

Review Article

Virtual Voice Assistant using Python

Ayush Agarwal¹, Abhineet Mishra¹, Saransh Tiwari¹, Rajan Singh¹

¹Department of Computer Science & Engineering, University Institute of Engineering, Chandigarh University, Punjab, India.

I N F O

Corresponding author:

Ayush Agarwal, Department of Computer Science & Engineering, University Institute of Engineering, Chandigarh University, Punjab, India.

E-mail Id:

19bcs1270@cuchd.in

How to cite this article:

Agarwal A, Mishra A, Tiwari S et al. Virtual Voice Assistant using Python. *J Adv Res Info Tech Sys Mgmt* 2022; 6(1): 25-29.

Date of Submission: 2022-03-10

Date of Acceptance: 2022-04-20

A B S T R A C T

In recent years, Artificial has shown enormous growth and future potential. One of the specific areas where that potential is growing each year is Natural Language Processing (NLP). With the help of artificial intelligence, voice assistants can communicate with users in natural language. They are extremely easy to use, and in today's growing digital world, their use is expected to continue to grow. It is no exaggeration to say that everything from home speakers to pocket-type smartphones has become established in everyday life. The purpose of this study is to understand the increasing use of voice assistants in our daily lives.

Keywords: Artificial Intelligence, Voice Assistant, Home Speaker, Smartphone

Introduction

Today we are living in the silicon world where everything we use in our daily life is in some way dependent upon this glorious invention. With the rise of the internet a new field emerged known as Artificial Intelligence. For centuries humans have been intrigued by the idea of creating an artificial machine which will do our basic tasks and this dream of ours may very well become a reality for all of us with the rise of Voice Assistants such as Google Assistant, Apple's Siri, and Amazon's Alexa. From simple math questions to searching for files and documents on the Internet, those limits are getting smaller every day.

In recent years with the rise of the smartphone economy every major tech corporation in the world is focussing their attention and resources towards creating a human-like voice assistant which can solve our everyday queries by just saying them to our device. Following are some of the tasks that VA(voice assistant) can do with ease:

- Setting timers or alarms
- Playing music
- Playing videos or any multimedia file
- Making calls and sending messages

- Giving weather information
- Answering queries given by the user
- Controlling other smart devices

Some important factors that distinguish voice assistants from simple programs are:

NLP: Ability to understand and process human language.

Machine Learning: Ability to identify patterns and adapt to new things.

Ability to understand stored information and use data, new conclusions using them is derived.

Methodology

The snowball technique by Wohlin (2014) was used to obtain sufficient and high quality work on the use of language assistants. This technique consists of the following steps:

1. First, perform a search on Google Scholar, IEEE Xplore, Scopus, and ACM Digital Library to collect the first starting group for related articles. The keywords used were "Voice Assistant," "Smart Speaker," "Amazon Echo," "Google Assistant," "Alexa," and "Siri."
2. In the first set of paper, repeat the snowballs before and after. Backward snowballing uses a reference

list to identify new articles to include, and forward snowboarding is related to identifying new articles that cite the article under review. With reverse and forward snowballs, the new papers identified at each step are piled up and proceed to the next iteration.

Home Usage

Childrens: Many studies have been conducted on how children over the age of 14 use voice assistants. A study was conducted by Beirl et al. (2019) regarding the usage of Alexa over the period of three weeks. The main purpose of the study was how the families use Alexa for their children. It was found that the main areas of interest by children towards voice assistants were storytelling, games and music.

Results showed that there was much enthusiasm towards alexa and it became part of their family ritual.

The study concluded that all of the above interactions contributed to social and emotional bonds, leading to further family cohesion. Another important finding in this study is that when young children are having a hard time following the rules of games and quizzes, families encourage and suggest why young children should say to Alexa. It played a role. The children also found that voice assistants were very helpful when playing video games and engaging in fun activities with friends. For example, in one instance a child aged 14 asked Alexa to order food for him and his friends while they were playing their favorite video game and according to him the entire process of ordering food via Voice Assistant took him only two to three minutes. It also created a deep impact in his mind that he can order everything via this newfound route only and now he did not require to open his phone to do this task.

Elderly: The use of voice assistants by the elderly (65+) is Savagoetal. Reported by (2019). The results of the surveys suggest that the wide range of features provided by voice assistants impressed participants. Participants also liked the idea that, as one participant mentioned, not only could there be mobility problems, but they also did not interfere with the process, so they could use only voice to complete certain tasks. rice field. It should be noted that voice assistants were primarily used as search engines, but more importantly, they were also used as translators and teachers. The authors also point out that further research is needed on the use of language assistants by older people.

Elderly people also use it to listen to their old songs and to search the web which according to them if done by their phone takes three times more time. Devotional songs are the top priority while searching for the songs with pictures about family coming in at close second position. Researchers also found out that this process increases their tendency to ask Siri or Alexa to do a particular task.

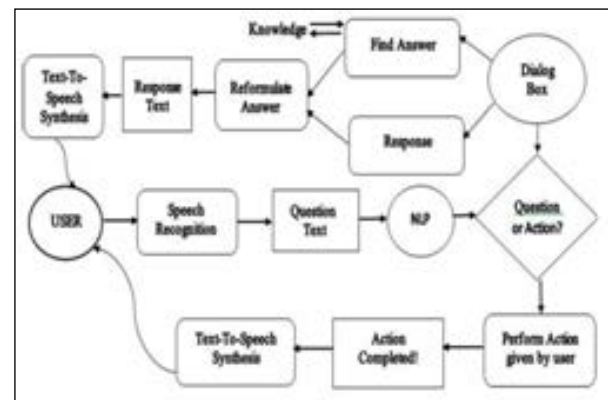


Figure 1. Proposed Model of Voice Assistant

Adults: Interesting results came from Sciuto et al. (2018) Here, the authors investigated how households integrate conversation agents into their lives. Specifically, the author analyzed the logs of 75 Alexa users to get a total of 278,654 voice commands. Participants who have owned an Alexa device for more than 6 months answered questions about using Alexa at home. Of the 75 participants, 26 reported having children, but the log file data did not tell which household member issued each order. Parents surveyed positively remembered how well their children interacted with Alexa, even before interacting with smartphones and other technology devices. In the Song (2019) survey, 433 adult participants completed an online survey to assess perceived usefulness, perceived ease of use, attitudes towards voice assistants, and behavioral intent to use them. did. The results suggest that the perceived usefulness has a significant impact on an individual's attitude towards language assistants and their behavioral intent to use this technology. In addition, consumers want to buy an easy-to-use device.

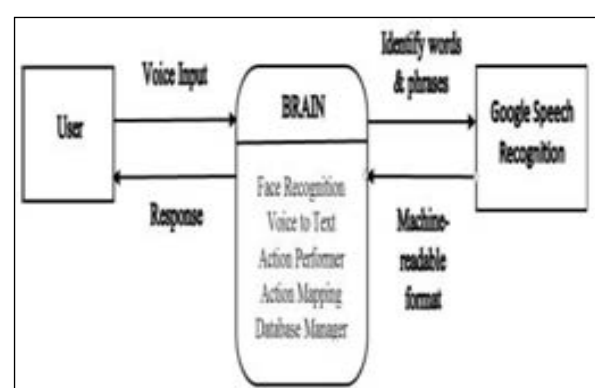


Figure 2.DFD Diagram Level 0

Data Flow Diagram

DFD is a graphic representation that provides a flow of information between input and output data. Also known as a "data flow diagram or bubble diagram". DFDs are often used as a preparatory step to create a system overview that can be explained in detail later.

Level 0 DFD: Users provide input in audio format. This voice command is recognized by the application. It is then checked to see if it is an authorized user and the action is taken according to the command specified by the user. The specified command is compared as an action and question form and answered in a dialog box or knowledge base search.

Level 1 DFD: Input is done by the user in audio format. The GoogleVoice API converts this voice data into text format, and the voice assistant takes action according to user-specified commands by comparing it to dialog boxes and knowledge bases.

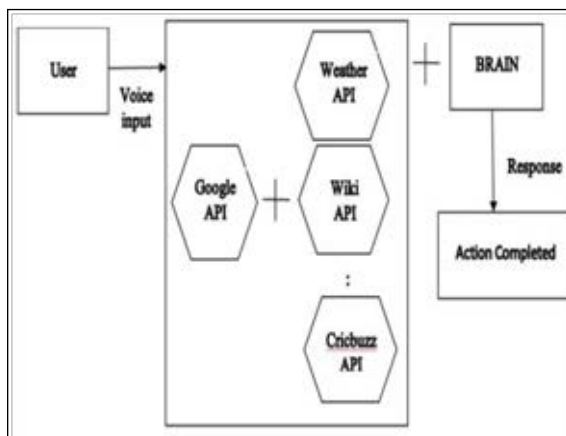


Figure 3

Flow Chart

The flow chart is a graphical representation of the algorithm. Various symbols are used to represent the flow chart. When the system boots, the first authorized user is authenticated, then the voice assistant continues to run in the background, waiting for available voice commands. When the user enters a command, the voice assistant provides the required output based on the conditions provided to the voice assistant. This output is sent to speech recognition, which converts the speech into a machine-readable format. Based on the input received, the Personal Language Assistant will perform the desired task.

Recommended System Features

Python provides a large standard library that includes areas such as Internet Protocol, string manipulation, web service tools, and operating system interfaces. Many commonly used programming tasks have already been written to the standard library, significantly reducing the length of code you write.

Python features a clean object-oriented design, offers advanced process control capabilities, powerful integration and word processing capabilities, and a unique unit testing framework, all of which is speed and productivity. Helps improve. Python is considered a viable option for building

complex multi-protocol network applications. The Text-to-Speech (TTS) system converts regular voice text to voice. Synthetic speech can be generated by concatenating the recorded speech parts stored in the database. The output is in audio format.

This voice assistant benefits a large number of users with universal eye-free and hands-free voice control on mobile devices. That framework will help shape future voice control devices.

Facial Recognition Technology (FRT) can eliminate the need for passwords, fingerprints and even keys.

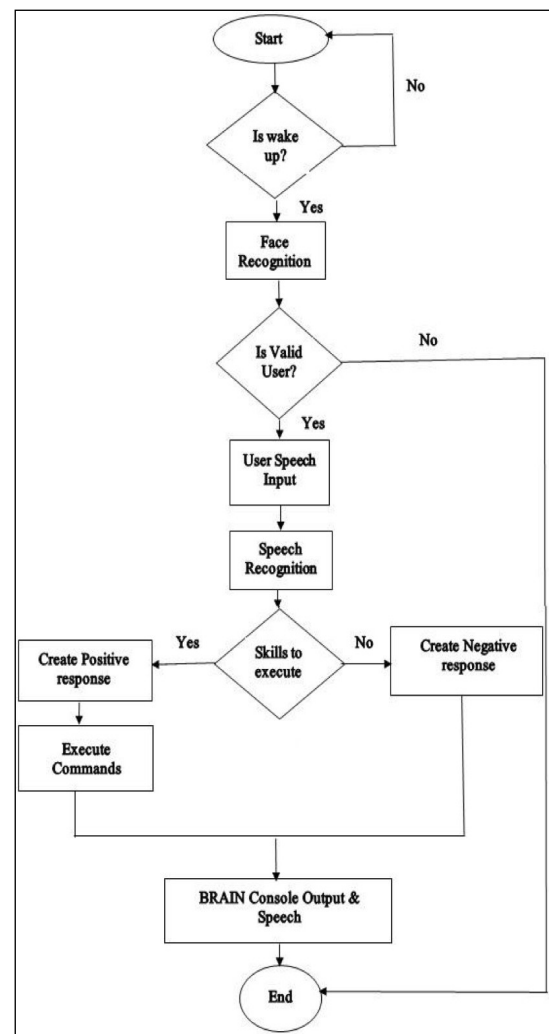


Figure 4

Security and Privacy Concerns

Although Voice Assistant acts as a great help in our modern life but with each blessing also comes a curse. Several concerns regarding the potential malicious use of voice assistant has been raised by many independent agencies and researchers.

Many users have been so dependent on this new technology that they are asking to do even very basic things from it.

Many non-users have stated that this new technology is not useful at all and it is just a fancy way for big multinational companies to make more money from their users. They found that the use of this technology is not as big as it is advertised. They also stated that they could have done the same thing by themselves with bare minimum effort.

Some parents reported that their children enjoyed using the voice assistant. Most of them adopted the technology early and bought the device to use it first. As a suggestion, the author proposes the ability to mute the device with different accounts and audio commands for children.

They also suggest that businesses need to introduce a secret mode similar to a web browser to prevent them from collecting data. McReynolds metal to address recording device privacy concerns. (2017) Propose indicators during recording.

The modality effect of voice assistants in the context of health information acquisition has been investigated by Cho (2019). Fifty-three undergraduate students at Southeastern University used mobile phones with the Google Assistant and Google Home Devices.

Participants were instructed to ask two types of health questions (ie, insensitive or sensitive health questions) using voice or text as input.

The results of showed that the interaction between voice and text significantly improved the perceived social presence over the voice assistant, but it contained less sensitive information and users reported privacy concerns. It was only when I did.

In contrast, when personal information confidentiality and privacy concerns are high, text was as effective as voice in evoking the human-like perception of voice assistants.

Furthermore, this result suggests that voice dialogue enhances the sense of social conversation between the user and the voice assistant, which leads to a positive evaluation of the voice assistant.

However, this only happened if the requested information was less confidential and individuals reported less privacy concerns.

Conclusion

Immersive technologies have truly changed how we communicate with each other in the 21st century. Voice assistants can provide new learning experiences but we also need to be wary about their downside. Research on this topic is fairly limited as it is a new technology that is in its baby stage and with each passing year we will get to know more about its upside and downside. One of the biggest challenges faced by voice assistants is its lack of availability of multiple languages which restricts its user base but many giant companies like google, Amazon

are already making big strides in solving this problem by investing tons of money in research and development. Adult, childrens and elderly people should be taught its advantages and how they can use it without causing much distress in their mental headroom. As a conclusion the role of this new technology is limited for now but as the world is moving towards becoming more digital we will witness one of the most transformative technologies within our lifetime.

Future Scope

Using the framework voice assistant technology can be updated further. Cyber security concerns are very real these days so it can be combined with the system to give further boost in security. Technologies such as voice authentication can be used to make sure that the command is given by the right user.

Work and advancements can also be made on how pitch and tone is judged. Accents from different regions can also be upgraded to make it truly a global technology. Further modifications like learning the answers of questions that are yet not known by voice assistant can be implemented.

References

1. Cho E. Hey Google, Can I Ask You Something in Private? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM 2019; 258.
2. Davie N, Hilber T. Opportunities and Challenges of Using Amazon Echo in Education. International Association for Development of the Information Society 2018.
3. Dousay TA, Hall C. Alexa, tell me about using a virtual assistant in the classroom. In EdMedia+ Innovate Learning 2018; 1413–1419.
4. Association for the Advancement of Computing in Education (AACE).
5. Ilhan K, Mušić D, Junuz E et al. Scarlet-Artificial Teaching Assistant. In 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), IEEE 2017; 11–14).
6. Juniper Research. Amazon Echo & Google Home to Reside in over 50% of US Households by 2022, as Multi-Assistant Devices Take Off 2017. <https://www.juniperresearch.com/press/press-releases/amazon-echo-google-home-to-reside>
7. Kukulska-Hulme A. Intelligent assistants in language learning: friends or foes?. In the World Conference on Mobile and Contextual Learning 2019; 127-131.
8. López G, Quesada L, Guerrero LA. Alexa vs. Siri vs. Cortana vs. Google Assistant: a comparison of speech-based natural user interfaces. In International Conference on Applied Human Factors and Ergonomics. Springer, Cham 2017; 241-250.
9. McLean G, Osei-Frimpong K. Hey Alexa... examine the variables influencing the use of artificial intelligent in-

- home Voice Assistants. *Computers in Human Behavior* 2019; 99: 28-37.
10. Nilsen JH, Røynealand K. «It knows how to not understand us!» A study on what the concept robustness entails in design of conversational agents for preschool children (Master's thesis) 2019.
 11. Rzepka C. Examining the Use of Voice Assistants: A Value-Focused Thinking Approach. 25th Americas Conference on Information Systems, Cancun, 2019.
 12. Trivedi N. ProblemPal: Generating Autonomous Practice Content in Real-Time with Voice Commands and Amazon Alexa. In e-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. *Association for the Advancement of Computing in Education (AACE)* 2018; 80-82.
 13. Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering, *ACM* 2014; 38). DOI: 10.1145/2601248.2601268
 14. Wiederhold BK. Alexa, Are You My Mom?" The Role of Artificial Intelligence in Child Development. *Cyberpsychology behavior and social networking* 2018; 21(8): 471-472.
-