INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

A Review of Voice Based Personal Assistants

Sohan Madalli[1], R.Manikandan[2], Vishal Pandita[3], Nishikant Surwade[4], Bahubali Shirgapur[5]

1. 2. 3. 4. 5 E&TC Department, Dr D Y Patil School of Engineering, Pune, Maharashtra - India

Abstract: Internet has developed over the years to make information accessible quickly and easily, revolutionizing how people communicate and interact with the world. Innovations in computing field has made technology more advanced and thus cheaper, nowadays it can be built into an increasing number of devices of various kinds. Personal Assistants are getting smarter by eliminating the traditional way of information access, storage and manipulation i.e. carrying out tasks such as saving notes, telling you the weather, notifications from social media, playing music, setting alarms, navigation and much more. Tremendous growth in technology have enabled us to connect the digital world to physical world, but as of now there are not much techniques to bridge the gap between digital world and physical interaction with the real world. Thus, this review gives a brief description to bridge this gap using voice based personal assistant.

Keywords: Personal Assistants, IPA,

I. INTRODUCTION

As we know, time is the sole factor distributed equally among all creatures. Due to proper utilization of time and resources only a few are able to taste success, whereas others are so busy with their activities that they have inadequate time to grab a golden opportunity. One of the most crucial quality that separates former from the latter is time management.

In this era where computers play an inevitable role in daily lives, an attempt to make their use much easier has been made. Here, is where voice based Intelligent Personal Assistant (IPA's) comes into picture. IPA's or chatbots are software that are meant to simulate conversation with a human in an instinctive manner. Chatbots mostly refer to databases so as to respond to the users query, in fact the communication takes place in a more humanely manner. Hence, an idea to evolve a user-friendly, personal assistant has been scaled to make the use of electronic gadgets more efficient and quicker.

In this study, we are trying for a review of all those available IPA's. Each of these technologies used in IPA has have some advantages along with some disadvantages. Based on this study, we can determine which personal assistant is suitable for making our life simpler.

II. RELATED WORK

Intelligent Personal Assistants (IPA) and their current software implementations exists in a vast and varied of applications, usually integrated in various Operating Systems (OS), like in personal computers, mobiles and in Internet of Things (IOT). Technology is constantly developing and changing aspect of learning.
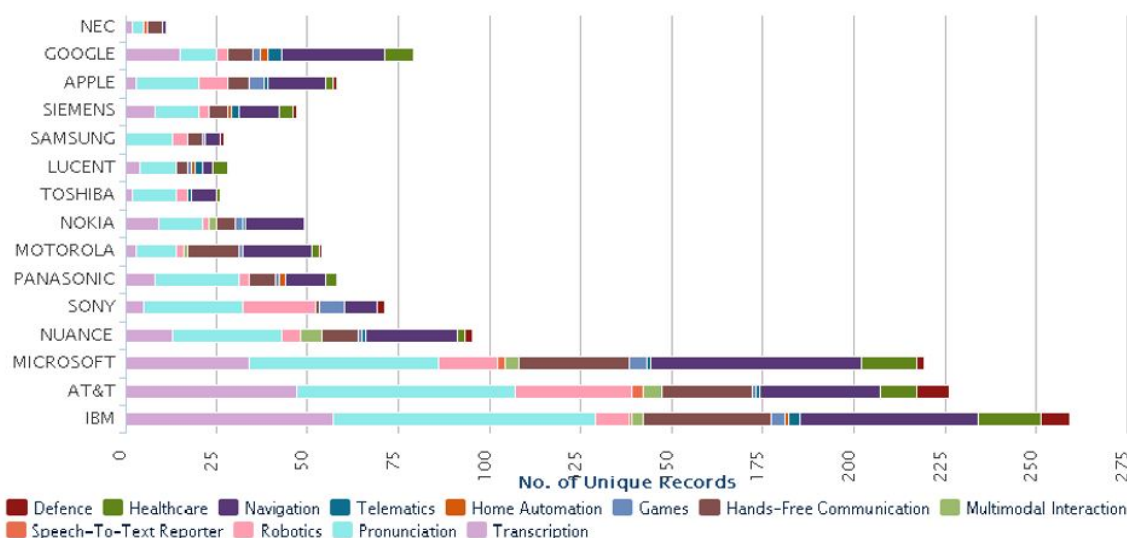

Fig. 1 Top companies across different Applications [1]

Above figure shows that Virtual Personal Assistants are significantly growing across all fields depending upon the operating sector (from defence to robotics).

Siri [2] (Speech Interpretation and Recognition Interface) is a part of Apple's iOS operating systems. Creators of Siri were the first to develop an artificially intelligent (AI) assistant and embed it into a smartphone. Under continuous upgradation Siri, uses an AI technique called deep learning to help it interact with humans more humanely: deep learning helps Siri monitor user's activities, so that it can give some personalised recommendations. It has some disadvantages like it is a proprietary software, which restricts developers to modify its code.

Google Now [3] emerged into the AI prowess in 2012. Android and iOS platforms have started using Google Now assistants. Machine learning (ML) is the base for Google Now, the devices using Google now become smarter depending upon the user's activities. Through ML, suggestions are provided based on user's frequent locations, booking a table in a restaurant etc. It has a more structured architecture than Siri. Its response to user's queries is quite accurate.

TABLE I

Top five speech recognition requests and proportion of each domain in a sample of Cortana user logs. [4]

Chat (21.4%)	Device Control (13.3%)	Communication (12.3%)	Location (9.2%)	Alerts (8.7%)	Weather (3.8%)
Tell a joke	Play music	Call	Where am I	Set alarm	In Celsius
Do you like Scooby-Doo	Show me live cricket scores	Call mom	Find a dentist	Show my alarms	Do I need a coat
Hello	Open Calendar	Call my wife	I'm hungry	Wake me up	What's the weather
Sing a song for me	Open WhatsApp	Text	Where I am	Wake me up in 30 minutes	What's the weather like
What's your name	Stop music	Tell me a recipe for chocolate cake	Take me home	Remind me	Will it rain today

Due to stiff market competition from Android and iOS, Microsoft created Cortana [5]. First launched in April 2014, it is now compatible on a wide range of devices from smartphones to laptops. Cortana has all functionalities which other IPA's provide, but what makes different in the market, is its transparency. Cortana is capable to read a user information, such as emails and give a personalised suggestions however, this task is done only after the consent of the user. This helps to confirm if the software is anticipating the user's needs correctly, and provides better privacy for the user. Cortana takes advantage of Microsoft's domination in the PC market, where users can have a fully opt mixed experience across different devices, including PCs, smartphones and tablets that run Windows 10. Major drawback in Cortana is that whenever a query is requested it performed a Bing search that isn't very accurate.

Amazon developed Alexa [6] solely for Amazon Echo devices to function as a personal assistant. Alexa can have a casual talk, play music, and read out news bulletins, makes to-do lists and much more. Additional feature that give Alexa a plus point is that it can used in home automation. Alexa can be merged to expand its potential by installing skills or installing apps. Developers can always use different tools available in Amazon for developing innovative applications. Alexa always needs apoer source for its operation as it doesn't have built-in batteries.

IBM pushed its boundaries far with the Watson [7] supercomputer developed in DeepQA which can reply to queries posed in natural language. Instead of launching its own AI human interface that interacts with consumers, it provided a platform for all the developers and start-ups to develop apps on IBM Watson Development Cloud. It features a dashboard that can be personalised according to users taste.

Bixby [8] a rebooted version of "S Voice" is only made available for Samsung's flagship phones. It includes Bixby Voice, Bixby Vision and Bixby Home. Devices can be triggered by calling out the keyword "Bixby", that is a part of Bixby Voice. Bixby Vision enables user to experience augmented reality. Additional features of Bixby are to translate text, recognize landmarks and scan QR codes. Various other widgets like weather, fitness tracker location etc. are provided by Bixby Home.

Mobvoi [9] was established in 2012, the company's Chumenwenwen (mobile voice search app) is a leading Chinese AI personal assistant. Google is a minor shareholder in the company. It gives accurate response using vertical search. Its watch could be triggered by saying "Hi Watch". The app is available on iOS and android for wearables such as Google Glass and Android Wear. The company also developed its own smartwatch "Ticwatch" which is meant to work with with iOS and Android devices.

Hound [10] by Soundhound has gained attention among those in the industry as it comprehends natural language and is able to handle sophisticated commands. Additional features that Houndify provides is emotion detection, iris detection and fingerprint identification.

Mycroft [11] is tagged as the world's first platform independent virtual assistant. It is basically a free software (open source) which can be altered based upon user's needs, could be combined with other projects or could be integrated with desktops and Raspberry pi.

Aditya Sinha et al. [12] presented a virtual voice based intelligent assistant for visually disabled users. This process is as follows: input voice first goes through speech recognition further on speech synthesis is carried out. Later on content extraction takes place after which the response is given back to the user. This project uses Java Sphinx libary for Speech analysis and text-to-speech part is performed using MaryTTS. It also uses neural networks that helps the user to improve task performance by its learning capacity.

Rishabh Shah et al. [13] suggested a chatbot based on Natural Language Processing (NLP). This paper shows educational systems that requires natural acquisition in the learning process. Due to inaccessibility of education this system has given a solution to the problem. This system incorporates tokenization of sentences and then query extraction based on N-gram division algorithm. This metadata is searched into its knowledge base and if a match is found then information is retrieved and given to the user.

Johann Hauswald et al. [14] proposed Sirius an application which takes input in the form of speech and pictures. Sirius is a database related application. It emphasizes on the design space for server architectures and also highlights use of FPGA's, CPLD's and GPU's. Some of the new modules in this system are ASR, IMM and QA. Using statistical models speech is translated into text. What makes Sirius different is its computer vision techniques that attempt to match the input image to its image database and return relevant information about the matched image using "The Image Matching" (IMM) service.

Er. Sheeba Babu et al. [15] devised "MIA-My Intelligent Assistant" that is meant to answer any questions raised. The questions supported are within the arithmetic, logical, and general category. MIA is brought to life using Raspberry Pi, and at its centre is Wolfram Alpha Search Engine which responsible for its intelligence. Speech synthesis is performed using ALSAMIXER software embedded into the Raspbian OS and is converted into.wav file which is further sent to Google API for conversion into text. Next, this text is fed as a query to Wolfram engine which thereby responds to the users query using TTS module of Google API.

Namita Mhatre et al. [16] proposed "Donna – an interactive chatbot". This paper devises a unique method which stimulates setting and initiates meetings of user with his clients that is done by web-based chatbot. Using pattern-matching algorithm, this chatbot schedules meetings between users by gathering some keywords in there calendar or to do lists.

Chirag Joshi et al. [17] designed a speech recognition system for visually disabled users. This project uses Jasper software for its implementation. Jasper is basically used for building controlled applications based on voice. Jasper is easy to integrate and configure with Raspberry pi as it specially operates for Raspbian operating system. Python Scripts were used for its operation to build several modules, for recording the sound via microphone and transferring the data to cloud for query processing. Next, queries were answered back to user using TTS through speaker.

Poonam Patil et al [18] developed AlexaPi a smart virtual assistant using Amazon Ecosystem platform. This paper showcases ways to monitor events of Home Assistant based on several conditions of "if and else". AlexaPi has been integrated with Amazon Web Service (AWS) and is based on Raspberry pi. AWS serves as a medium for processing of natural language and content extraction.

III.CONCLUSIONS

Convergence of communication technologies and information have led into developing machines that understand us, do what we tell them to, and even anticipate our needs. This review paper gives a glimpse of different IPA's available in the market. The application of IPA's will simplify the process of using a device (computer or any handheld device), by reducing the need of mouse or keyboard in certain cases and it makes the process a little more acceptable to most people who are aged as they are not acquainted with new age technology. Furthermore, the whole process of launching programs, keeping reminders etc. will go at fast rates thereby saving the user's time and in turn improving their productivity. This application will also enable disabled user's (e.g. users who can't use the keyboard/mouse or users without sight) use a computer.

Therefore, in the next few years Machine Learning along with AI would be integrated with personal assistants to make them work in a more humanely manner.

REFERENCES

[1]   Voice Recognition webpage– Controlling mobile devices using voice command. [Online] Available: http://patseer.com

[2]   Apple Siri webpage. [Online]. Available: https://www.apple.com/ios/siri

[3]   Google Now webpage. [Online]. Available: https://www.google.com/search/about.

[4]   Jiepu Jiang, Ahmed Hassan Awadallah, Rosie Jones, "Automatic Online Evaluation of Intelligent Assistants", Tech. Rep.

[5]   Microsoft Cortana webpage. [Online]. Available: https://www.microsoft.com/en-us/windows/cortana

[6]   Amazon Alexa webpage. [Online]. Available: https://developer.amazon.com/alexa

[7]   IBM Watson webpage. [Online]. Available: https://www.ibm.com/watson/developer/

[8]   Samsung Bixby webpage. [Online]. Available: https://www.samsung.com/us/explore/bixby/

[9]   Mobvoi webpage. [Online]. Available: https://www.mobvoi.com/pages/about-us

[10] Hound webpage. [Online]. Available: https://www.houndify.com/

[11] Mycroft –AI webpage. [Online]. Available: https://mycroft.ai/

[12] Aditya Sinha, Gargi Garg, Gourav Rajwani, Shimona Tayal, "Intelligent Personal Assistant". International Journal of Informative & Futuristic Research, Vol. 4, Issue 8, April 2017.

[13] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, "An Intelligent Chat-bot using Natural Language Processing". International Journal of Engineering Research, Vol. 6, pp. 281-286, 1 May 2017.

[14] Johann Hauswald, Michael A. Laurenzano, Yunqi Zhang, Cheng Li ,Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci1, Lingjia Tang, Jason Mars, "Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers".ASPLOS'15, March14–18,2015, Istanbul, Turkey.

[15] Er. Sheeba Babu, Akhil S Nair, Akhil S R, Allan Abraham, Ashish Anil Abraham. "MIA-My Intelligent Assistant". International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2016.

[16] Namita Mhatre, Karan Motani, Maitri Shah, Swati Mali, "Donna Interactive Chat-bot acting as a Personal Assistant". International Journal of Computer Applications (0975 – 8887). Vol. 140 – No.10, April 2016.

[17] Chirag Joshi, Pradnya Jawale, Dr.Milind Nemade, "Speech Recognition system for Blind people using Raspberry pi and Jasper platform". International Journal of Engineering Sciences & Research Technology, May 2017

[18] Poonam Patil S, Rudrappa B Gujanatti, "AlexaPi on Amazon Ecosystem for Home Assistant Environment and IFTTT recipes". International Research Journal of Engineering and Technology, Vol. 04, Issue. 08, Aug -2017.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⓒ (24*7 Support on Whatsapp)