

A Survey Paper on Voice Assistants

N M Nithin Hegde¹ Prajwal S² Hemanth A³ Nithin D S⁴

^{1,2,3,4}I SE, SJBIT, Bangalore, India

Abstract— Everything in this world is invented with just one aim, to make life easier, from a mere toothbrush to an extravagant car. With a similar aim in mind to fulfil the thirst for knowledge, internet and search engines were created, where world's vast knowledge was within the reach of our fingers, but that wasn't enough and as time passed, voice recognition tools came into play which was later embedded into search engines to create voice assistants. Voice assistant is a fast-growing feature and it has the capability to change the way people are living. It has already made a huge impact through devices like Amazon Echo, Google home and Apple home pod. Voice assistant is widely used in Desktop and mobile phones through Siri, Cortana and Google assistant. It helps in making the life of a visually impaired person a little less difficult. Voice assistant is also combined with internet of things (IOT) to create a home assistant, to access all the appliances using voice commands. With so many boons it is bound to be one of the leading technologies used in the future. The goal of this paper is to show best ways to implement voice assistant-based device and to present people's opinion on voice assistant.

Keywords: Voice Assistants, Internet of Things (IOT)

I. INTRODUCTION

As the Internet came into existence a lot of things were heavily dependent on Internet. But the drawback was we were supposed to access the information by typing it in the browser. The introduction of voice assistant played a vital role as it works through voice command.

The system proposed speaks intelligently with the user in natural languages. It has a feature of providing a spoken solution to the real-world problems; these problems are solved using AI. This mainly helps visually impaired people in operating devices. The Voice assistant communicates with the user through voice and also provides a solution to the user through voice, so that it can be easily accessed by all kinds of people, ranging from small kids to elderly people, and it is also useful for physically handicapped people.

There are many digital assistants like Microsoft's Cortana, Google's Google Assistant and Apple's Siri. People's life has been dramatically altered by the existence of information technology. Nowadays almost all the users are dependent on voice assistants for their needs, like opening an installed app or setting an alarm. Voice assistant is a medium to access such technology. Artificial Intelligence is one of the most advanced platforms in the computing industry. Currently, voice recognition technology is rapidly changing and it is also creating an impact on daily day to day activities. Voice assistant is based on Artificial Intelligence and it provides assistance through voice recognition or text mode operation. Rather than serving as a standalone concern device, it also communicates with humans and also expands the human potential. A businessperson, for example, must make hundreds of critical decisions every day, while politicians must approve large projects based on expert

opinions. Artificial Intelligence is a useful tool for assisting in such decision-making. Voice Assistant also provides services to the user like mail exchange, music player, checking weather, playing videos on YouTube, Wikipedia search engine to the users.

The main use of voice assistant is that it can be used in home healthcare using which patients can monitor and know their health issues. For example, heart sensors monitor patients' heart 24/7. This is really useful to elderly people as they can get the updates without visiting hospitals and also for those people who are unable to use web applications, this voice assistant plays an important role. These voice assistants also help in reducing the cost of labour. Voice assistants are also used in internet-connected devices, in smart homes where we can control devices through the voice commands.

Figure 1 shows the General Block Diagram of Voice Assistant:

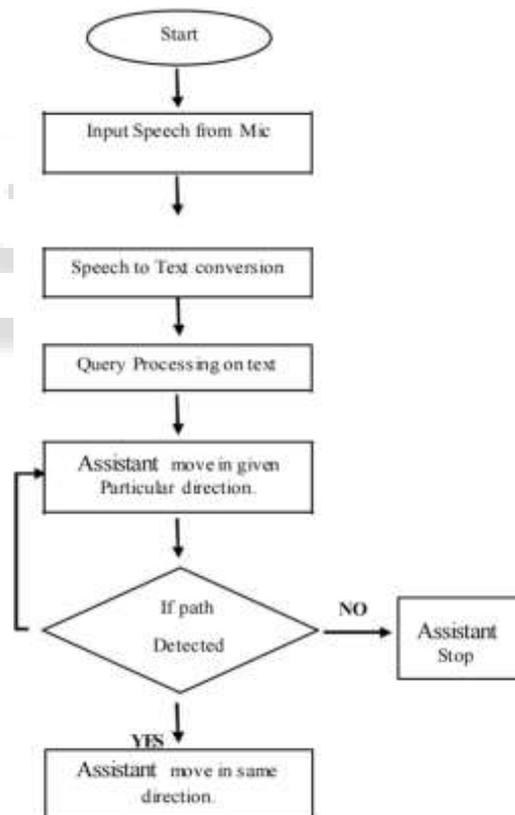


Fig. 1. General Block Diagram of Voice Assistant

II. OBJECTIVE

The main objective of the voice assistant is to provide a solution to the questions that the users may have. It is also used to access various data sources available on the web. Virtual assistants can assist the user with some of the tasks which are time-consuming and repetitive in nature and drastically save their time. Usually, human beings have a habit of forgetting important events of their daily life like

birthdays and anniversaries, voice assistant plays a very important role in this process. If you provide information to the voice assistant in advance about certain events, it reminds you on those particular dates. Most of the people are already aware of voice assistants, and they are also comfortable with it. Voice assistant helps a lot when the user is busy or running out of time. This shows us that the people are clearly moving away from screen interaction and are attracted towards voice assistants.

III. LITERATURE VIEW

[1]. This paper shows us that it was the first system to intelligently communicate in natural language with the users in LINUX and it is based on the character JARVIS, who served Tony Stark in the Marvel studio productions Iron Man in which the role was played by Edward Jarvis. It was made possible by Artificial Intelligence and thanks to the Artificial Intelligence, the most advanced platform in the computer industry is being used. It is built using gTTS, AIML and Python-based state-of-the-art technology. In this the voice of a male has been incorporated from gTTS libraries, inspired by the Marvel World, JARVIS. It combines the power of AIML with the industry-leading Google platform for text-to-speech conversion. [2]. This paper shows us that the voice assistants are user-independent, and it receives human voice commands in English language. The user can access installed programs, send and receive e-mails, and perform other tasks using the browser in the web. This model can also read news from news portal and allows the user to search for information and voice commands, it is also used to control household appliances. [3] This paper tells us that voice assistant is one of the fastest growing technology and it shows how the intelligent personal assistant can perform the mental tasks like turning on or off smart phone applications using voice user interface. The voice assistant listens to voice commands by the user and responds to the user by saying the phrase "done listening". It uses gTTS package that helps the voice assistants to speak like normal people and it is also used to convert the audio into text format and the advantage of these voice assistants is, it remembers the name of the person till the end of the conversation, it also performs the tasks requested by the user like playing songs from youtube and searching things from the google. So, it plays very important role as it provides hands free operation. [4]. This paper tells us that Intelligent Personal Assistants are developed to help people to perform simple digital tasks. The IPA is a computer program that works with the help of Artificial Intelligence. Here the IPA can process the commands in two languages, that is in English and Bengali, to complete the tasks which has been requested by users. The speech recognition module of this model is built using Sphinx-4 engine and a modified finite state automation handles the language processing. IPA interacts with the user through voice and here are the steps that describes how it works, in the first step it listens to the human's speech and in the next step it understands what is the input given by the user and in the third step it performs the action requested by the user. It uses Natural language processing [NLP] to determine what has been inferred in the sentence by the user and to decide on a course of action in response.[5]. This paper shows, when the internet was

invented, obtaining the knowledge from the internet became part of our daily day-to-day activities. It caused storage problem and hence the cloud was used to overcome the problem. The Dessy system is designed to be energy efficient. Dessy is written in Java language to allow cross-platform compatibility and to provide quick designing. The architecture of Dessy is modular. Keywords, property-value pairs, file metadata, and user defined tags are used by the client to find a file. Cloud service interfaces can be found on mobile device itself, or they can be accessed remotely via another Dessy host. The searches are executed in parallel on all the connected devices and local search in mobiles can be disabled to conserve the life of a battery [6]. The popularity of a personal digital assistant has risen due to development in wireless network access and computer performance development. They compare one terminal to others and the effect of QoS degradation on QoE along PDA. They target web services that have the same web pages regardless of the type of terminal, they chose two web services, one from a Japanese stationary and the other from a Japanese town. In the experiment, they treat TCP throughput and satisfaction as QoS and QoE parameters. A 5-point rating scale is used to assess satisfaction. [7]. This paper is about the assistant based on an Evidence theory and this assistant is applicable to enhance human capability in daily life decisions. This assistant helps in making decisions in various situations by using few keywords. Evidence Theory is best applicable in uncertainty reasoning and it is good in solving variety of problems with imprecise and incomplete information. Mass functions and belief are main concepts to work on uncertainty reasoning. To make use of the decisions easily, this assistant is constructed using three modules that is reasoning module, acquisition module and core module. [8]. Here in this paper users can communicate with virtual assistants such as Siri in a digital world to complete tasks. Links between natural language requests and their concrete realization are defined in these systems during the conception process. When the assistant is unfamiliar with a job, a more adaptive solution would be to enable the user to provide natural language guidance or presentations. An adaptive solution should allow virtual assistant to manage a much larger digital environment with multiple application domains and provides better matching user needs. [9] This paper aims to create an assistant using the Raspberry Pi processing chip, which includes components such as IR sensors, a camera, a microphone, and a motor driver. The voice assistant here is controlled by voice commands and uses artificial intelligence. To operate, it receives a consistent signal from the IR sensor. It uses Pi camera module to tell the difference between handwritten or printed material from an image, and then communicates the information to the client through a built-in speaker. The system's limitation here is that it can only provide predefined voice instructions and store a limited number of commands. As a result, the client is unable to access full data in an easily understandable format. The restricted assignment is carried out by these systems, which are either voice activated or optical character recognition.[10] A voice control system was created using Google Assistant, an open API artificial intelligence representative service, and IFTTT (IF This, Then That), a conditional auto-run system. IFTTT is a framework that helps you to connect all of your apps with your devices.

It uses Google Assistant to identify voices and the IFTTT platform to manage voice commands. To Open API, Google Assistant was used. Voice commands are sent via the Google Assistant app. After being recognized, voice data is sent to the Google server and it is converted to text and sent to the conditional auto-run mode IFTTT service. A single operation is performed by a single IFTTT recipe. [11] Here in this paper "Let's Talk!" is a new VOCA i.e. (Voice Output Communication Aid) application. VOCA is a communication assistant tool for autistic children. i.e., (children with communication disorder). Here tools such as Drop Talks, Voice4u and Tap to Talk come under VOCA. Devices with these capabilities have higher cost and is complicated to use, it not suitable for general use. As an alternative this paper, provides a new communication assistant tool, "Let's Talk!". When compared to existing VOCA, it has better functionalities and offers numerous benefits in low cost and while focusing on a straightforward and clear manipulation.[12] This paper represents that with the assistance of helpful software and resources, patients can easily access medical records, track and interpret health conditions it is particularly helpful for aged patients. Voice assistants are gaining popularity in home healthcare, enhancing patient care by increasing productivity and providing an innovative performance by using a broad set of resources. This serves as a proof-of-concept for a voice-enabled healthcare system that aids in continuous monitoring of a person's heart condition. Technology, on the other hand, is both an obstruction and a helper hand, to the elderly people who are most in need of heart related treatments.[13] An android-based personal assistant app for mobile phones that supports voice control in Serbian language. The objective of this application is to enable simple, natural communication between the phone and the user by providing a set of critical commands for quick and efficient device use in a variety of tasks, such as texting, calling, and manipulating contacts. In today's world, finding a need for such an application is easy. One of the most key components for this application is that it performs well regardless of the user or the environment. Since the circumstances are unlikely to be optimal all the time when the method is used, the device is resistant to the factors such as noise and speech characteristics.[14] Many devices and systems now have voice assistants. Measuring the consistency of the user interface is therefore crucial for evaluating and improving voice assistants. In past few years, this technology has advanced to the forefront of a growing number of sophisticated voice service applications. The healthcare, automobile, authentication and identification, voice commerce and customer support, and smart home industries are among the most common use cases. User experience can be an effective way to assess the efficiency of voice assistance systems. The value of such metrics is demonstrated by the wide variety of applications and promising potential prospects for these systems in both the corporate and the consumer sectors.[15] Here the aim of this paper is to evaluate the effectiveness of brands use of voice-activated digital assistants in customer communication. Communication is limited to sharing clues or conducting any operations necessary to achieve a particular result. Voice assistant's ability to mimic emotional and social responses allows an individual to construct contact with them according

to the established patterns. As a result, an investigation was conducted that uses user experience analysis to expose the nature of a brand's communication issues, that engages in a dialogue with customers through voice assistants, and it determines the shortcomings of this communication. [16] The principle of voice assistant operations are described in this paper along with the main shortcomings and limitations. A method to create voice assistant without the use of cloud services is described, which allows for the future applicability of such devices to be significantly expanded. The use of local systems expands the range of tasks in which they can be used in IoT systems, smart home systems, healthcare, security, and systems with an embedded processor. [17] Here the application was created to make email writing easier to every user. The proposed work aims to create a mechanism that converts Speech to Text (STT) for email content and Text to Speech (TTS) for email reading for speech recognition, this application makes use of the Google Web Kit API. The application's effectiveness has been demonstrated through research. Email has become a common tool in both professional and personal life. This application mainly consists of 3 modules, Inbox, sent mail and trash. Inbox consists of all received mails. Sent mail, where the conversion of speech to text takes place. Finally, trash, which has track of all deleted mails. [18] This paper was funded by the European Commission's Horizon 2020 Programme, as a part of the OPERANDO project based on work from COST Action CRYPTACUS, as well as European Cooperation in Science and Technology. In this paper smart mobile devices are equipped with numerous sensors; they enable applications to be context-aware. To enhance user experience even further, most mobile operating systems and service providers are gradually moving towards smart devices with voice-controlled intelligent personal assistants, entering a new level of human and technological interaction. This paper talks about potential risks about voice assistant. According to their independent research, voice assistant poses a very real and dangerous threat, as it is enhanced by the underlying mechanisms of the fundamental operating systems, also this paper briefly discussed and demonstrate how to launch attacks through voice assistant, as well as analyse their effect in real-world scenarios. [19] This paper tells us about the development of control systems in the electronic devices like monitoring and control systems using voice commands. Google Assistant was one of the largest and most comprehensive interface platforms. It was a fantastic opportunity because of its widespread use and small size, Google Applications were created. It was made possible by Dialog flow API (Application Programming Interface), which was created to process human communication using artificial intelligence. A cell phone was needed to check and monitor the appliance through voice communication. The sound was sent from the phone to the Google Assistant server via the internet. The server changed the voice into content and the Dialog flow API took over it. The material was prepared using this API in accordance with the topic characterization or strategy. The test results showed that developing an electronic gadget application based on voice control had a lot of potential. It provided 75% success rate. [20] This paper shows us that the home automation is gaining a lot of popularity in people's life because of its convenience. The

aim of this study was to monitor and protect various electrical appliances in critical situations while also efficiently utilizing the power, it also explains about the fire alarm system and also shows us how we can access house hold appliances with the help of smart phones. Here is the explanation about how it works, at first the input is given to Alexa through voice commands and then the Alexa decodes the input given by the user and sends it to IFTTT (IFTTT is a software that connects different applications through different developers to integrate multiple services of automation) and inside the IFTTT there are several applets which are wirelessly connected to esp8266 (esp8266 is useful for building IOT products) when the esp8266 receives the command it operates light or fan and other home appliances and in case of any fire incident or smoke in the house the sensors scans the temperature of the home and sends the command to the registered user through the IFTTT web cloud service. In this application we can operate all the hardware setup through real time database mobile appliances.

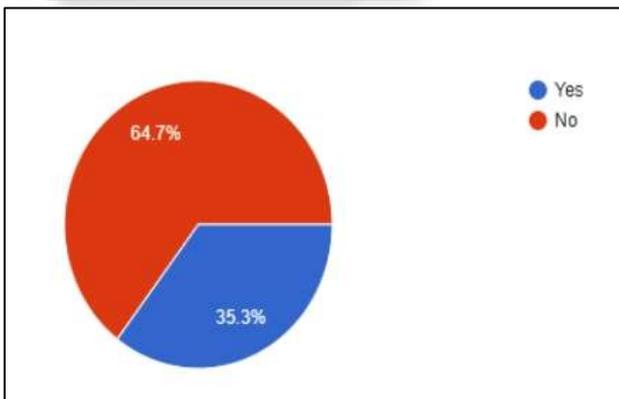
IV. RESULTS

We conducted a survey on the voice assistant to obtain the collective information among the people about the voice assistant. The main aim of this survey was to determine opinion of the people about voice assistants.

The Questions were:

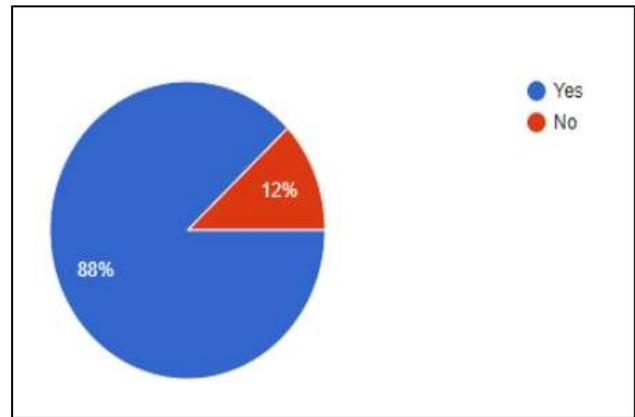
A. Are You Concerned About the Privacy Issues Regarding Voice Assistants?

For this question 132 people responded and out of those, 64.7% people responded that they do not have any privacy issue regarding voice assistants. Few of them responded that they have privacy issues regarding voice assistants and that collected out to be 35.3%. Majority of the people responded that they do not have any kind of privacy issue with the voice assistant.



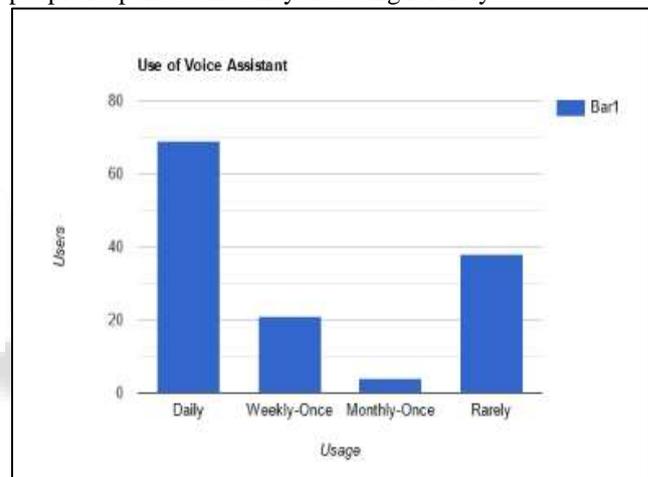
B. Do you use voice assistant?

For this Only 12% of surveyed people responded that they were not using voice assistants and 88% of the people responded that they are using voice assistants which shows us that the voice assistants are popular among people and also, they are well known. Today's voice assistants are pretty powerful, allowing us to make general web queries, send messages, make calls, set reminders, access music controls, and more.



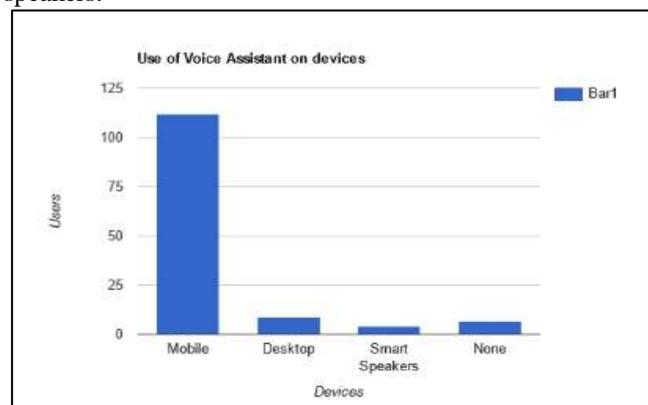
C. How Often Do You Use Voice Assistant?

To this question more than 52.6% out of 132 people responded that they are using voice assistant daily in their life which shows the significant of the voice assistant among the people's life .15.8% people responded that they are using voice assistant once a week and remaining 28.6% of the people responded that they are using it rarely.



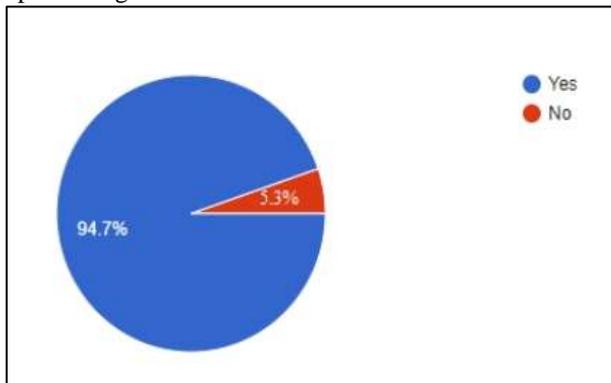
D. On Which Device Do You Use Voice Assistant Feature More Often?

For this question 85% people responded that they are using it in mobile phones as it is very easy to handle and to carry. 6.8% of the people responded that they are using it on desktop and 3% of the people responded that were using it on smart speakers.



E. Do You Think Voice Assistant Is the Future?

For this majority of the people i.e. 94.7% agreed that the voice assistants are the future and remaining 5.3% people responded against it.



V. CONCLUSION

All the above papers give the idea of how voice assistants are helpful in our daily day to day activities and it also shows us how voice assistants are built using different technologies and different modules. The demand for Voice assistants is growing day by day and it is also advancing at high rate in healthcare industries and also in household appliances, thus creating a lot of importance in people's life. As a result of voice assistant industrial and economic landscapes are changing drastically. Voice assistant will become an essential part of human existence, as computers and smartphones in future. And further upon conducting more survey and gaining more knowledge about voice assistants, we conclude that voice assistants can be helpful for people throughout the world.

REFERENCES

- [1] JARVIS: An interpretation of AIML with integration of gTTS and Python. Published year:2019 Accessed year:2020
- [2] ABYS (Always by Your Side): A Virtual Assistant for Visually Impaired Persons. Published Year:2019 Accessed Year:2020
- [3] Artificial Intelligence-based voice assistant. Published Year:2020 Accessed Year:2020
- [4] Domain Specific Intelligent Personal Assistant with Bilingual Voice Command Processing. Published Year:2018 Accessed Year:2020
- [5] Cloud-Assisted Mobile Desktop Search. Published Year:2018 Accessed Year:2020
- [6] Study on Relationship between QoS Degradation and QoE for Web services with Personal Digital Assistant. Published Year:2016 Accessed Year:2020
- [7] A Decision Assistant Based on Evidential Reasoning. Published Year:2006 Accessed Year:2020
- [8] Usage-Based Learning in Human Interaction with an Adaptive Virtual Assistant. Published Year:2019 Accessed Year:2020
- [9] Raspberry Pi based voice-operated personal assistant (Neobot) Published Year:2019 Accessed Year:2020
- [10] Short Research on Voice Control System Based on Artificial Intelligence Assistant. Published Year:2017 Accessed Year:2020
- [11] Voice Output Communication Aid Application for Personal Digital Assistant for Autistic Children. Published Year:2012 Accessed Year:2020
- [12] Interactive home healthcare system with integrated voice assistant. Published Year:2019 Accessed Year:2020
- [13] Voice Assistant Application for The Serbian Language. Published Year:2015 Accessed Year:2020
- [14] Measuring User Experience Quality of Voice Assistants. Published Year:2020 Accessed Year:2020
- [15] The Voice-Activated Digital Assistants in the Brand Consumer Communication. Published Year:2020 Accessed Year:2020
- [16] Investigation and Development of the Intelligent Voice Assistant for the Internet of Things Using Machine Learning. Published Year:2018 Accessed Year:2020
- [17] Human computer interaction (HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU) Published Year:2020 Accessed Year:2020
- [18] Monkey Says, Monkey Does: Security and Privacy on Voice Assistants. Published Year:2017 Accessed Year:2020
- [19] Voice-Based Monitoring and Control System of Electronic Applicants Using Dialog Flow API Via Google Assistant Published Year:2020 Accessed Year:2020
- [20] Power Effective Smart Home with Voice Assistant Published Year: 2020 Accessed Year:2020