducted by De Barcelos, it was discovered that VA have a general design that is summarized.[17]Because home automation systems only require small-scale installations as opposed to systems for commercial industrial automation, voice command databases are comparatively modest and can even be processed by hardware like the Raspberry Pi. This is helpful as it returns the control back to the individual user and mitigates the leaking of personal voice information. So, the goal of this research project is to control several Internet of Things (IoT) devices, develop an AI application that can be voice-controlled and run on a Raspberry Pi.[18]To identify a user, speech-based authentication systems use specific aspects of the human voice. Yet, research indicates that impersonation and replay attacks can compromise voice authentications. Inaudible voice instructions that are difficult for humans to understand but which speech recognition algorithms can understand can be used to trick voice assistants. It has been demonstrated that even voice processing systems like Google, Bing, IBM, and Azure speech APIs are vulnerable to concealed voice commands.[19]The use of IVAs throughout the customer journey because it is a rather recent occurrence in marketing. As a result, this study aims to fulfil an exploratory goal to investigate the fundamental customer behavioural patterns. To learn how the respondents will use their IVA during the decision-making process, qualitative, one-on-one in-depth interviews were used. Because the data collection was semi-structured, the questions and their order were flexible, allowing for a naturalistic and interactive research approach.Between July and August 2019, eight IVA users were questioned. Face-to-face interviews and online video interviews were used to gather the primary data.[20]Recent studies have demonstrated that a variety of private variables, such as self-sufficient, private inventiveness, technological anxiety, and technology knowledge, have an impact on the adoption process and sustained usage of technologies. According to authors in, the biggest obstacles to implementing new technology are sluggish uptake and scepticism. Anytime consumers are opposed to a technology, the adoption and utilisation scenario will be delayed. In fact, according to the diffusion of innovation curve, almost 50% of customers fall into this category of late adopters.[21]The most widely used virtual assistants (VAs) currently available are Siri from Apple, Alexa from Amazon,Cortana from Microsoft, and Google's Assistant. These assistants, which are frequently found in gadgets like mobile phones, can each be considered a "speech-based natural user interface" (NUI), a device that a user can control through instinctive, natural behaviour. Speech recognition, Natural Language Processing (NLP), and interaction storage are all cloud-based processes. The PAAS provider Amazon Web Services hosts Amazon's speech recognition and natural language processing (NLP) service, which is referred to as Amazon Voice Services (AVS) (AWS). Along with hosting AVS, AWS also manages the cloud storage where audio and data logs of voice exchanges are kept. The user endpoint and AVS exchange data through JavaScript Object.[22]The Cloud Security Alliance and the Open Web Application Security Project (OWASP), sometimes known as OWASP, have jointly produced a safe medical device deployment guide. It offers a thorough overview of the controls and safety measures that should be implemented to strengthen the security of a medical IoT environment, including perimeter defence mechanisms, network security controls, device and OS update guidelines, device security controls, security testing plans, and appropriate incident response plans. The protection of data during storage or transmission comes from data encryption. Even if the hacker obtains access to the medical database or transmission medium, solid data encryption will make it difficult for the attacker to read sensitive health data.[23]During the phases of its distribution, it encounters a number of challenges, including latency, bandwidth, quality of service (QoS), trust, security, privacy, and threats and attacks. Thus, the primary difficulties for the cloud computing paradigm are privacy and security. As a bridge between cloud data centres and edge devices or Internet of Things devices, fog computing was introduced as a new computing paradigm. The key compute components in the Fog environment, which is typically not audited by any security standard, are user devices and end devices. Consequently, the primary goal of this work is to create a thorough evaluation of contemporary methods and strategies for addressing the security of fog computing.[24]Threats and loopholes are the primary categories of security issues. A security threat is a risk that could compromise a system or an organisation. The STRIDE model identifies a

number of computer security threats, includingspoofing, fiddling, repudiation, information disclosure, denial of service, elevation of privileges, and many other. Each security threat should be mitigated by protective mechanisms that guarantee the following properties, including authenticity, integrity, non-repudiation, confidentiality, availability, authorization, and many others. Secure messaging can be divided across two domains. The first aspect has to do with data transfer security, specifically the secure delivery of messages, voicemails, and photographs to the hosting chatbot server. How the user's data is managed on the server is the subject of the second domain (backend), including how it is saved, processed, and shared. The user's domain's lifespan is covered by both domains. Threats to user communications can be found in the first message.**[25]**The major obstacles to the use of home automation systems in practise are privacy and security, and numerous studies have identified the essential elements of IoT security. We continue this research effort to address these issues against the backdrop of an extensive study on the security threats and vulnerabilities of an IoT network. First, we put our previously created end-to-end security paradigm into practice. To control different smart home devices, we created voice-based commands in Python and customized the codes for each sensor. To customize the home automation system, the AI speech recognition engine was then taught to learn different voices from different home users.**[26]**Voice-activated automated assistance tools. Amazon developed a novel version of its voice assistance that can help guests interact with the hotel much more effectively than in the past. The customer can also use Alexa in the room for services like calls, weather updates, ask questions, housekeeping, maintenance, etc. Inns are well known for implementing new technology to cater to the requirements of their guests. This technology has already been adopted by many hotel sectors, such as Marriott lodging, which is currently working use Amazon to implement the technology in their rooms. Additionally, Alexa will help hotels determine visitor engagement through analytics and announcing, allowing them to modify their services in accordance with customer demands.**[27]** The algorithm has been shown to satisfy the security requirements for mutual authentication, session key formation, message confidentiality, message integrity, and message freshness.Additionally, it has been demonstrated that the scheme cuts down on the time needed to perform the same process. Mutual authentication, session key establishment, communication confidentiality, and integrity are examples of security features. In essence, the limitations of the cryptographic techniques used are the only thing stopping the malicious organization. It was suggested to use an anonymous secure framework that could be installed in smart houses. The suggested framework achieves anonymity and unlikability for the smart devices in addition to integrity and authentication for the system by using lightweight operations. The session secret that was used was also dynamic for voice assistants.**[28]**In addition to Google's voice search, Siri, Amazon, and Cortana are some examples of voice search assistants that can be used. Voice search is the transformation of an analogue command into a digital one, or, to put it another way. Voice search optimises and expedites the web search procedure for any necessary information, reducing the need for typing and haptic efforts. They attempt to incorporate the voice assistant into their daily lives by using an Alexa-enabled Echo Dot. Experts stated that it is seen applications and studies demonstrating theoretical knowledge and comprehension. The respondents haven't, though, utilizing the technology in their tourism areas. The experts concurred that a chatbot or voice search tool should be used.**[29]**Echo Plus, an Alexa-based Amazon assistant, uses cloud-based speech services to carry out tasks like responding to common queries or managing home automation equipment Echo Plus, an Alexa-based Amazon assistant, uses cloud-based speech services to carry out tasks like responding to common queries or managing home automation equipment. This knowledge base is regularly updated with exploitation techniques associated with threats explained in terms of tools and actions to take. There are actual attacks that can be carried out using each exploitation method. The two primary sub-phases of the penetration testing planning stage, which is based on the threat model gathered in the preceding phases, are the identification of the threatened assets and the subsequent threat testing activities planning. It's important to keep in mind that the risk value determined during the assessment for risk step may influence threats and asset test prioritization in some cases, but not always.**[30]**The Echo is a smart device that uses a voice assistant. Before plugging the Echo into an outlet for setup, customers must first download the app onto a smartphone, tablet, or other mobile device. People using could then input their voice to the app and start linking it to their Wi-Fi by using the app. Users can configure other features, such as their lists, audio, news, or settingandB as calling and messaging through the program. When the Echo is online, it can respond to queries and carry out duties like setting reminders. Compared to other speech assistants, Alexa is more effective at learning. Since more than 7,400 brands of smart products are compatible with Amazon's Alexa, compared to about 1,000 brands for Google Assistant and 50 for Siri, it outperforms its rivals. The drawback is that Alexa is not preinstalled on devices like Google Assistant and Siri are. The Amazon Alexa app allows users to access voice assistance, but Google Assistant and Siri from Apple users can omit this stage.**[31]**The two operations that make up the recommended scheme for voice encryption's structure are masking and permutation of the speech signalwith the suggested methods. Receiver will then receive the encrypted speech signal via a channel and decrypt it to retrieve the initial statementsignal in accordance with the chaotic map. A key and a sound fileare used as the input for the suggested strategy. The audio recording is represented as a series of bytes. The sound header may be left alone so that the clientswill still be able to hear the jumbled noises when the audio data is run once again. The encryption method may only use the sound data input into a stack. The receiver will then receive the encrypted speech signal over a channel and decrypt it to retrieve the initial indication for speakingin

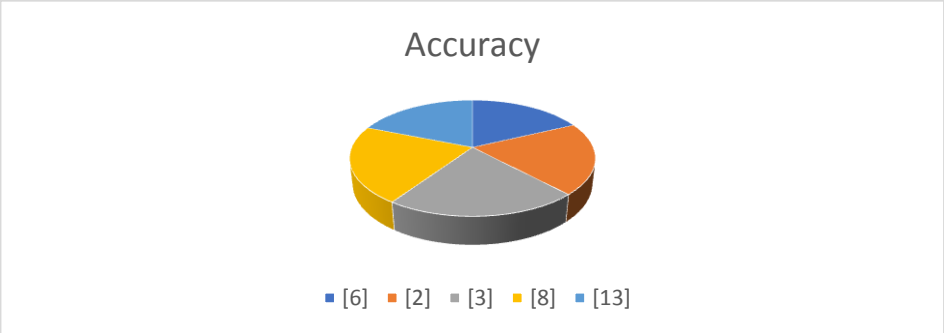
accordance with noisy map. Asound file and a key are used as the input for given strategy. The audio recording is represented as a series of bytes. The sound file header can be left alone so that the users can still hear the jumbled noises when the sound data is played again. The encryption method may only use the audio data stack as an input. Multiple metrics are used to assess how well the suggested voice system is performing. One tool used in numerous disciplines, for instance, is the distribution of data histograms. If the encryption has been operating as planned if the distributions of the integers used to represent encrypted data are close. In other words, the distributions of encrypted data are closer the higher the encryption level are to one another. Different measures that we are also used for quantitative evaluation to statistically analyse our findings.**[32]**Based on the results of two verification questions that was asked to the listening exam and the completion codes that each listener had to input on the Prolific platform, we have made a first effort to weed out unreliable listeners from our test results. Only after completing the entire test was the right completion code accessible. Additionally, you can judge a participant's reliability by looking for regular ratings. According to the transitivity of preferences principle, when aapplicant chooses scenario A to scenario B and scenario B to scenario C, possibly would also prefer A to scenario C. If not, the trio is referred to as circular or intransitive. In a broader sense, refers to the transitivity characteristic as consistency in preferences. The following includes a uniformity check for each.**[33]**Utilizing voice samples, we created an audio-based authentication protocol that can identify a person. To create a strong authentication method, the MFCC was matched of each common language input instead of using random text. Identity registration and identity authentication and validation are two of the two sub-modules that make up the identity and access management system. The user must finish the process of signing up to access MIRA services utilising the characterrecording module. For this, MIRA collects a smart device identification in addition to private information like username, the users address, the users gender, age, and voice samples. The intelligent device, the process of signing up to access MIRA services utilising the identity registration sub-module. For this, MIRA collects a smart device identification in addition to private information like name, address, gender, age, and voice samples. The intelligent device identifier and voice samples are among the data gathered to enable authentication. Age, gender, and medical history are factors in the personalized suggestion.**[34]**The students can access the key days for the course (delivery of practices, completion of surveys, project delivery, via a (mobile or computer) device, or a particular application was created. this software has a client-server architecture built into the Amazon Web utility for the Alexa utility system. (AWS). To access the computer program "UBU Voice Assistant," students must first authenticate their identity. UBU Virtual, the University of Burgos' learning management system (LMS), is utilized in this procedure. To authenticate their identity and gain access to the platform, students must present legitimate identification. The learner is then permitted to continue after these credentials have successfully been validated. IPAs can also be utilized to provide very helpful results for students who need visual or auditory support as well as providing knowledge to students on administrative topics. These users value the flexibility of access in specific Information looking. Recent studies have however also shown that each IPA needs to be modified to meet the requirements of each use, both in terms of the interface and usefulness. **[35]**Utilizing acoustic channels to attack an ASR. It is suggested utilising channels that are equipped to playback audio files, not hidden and can be understood by a human listener for audible voice command assaults.Then Tavish et al. introduced Cocaine Noodles, an attack that can be detected by a speech recognition system rather is difficult for humans to understand because it takes advantage of the distinction between synthetic and natural sound. Based on understanding of speech recognition processes, a white box technique was used to produce a better result.Attacking ASR with Misinterpretation provided an empirical analysis of speech misinterpretation-based vApp squatting attacks. Additionally provides a similar method for attacking the way a skill is invoked in a concurrent job by utilising a malicious skill with a name that is similarly pronounced or paraphrased to hijack.**[36]**et al. revealed that half of 116 VA owners were unaware that their recorded audio files are being saved by the device manufacturers through semi-structured interviews with 17 Virtual Assistant users. They also discovered that users have a constrained understanding of the Virtual Assistant ecosystem and related data. Similar to this, authors interviewed owners of smart homes to analyse user mental models and comprehend how they perceive the privacy of IoT devices. Geeng et al. looked into conflicts and difficulties that can occur when there are several users in a smart home environment. Interviews with VA users and non-users by Lau et al. suggested that privacy issues may be the primary barrier to new users. According to a survey, approximately 71% of applications without privacy policies ought to have one, and many apps demonstrate possible discrepancies with privacy requirements.**[37]**Natural language processing, a technique for converting speech to text, is a component of Virtual Assistant. In addition to becoming more conversational and environment conscious, it effectively responds to most questions. With Alexa and Siri, it's important to essential command right toget the desired response, but it is also really good at comprehending spoken language. The Raspberry Pi board is used to create the voice-controlled personal assistant, which executes the concept and logic it was designed with.**[38]**For the purpose of detecting speech replay or synthesis assaults that use loudspeakers, a number of strategies have been put forth. Wearable gadgets, like eyeglasses, earbuds, or necklaces, are suggested by Feng et al. and Liu et al. to detect speech liveness. These methods look for users' physical presence and muscle gestures to confirm that device owners have actually provided voice instructions that are being processed.**[40]**

3.Comparison and results

Table 1: Overview review of comparisons between various parameters

Reference papers	**[6]**	**[2]**	**[3]**	**[8]**	**[13]**
Space	Large space	Medium space	Less space	Less space	Medium space
Accuracy	Medium	High	High	High	Medium
Sensitivity	Medium	Medium	Low	Medium	High
Specificity	-	High	-	High	Medium

Fig.1.Graphical representation of accuracy parameter across different sources



The pie chart in Fig.1.shows the distribution of accuracy scores among different references and overall review. Each section of the pie chart represents one source of data and is labelled with both the name of the reference or project and the percentage of accuracy achieved. The blue section represents Reference [6], which achieved medium accuracy. The yellow section represents Reference [2] , which achieved high accuracy. The grey section represents reference [3], which achieved high accuracy. The Orange section represents overall review which achieved high accuracy [8]. The light blue section represents reference [13].

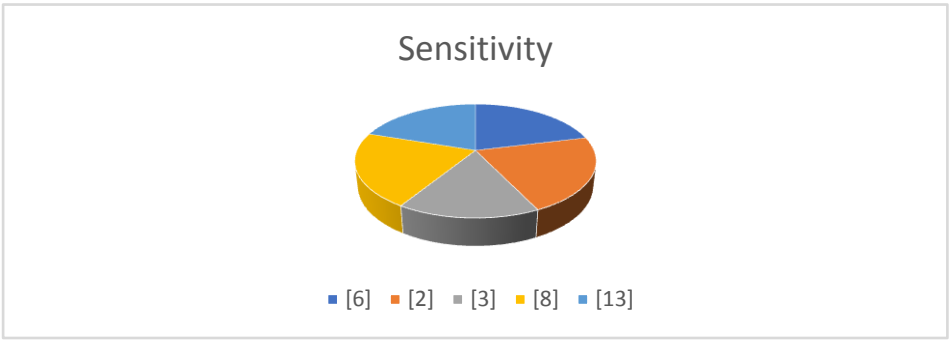


Fig.2. Graphical representation of accuracy parameter from different studies

This pie chart in Fig.2. shows the distribution of sensitivity scores among different references and overall review. Each section of the pie chart represents one source of data and is labelled with both the name of the reference or project and the sensitivity achieved. The blue section represents reference [6], which achieved medium sensitivity. The yellow section

represents reference [2] , which achieved medium sensitivity. The grey section represents reference [3], which achieved low sensitivity. The Orange section represents reference[8] review which achieved medium sensitivity. The light blue section represents reference [13].

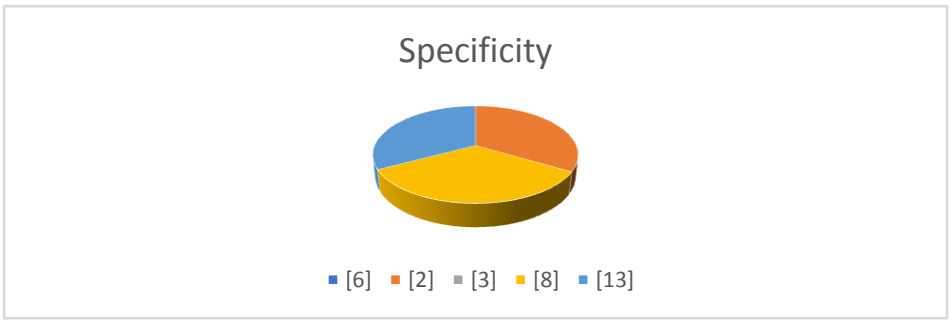


Fig. 3. Graphical representation of specificity for different cases and its effects on sensitivity parameter.

The pie chart in Fig.3. shows the distribution of specificity scores among three different references and overall review. Each section of the pie chart represents one source of data and is labelled with both the name of the reference or project and the specificity achieved. The yellow section represents reference [2] , which achieved high specificity. The Orange section represents reference [8] which achieved high specificity. The light blue section represents [13] having high specificity.

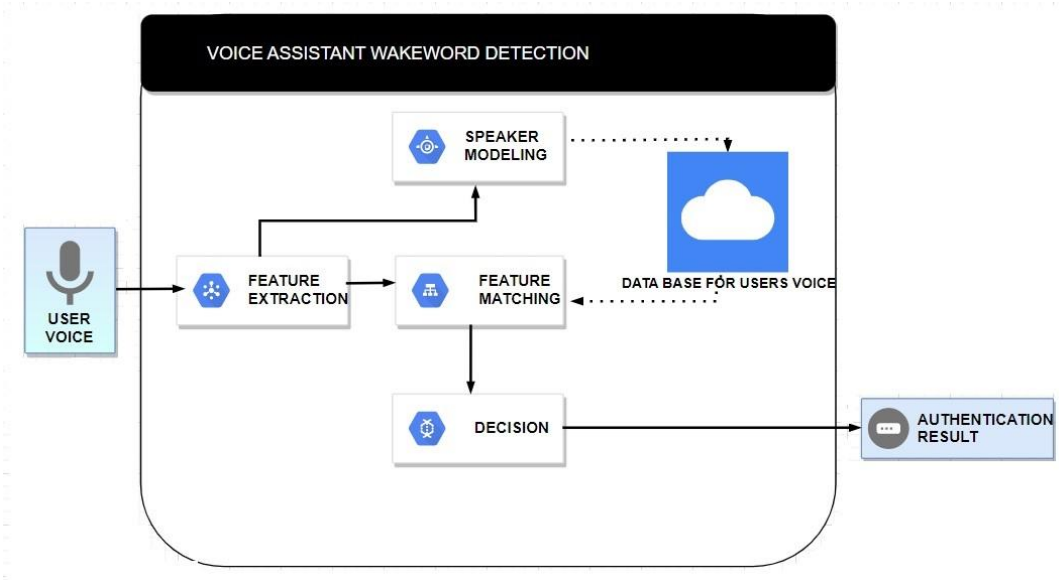


Fig.5. Diagrammatic representation of the authentication architecture in general

In Fig.5. the first block, "user voice," represents the user speaking to the device. The voice assistant then performs feature extraction on the user's voice, breaking it down into its component parts and analysing them. Next, speaker modelling takes place, where the voice assistant creates a profile of the user's speech patterns, accent, and other characteristics. This information is then used to perform feature matching, in which the voice assistant attempts to match the extracted features with those in its database. Finally, the voice assistant makes a decision based on the results of the feature matching, determining whether or not to respond to the user's input

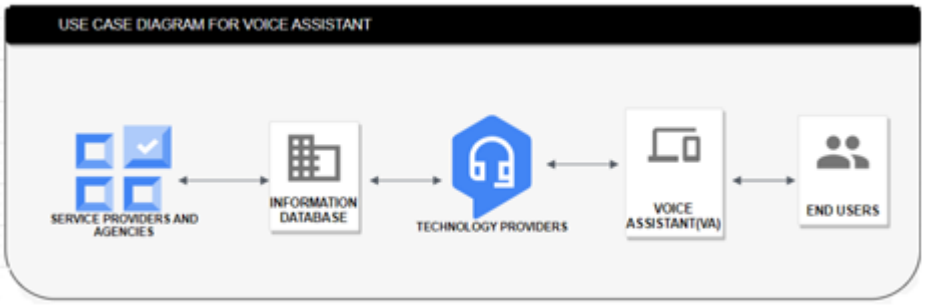


Fig.6. Representation of general use case scenario of voice assistant.

In Fig.6. the voice assistant is shown in the middle of a schematic for voice assistant usage, surrounded by service providers, agencies, technology suppliers, and database providers. The assistant's supporting gear and software are created by technology companies, and its data is provided by database companies. The experience is delivered to end users by service providers, and it may be promoted by agencies. The assistant's components all work together to guarantee proper operation and reliable information delivery to users.

Conclusion and future scope:

We researched in order to ensure the safety and secrecy of Voice Assistant. We highpoint primary safety and encounters in the literature survey from using 40 publications and found relevant papers to research questions. The writers of this survey were able to come up with alternatives in Voice Assistant security and privacy. To sum up, developing a voice assistant with a safe framework is a difficult but crucial task to guarantee the security and privacy of user data. A voice assistant security architecture should include several levels of protection, such as encryption of data in transit and at rest, safe authentication methods, and ongoing system audits and monitoring for security risks. It is crucial to take into account the security threats connected with speech recognition technologies, such as voice spoofing and unauthorized access to the device's microphone, while developing a safe voice assistant. Strong security measures may be put in place to reduce these risks and safeguard user data, such as multi-factor authentication, user verification, and access limitations.Ultimately, a safe voice assistant may provide people a practical and effective method of interacting with technology while protecting their security and privacy. To maintain the greatest degree of security for consumers as technology develops, it is crucial to keep up with the most recent security practises and include them into voice assistant frameworks.

Voice assistants that use ML technology for wake-word detection will become much more commonplace in our daily lives in the future. These voice assistants will be much more proficient at understanding user orders and precisely detecting their wake-word thanks to machine learning. Such interactions will be significantly more practical and natural if a wake-word can be reliably detected. Virtual personal assistants are a fascinating area where voice assistants with wake-word detection using ML can be used. These helpers might be created to do a variety of duties.As this technology advances, people with impairments or restricted mobility may be able to utilise voice assistants to control numerous gadgets all around their home or place of employment, considerably enhancing their quality of life. Overall, the future of voice assistants combining wake-word detection and ML technology looks very bright, and in the upcoming years, we can anticipate the emergence of many more cutting-edge applications.

References:

[1] Marina Maayah1· Ahlam Abunada· Khawla Al-Janahi1 · Muhammad Ejaz Ahmed2 · Junaid QadirJunaid Qadir, jqadir@qu.edu.qa; Marina Maayah, mf2000610@qu.edu.qa; AhlamAbunada, 199955256@qu.edu.qa; Khawla Al-Janahi, kj1304502@qu.edu.qa; Muhammad Ejaz Ahmed, Ejaz.Ahmed@data61.csiro.au | 1 Department of Computer Science and Engineering, Qatar University, Doha, Qatar. 2 CSIRO's Data61, Sydney, Australia.Discover Artifcial Intelligence (2023) 3:8 | https://doi.org/10.1007/s44163-023-00051-x

[2]StealthyIMU: Stealing Permission-protected Private Information From Smartphone Voice Assistant Using Zero-Permission SensorKe Sun, Chunyu Xia, Songlin Xu, Xinyu Zhang University of California San Diego kesun@eng.ucsd.edu, cxia@ucsd.edu, soxu@ucsd.edu, xyzhang@ucsd.eduNetwork and Distributed System Security (NDSS) Symposium 2023 27 February - 3 March 2023, San Diego, CA, USA ISBN 1-891562-83-5 https://dx.doi.org/10.14722/ndss.2023.24077

[3] William Seymour, Mark Coté, and Jose Such

CHI '23, April 23–28, 2023, Hamburg, Germany © 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), (April 23–28, 2023), Hamburg, Germany, https://doi.org/10.1145/3544548.3580967.

[4] Cayetano ValeroJaimePérezSonia Solera-CotanillaMario Vega-BarbasGuillermo Suarez-TangilManuel Alvarez-CampanaGregorio López. Analysis of security and data control in smart personal assistants from the user's perspect. (10 February 2023) https://doi.org/10.1016/j.future.2023.02.009

[6]Shimaa Ahmed University of Wisconsin-Madison, Ilia Shumailov University of Cambridge, Nicolas Papernot University of Toronto and Vector Institute, Kassem Fawaz University of Wisconsin-Madison ,Towards More Robust Keyword Spotting for Voice Assistants August (10–12, 2022) USENIX Association 31st USENIX Security Symposium 2655

[7]Anouk Mols, Yijing Wang, and Jason Pridmore Erasmus University Rotterdam, Rotterdam, The Netherlands Household intelligent personal assistants in the Netherlands: Exploring privacy concerns around surveillance, security, and platforms(2022).Convergence: The International Journal of Research into New Media Technologies 2022, Vol. 28(6) 1841–1860 © The Author(s) 2021 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/13548565211042234 journals.sagepub.com/home/con

[8] Fulya Acikgoza and Rodrigo Perez Vega The Role of Privacy Cynicism in Consumer Habits with Voice Assistants: A Technology Acceptance Model Perspective(2022) Fulya Acikgoz fulya.acikgoz@bristol.ac.uk School of Management, University of Bristol, Bristol, Howard House, Queens Ave., Clifton, Bristol, BS8 1SD, UKhttps://doi.org/10.1080/10447318.2021.1987677ß2021TheAuthor(s).PublishedwithlicensebyTaylorFrancisGroup,LLC

[9]Konrad Kollnig, Anastasia Shuba, Reuben Binns, Max Van Kleek, and Nigel Shadbolt Are iPhones Really Better for Privacy? A Comparative Study of iOS and Android Apps DOI 10.2478/popets-2022-0033 Received 2021-08-31; revised 2021-12-15; accepted 2021-12-16.

[10]Guillermo Suarez-Tangil IMDEA Networks Institute Madrid, SpainJide Edu King's College London London, UK, Xavier Ferrer-Aran King's College London London, UK, Jose Such King's College London London, Measuring Alexa Skill Privacy Practices across Three Years © 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-9096-5/22/04. https://doi.org/10.1145/3485447.3512289

[11]Nathan Malkin and David Wagner, University of California, Berkeley; Serge Egelman, University of California, Berkeley & International Computer Science Institute Runtime Permissions for Privacy in Proactive Intelligent Assistants USENIX Symposium on Usable Privacy and Security (SOUPS) 2022. August 7–9, 2022, Boston, MA, United States.

[12] William Seymour william.1.seymour@kcl.ac.uk King's College London London, UK Mark Coté mark.cote@kcl.ac.uk King's College London London, UK Jose Such jose.such@kcl.ac.uk King's College London London, UK. Can you meaningfully consent in eight seconds? Identifying Ethical Issues with Verbal Consent for Voice Assistants. In 4th Conference on Conversational User Interfaces (CUI 2022), July 26–28, 2022, Glasgow, United Kingdom. ACM, New York, NY, USA. https://doi.org/10.1145/3543829.3544521.

[13] Mavrina L, Szczuka J, Strathmann C, Bohnenkamp LM, Krämer N and Kopp S (2022) "Alexa, You're Really Stupid": A Longitudinal Field Study on Communication Breakdowns Between Family Members and a Voice Assistant. Front. Comput. Sci. 4:791704.doi: 10.3389/fcomp.2022.791704

[14] Maria Vernuccio | Michela Patrizi| Alberto Pastore Delving into brand anthropomorphisation strategies in theexperiential context of name-brand voice assistant. Department of Management, SapienzaUniversity of Rome, Rome, Italy.(5 August 2021) DOI: 10.1002/cb.1984

[15] Michela Patrizi - Maria Vernuccio - Alberto Pastore. "Hey, voice assistant!" How do users perceive you? An exploratory study1(1 Feb 2021)

[16]RanyaAloufi, Hamed Haddadi, David Boyle Imperial College London. A Tandem Framework Balancing Privacy and Security for Voice User Interfaces https://doi.org/10.48550/arXiv.2107.10045

[17] Adrián Valera Román,DenisPatoMartínez,Álvaro Lozano Murciego,Diego M. Jiménez-Bravo andJuan F. de Paz. Voice Assistant Application for Avoiding Sedentarism in Elderly People Based on IoT Technologies (20 April 2021) https://doi.org/10.3390/electronics10080980

[18] Sitalakshmi Venkatraman,Anthony Overmars,Minh Thong. Smart Home Automation—Use Cases of a Secure and Integrated Voice-Control Systemhttps://doi.org/10.3390/systems9040077

[19] Shaohu Zhang szhang42@ncsu.edu North Carolina State University Raleigh, NC, USA Anupam Das anupam.das@ncsu.edu North Carolina State University Raleigh, NC, USA HandLock: Enabling 2-FA for Smart Home Voice Assistants using Inaudible AcousticSignal https: //doi.org/10.1145/3471621.3471866

[20] Karien Oude Wolbers ,Nadine Walter. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. https://doi.org/10.1016/j.dss.2007.07.001

[21] DEBAJYOTI PAL , MOHAMMAD DAWOOD BABAKERKHELL , AND XIANGMIN ZHANG3 Exploring the Determinants of Users' Continuance Usage Intention of Smart Voice Assistants *DOI:* 10.1109/ACCESS.2021.3132399

[22] Tom Bolton , Tooska Dargahi , Sana Belguith , Mabrook S. Al-Rakhami , Ali Hassan SodhroOn the Security and Privacy Challenges of Virtual Assistantshttps://doi.org/10.3390/s21072312

[23]Mohamed Elhoseny , Navod Neranjan Thilakarathne , Mohammed I. Alghamdi , Rakesh Kumar Mahendran , Akber Abid Gardezi , Hesiri Weerasinghe , Anuradhi Welhenge. Security and Privacy Issues in Medical Internet of Things: Overview, Countermeasures, Challenges and Future Directions https://doi.org/10.3390/su132111645

[24]Abdullah Al-Noman Patwary , Ranesh Kumar Naha , Erfan Aghasian , Muhammad Bilal Amin , Mingwei Gong ,Saurabh Garg , Sudheer Kumar Battula , Md Anwarul Kaium Patwary , Aniket Mahanti. Towards Secure Fog Computing: A Survey on Trust Management, Privacy, Authentication, Threats and Access Control https://doi.org/10.3390/electronics10101171

[25] MartinHasa ,JanaNowaková , KhalifaAhmedSaghair , HussamAbdulla , VáclavSnáše , LidiaOgiela. Chatbots:Security,privacy,dataprotection,andsocialaspectshttps://onlinelibrary.wiley.com/doi/epdf/10.1002/cpe.6426

[26] Sita Lakshmi Venkatraman Anthony Overmars and Minh Thong Smart Home Automation—Use Cases of a Secure and Integrated Voice-Control System Systems 2021, 9, 77. https://doi.org/10.3390/systems9040077 https://www.mdpi.com/journal/systems.

[27] AnsumanSamal Woolley, I. 2019. AI and Data Security: A Help or A Hindrance? Accessed 23 October 2019 https://doi.org/ai-and-data-security-123481333/.

[28]K Jeremy Varghese and A Vinnarasi IOP Conf. Series: Materials Science and Engineering 912 (2020) 062014 IOP Publishing doi:10.1088/1757-899X/912/6/062014.

[29] Anna Victoria Rozumowski, Wolfgang Kotowski& Michael Klaas. Privacy based on voice search https://doi.org/10.30958/ajt.7-4-1.

[30] Massimiliano Rak, Giovanni Salzillo,, and Claudia Rome . Voice Assistant privacy on home automation(2020) http://dl.acm.org/citation.cfm?doid=2806777.2806935.

[31]Calandra E. Weaver, Edward J. Lazaros, Jensen J. Zhao, Christopher B. Allen D. Truell .Security and Privacy on Voice Assistants (2020) https://doi.org/10.1109/ACCESS.2017.2747626.

[32]Ibrahim Yasser Mohamed A. Mohamed , Ahmed S. Samra and Fahmi Khalif Private Protocol for voice assistants(2020) doi:10.3390/e22111253 www.mdpi.com/journal/entropy.

[33] Perception of Privacy Measured in the Crowd – Paired Comparison on the Effect of Background Noises Anna Leschanowsky, Sneha Das, Tom Backstr ¨ om, Pablo P ¨ erezZarazaga.(2020) http://dx.doi.org/10.21437/Interspeech.2020-2299.

[34] Real-Time Assistant Ubaid Ur Rehman, Dong Jin Chang , Younhea Jung , Usman Akhtar , Muhammad Asif Razzaq and Sung young Lee (2020) doi:10.3390/app10072216 www.mdpi.com/journal/applsci.

[35] Effectiveness of Using Voice Assistants in Learning: Sáiz-Manzanares, Raúl Marticorena-Sánchez and Javier Ochoa-Orihuel (2020) doi:10.3390/ijerph17155618 www.mdpi.com/journal/ijerph.

[36] Yangyong Zhang, Lei Xu, Abner Mendoza, Guangliang Yang, PhakpoomChinprutthiwong, Guofei Gu SUCCESS Lab, Dept. of Computer Science & Engineering Network and Distributed Systems Security (NDSS) Symposium 2019 24-27 February 2019, San Diego, CA, USA ISBN 1-891562-55-X https://dx.doi.org/10.14722/ndss.2019.23525 www.ndss-symposium.org

[37] Song Liao∗ , Christin Wilson∗ , Long Cheng, Hongxin Hu, Huixing Deng School of Computing, Clemson University, USA. Measuring the Effectiveness of Privacy Policies for Voice Assistant Applications. ACSAC 2020,(December 7–11, 2020), Austin, USA © 2020 Association for Computing Machinery. ACM ISBN 978-1-4503-8858-0/20/12 https://doi.org/10.1145/3427228.3427250

[38] Piyush VashisthaJuginder Pal Sing Pranav Jain Jitendra Kumar. Raspberry Pi based voice-operated personal assistant (Neobot) (2019). Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology [ICECA 2019] IEEE Conference Record # 45616; IEEE Xplore ISBN: 978-1-7281-0167-5

[39] Tamino Huxohl, Marian Pohling, BirteCarlmeyer, Britta Wrede, Thomas Hermann Ambient Intelligence Group CITEC - Bielefeld University Bielefeld, Germany. Interaction Guidelines for Personal Voice Assistants in Smart Homes./(2019) 978-1-7281-0984-8/19 ©2019 IEEE

[40] Il-Youp Kwak, Jun Ho, Seung TaekHan , South Korea s.t.han@samsung.com Iljoo Kim, Jiwon Yoon Voice Presentation Attack Detection through Text-Converted Voice Command Analysis CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk © 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-5970-2/19/05 https://doi.org/10.1145/3290605.3300828