THINK INDIA JOURNAL

ISSN:0971-1260

Vol-22-Issue-41-December-2019

Current State of Virtual Personal Machines

Sahejpreet Singh Anand

Information Technology, Bachelor of Technology, Mukesh Patel School of Technology Management & Engineering Shirpur

ABSTRACT

This report discusses about a Virtual Personal Assistant (VPA). It discusses the working of an VPA in brief. Here we discuss how data is received from the user and how we process that data to perform the required tasks. The report also discusses the industry leaders of VPA that use Natural Processing Language (NPL) and Deep Learning to process data. It will show what various functions and services are provided by each of those leading VPA's and how they lack. The report discusses the history and the current advancements in the field of VPA. We also see how the mindset of the people changed over VPA and challenges it faced to succeed. The report also discusses the risks involved in using a VPA and some ways to overcome those risks.

INTRODUCTION

As the world is moving towards smart future, no data is of waste and can be used in one or the other way. We are creating new technologies which process our data to gives us required results. Virtual Personal Assistants (VPA) is a software application program that uses Natural Language Processing (NLP) to understand the natural language voice commands and then execute the commands based on them. Generally, VPA's use three methods of interaction:

1) Text (Used in Messaging apps and Online Chat bots).
2) Voice (Used in Amazon Alexa, Apple Siri and Google Chrome Cast).
3) Photos (Used in Samsung Bixby).

Virtual personal assistants are generally called passive listeners which basically means they listen to every sound you make until they hear their familiar command or greeting also known as "wake" word. This concept creates a privacy concern in the minds of the user as all the data is stored. Many companies such has Facebook have implemented end to end encryption where a user's privacy is guaranteed. Other companies such as Amazon, Google and other have also started to implement end to end encryption in their data storage services. The VPA must be connected to the internet as it continuously searches data and uses it to train itself and hence execute commands faster and with more accuracy. There was a reported case in January 2017 where an ECHO (Amazons smart speaker) was a witness to a double murder. This bought some appreciation for the VPA's, but many are still concerned about its use.

There are some things that an VPA must perform, functions of VPA are:

1) Adding, viewing and modifying events.
2) Add things in To-do list.
3) Control smart home devices.
4) Make and receive phone calls.
5) Get directions.
6) Weather forecast.
7) Book means of transport.
8) Book an appointment.
9) Send E-mails.

ISSN:0971-1260

Vol-22-Issue-41-December-2019

10) Open webpages.

In this world everyone is busy, no one has the time to do things manually. Artificial Intelligence is at its highest and is constantly going up. Someone wisely said that, "AI is the best and the last invention of mankind". Any random nobody can now impact the industry and make his mark in the whole world. AI is a growing field which makes it easier for upcoming researcher to get noticed easily. AI is getting the attention it truly deserves. Researchers are now valued a lot in the industry. We hear salaries of millions of dollars given in the field of AI, which in my view is well deserved. The most expensive thing that money can't buy is time, if the technology saves your time and performs functions as per your needs, you will use it. VPA is that technology that is much underappreciated right now but that is due to the knowledge gap.

Fields where VPA has achieved the most are:

1) Self-driving cars.
2) Healthcare.
3) Finance and banking.

METHODOLOGY

VPA's being passive listeners must understand which data is command and which is not. This is generally tackled using a "wake "word. "Hey Siri", "Alexa", "Ok google", "Cortana", etc. are some examples. Alexa gives an option to change your wake word which is cool as it can be changed as per your needs. Wake word is essentially the start of your conversation with the VPA. There are some things to be considered when choosing your own personal wake word. They include:

1) Avoid single word phrases.
2) Avoid long phrases.
3) Choose words with distinct sounds.

Once the "wake" word is identified, the important part of the VPA now is how effectively and efficiently can it detect the input. The working of an VPA can be broken down in three stages:

1) Convert Speech to text.
2) Understand intent from the text.
3) Perform the action in response.

Speech Recognition, Spoken Language Understanding, Dialog Manager, Database, Natural Language Generator and Text to Speech are the series in which voice commands are processed.

The first step is to convert voice command to a text input that your device can understand. Speech recognition (ASR) can be done using various Speech to Text software's like Braina Pro, Speech Notes, Voice Finger, Apple Dictation, Google Docs voice typing, Dragon Professional Individual, Gboard, ListNote, Speech Recognizer, etc.

The first job done is that the speech is broken into tiny parts known as phonemes. These are total 44 phenomes. This allow the software to recognize the words said. The words that sound similar like two and to are differentiated by analyzing the syntax and the whole context of the sentence. That is why sometimes there are mistakes in analyzing words that have similar pronunciation.

Most VPA's use Voice Biometrics to authenticate the user to its services. Voice biometrics works by making a digital profile of a person's speech to make a personalized voice print or template. To ensure the accuracy of the biometrics we are generally asked to say some phrases or series of numbers. This is repeated until a template of the user is made and stored in the database. When the user tries to access the VPA, his phrase is compared to his premade personal template and is verified. Noise reduction is used to determine the speech clearly as there is hardly any quiet places these days. Speech compression engine is used to make the speech

ISSN:0971-1260

Vol-22-Issue-41-December-2019

a better traveler which reduces Round Trip Time (RTT). RTT is the time taken for the message to travel from the sender to receiver and back.

The second part is text to intent which interprets what the user wants to say. Syntax and Semantic processing are done here. Syntax analyzing also known as parsing is using knowledge about grammar to analyses and identify structure or formation of sentences. Semantic analysis is understanding the meaning of those sentences formed after parsing. Pragmatic analysis is the process of extracting meaningful information from text. Pragmatics analysis focuses on the finding the actual meaning or context of the text. Spoken language understanding (SLU) and Dialog manager is used like this.

The third part is intent to action which basically means doing what the user desires. Natural language generator (NLG) and Text to speech (TTS) is used to generate the voice and speak the intended information or in other cases give confirmation about the work being complete.

IMPACT ON SOCIETY

VPA is one of the most important part of technology that has impacted the AI industry at large. The start of VPA went unnoticed, its name is Radio Rex. It was the first ever voice activated toy released by Elmwood Button Co. Radio Rex is a great piece of speech-recognition history, but it has a terrible false-rejection rate. This was due to the non- accepting nature of human beings. That technology was way ahead of its time.

Nowadays, the technology biggest concern is user privacy. VPA are highly judged based on its basic requirement i.e. passive listening. They are always listening to every conversation whether it is with them or with others. There are various measures that can be taken to ensure privacy. They are as follows:

1. Using a good wake word to avoid unexpected activation.
2. Turning off VPA when not in use.
3. Not providing passwords and other personal information.
4. Erase old recording.
5. Increase Security checks.

REFERENCES

1. https://en.wikipedia.org/wiki/Virtual_assistant
2. https://searchcustomerexperience.techtarget.com/definition/virtual-assistant-AI-assistant
3. https://www.zirtual.com/what-we-do/virtual-personal-assistant
4. https://www.predictiveanalyticstoday.com/top-intelligent-personal-assistants-automated-personal-assistants/
5. https://www.calcalistech.com/ctech/articles/0,7340,L-3741017,00.html
6. https://www.lifewire.com/virtual-assistants-4138533
7. https://www.theatlantic.com/technology/archive/2016/05/the-privacy-problem-with-digital-assistants/483950/
8. https://nakedsecurity.sophos.com/2017/01/27/data-privacy-day-know-the-risks-of-amazon-alexa-and-google-home/
9. https://www.eweek.com/security/five-ways-digital-assistants-pose-security-threats-in-home-office
10. https://www.futurity.org/metabolomes-store-data-2098912-2/
11. https://www.independent.co.uk/life-style/gadgets-and-tech/news/amazon-echo-alexa-evidence-murder-case-a8633551.html
12. https://picovoice.ai/docs/choose-wake-word/index.html

ISSN:0971-1260

Vol-22-Issue-41-December-2019

13. https://www.globalsecurity.org/security/systems/biometrics-voice.html
14. https://www.marshall.usc.edu/blog/how-do-digital-voice-assistants-eg-alexa-siri-work
15. http://scienceline.org/2014/08/ever-wondered-how-does-speech-to-text-software-work/
16. https://www.oreilly.com/library/view/python-natural-language/9781787121423/03d0bf64-f5c3-474a-904e-a4781b02606e.xhtml