

# Virtual Agent Using Artificial Intelligence and Python

IYER ANISH SUBRAMANIAN

*Keraleeya Samajam's Model College, Dombivli East, Mumbai, Maharashtra, India, First Semester, Department of Information Technology, University of Mumbai, Mumbai*

**Abstract**— An intelligent virtual assistant (IVA) or intelligent personal assistant (IPA) can be a software agent that performs tasks or services for a supported private command or question. Sometimes the term "chatbot" is used to refer generally or specifically to virtual assistants accessed through an online chat. In some cases, online chat programs are for entertainment purposes only. Some virtual assistants are ready to interpret human speech and respond through synthesized voices. Users can ask their assistant questions, control home automation devices and media playback by voice, and manage other basic tasks like email, to-do lists, and calendars with voice commands.

**Indexed Terms**— NLP, VLP, Artificial Intelligence, Speech-To-Text, Chatbot, Neural Network, Deep Learning, Machine Learning

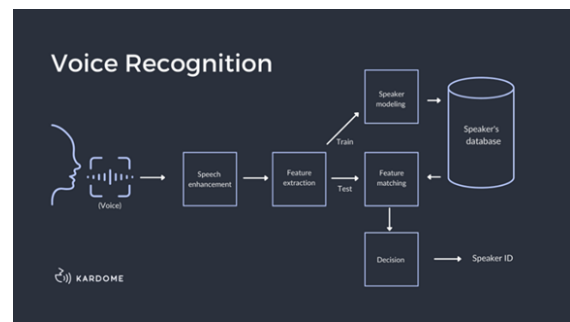
## I. INTRODUCTION

Gone are the days when people depended on other people for help or services. The digitization of the world has ensured that people no longer need to ask anyone for help, they can rely on a much more efficient and reliable device to take care of their daily needs. Computers, cell phones, laptops, etc. have become part of us and our daily lives. They can range from simple calculations to complex programs to reduce drudgery and labour waste. Virtual Personal Assistant has become almost a basic requirement in all electronic devices to perform required problems with ease.

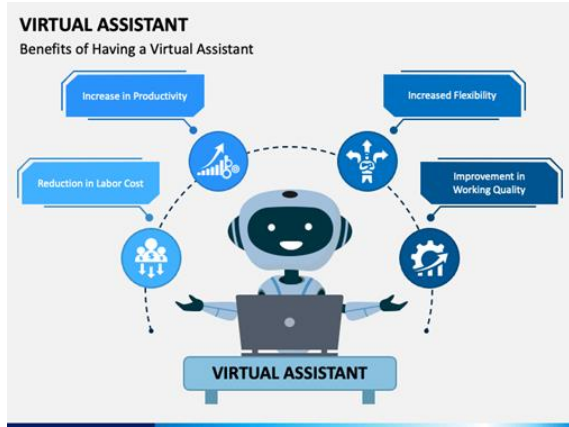
VPA is more than just a bot, VPA can make life easier for the user in several ways. Speech recognition is one of the relatively new integrations in the VPA. But although moderately efficient, it is not very useful and is not used by the user due to its large number of bugs. Although the failure rate of upcoming VPAs is around 5 percent, it is not yet up to the task of becoming a permanent part of users' lives. Therefore, the goal of the project is to build a VPA with speech recognition that has a minimal error rate.

Voice recognition is a complex process that uses advanced concepts such as neural networks and machine learning. The auditory input is processed and a neural network is created with vectors for each letter and syllable. This is called a record. When a person speaks, the device compares them to this vector and extracts the different syllables they most closely match.

Another market driver for the electronic assistant is that the car has become the mobile office, making security an increasingly important issue. Using voice commands instead of tones is not just a convenience, but is perceived by consumers as a safety necessity. The global unified messaging services market is expected to account for a significant share of telecommunications applications. It is estimated at several billion in 2007.



While indirect revenues for the carriers will be several folds. A few companies has started offering converging products in the VPA direction, e.g. Conita, WildFire, VoxSurf, VoiceGeneie, and VoiceTel and Mitel Networks, though one or two provide solutions for mobile carrier environment.



## II. FEATURES OF VIRTUAL BOT

### A. Tasks

A task is a personal or work-related task that you want to track to completion. A task can occur once or be repeated (a recurring task). A recurring task can be repeated at regular intervals or based on the date you mark the task as complete. For example, you might want to send a status report to your manager on the last Friday of every month and get a haircut if it's been a month since your last haircut.

### B. Internet Applications

The Virtual bot allows people to access , customize and use the internet to help them source information ranging from weather, schedule , to do some task which are internet specific, stock performance, competitive data and news using simple conversational voice commands.

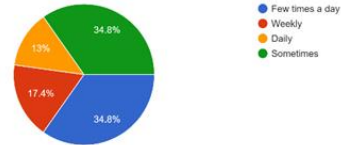
The combination of richness of internet and mobility of handheld device is forming a vast network in which it creates a functionality for user just by using human voice. A voice portal can be defined as “speech-enabled access to Web-based information.” In simple words a voice portal is a software that allows a person to access information, get information or to do some specific task by software using human voice.

The emerging of voice portal helps to make lot of functionalities for user through which some tasks which can be done with the help of by using just voice commands. This coming up technology creates a new model for business applications knows as V-Commerce.

### Research Approach:

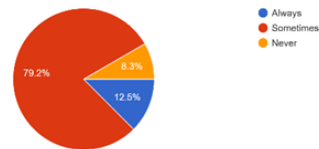
We are Trying to find out how many people use virtual agent and what problems they face and what features they expect in virtual agent.

How often do you use Voice Assistant?  
23 responses



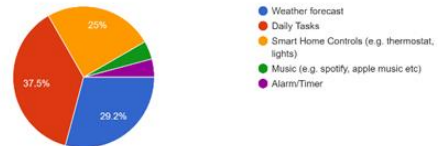
In survey we found that 34.8% people use voice assistant Few times a day and Sometimes in a day ,13% people use Daily and 17.4% people use weekly.

How often do you think that virtual agent is providing the solution as per your queries?  
24 responses



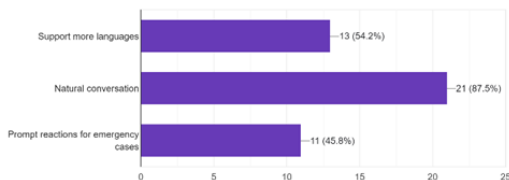
79.2% people told Sometimes virtual agent is providing solution as per queries, 12.5% people told Always and 8.3% people told never.

For what purpose do you use a Voice Assistant?  
24 responses



37.5 % people told they use virtual assistant for daily tasks, 29.2% people told they use for Weather forecast, 25% people told they use for smart home and 1% people told for Alarm timer and Music.

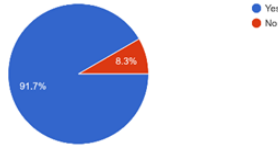
7. What do you expect from AI Voice Assistant in your near future?  
24 responses



87.5% people told they want Natural conversation, 54.2% people told they want support of more languages and 45.8% people told they want prompt reactions for emergency cases.

45.8% people rated 3 for hearing by Voice assistant is wrong, 33.3% people rated 4, 8.3% people rated 5, 2% people rated 2 and remaining rated 1.

6. Do you have any privacy concerns?  
24 responses



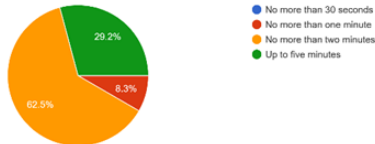
91.7% people have privacy concern issue and 8.3% people don't have issue.

Do you think that voice assistant must be modulated in such a way that it can understand different accents  
24 responses



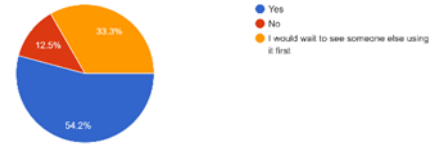
100% people told they want voice assistant to understand different accent.

How long are you willing to spend queuing for solution through Virtual agent?  
24 responses



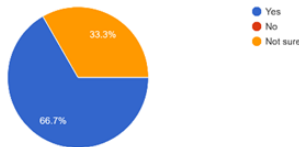
62.5% people said they will wait for no more than 2 minutes, 29.2% people told they will wait for five minutes and 8.3% people told they wait no more than one minute.

Would you speak to a retailer's artificial intelligence-based "virtual assistant" kiosk in a store?  
24 responses



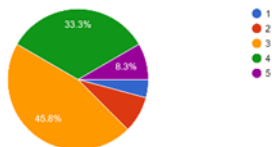
54.2% people told yes they will use ai based voice assistant kiosk in a store, 33.3% people told they will wait to someone using it first and 12.5% people told they said no.

Do you think you would buy more if an artificial intelligence-based "virtual assistant" made recommendations based on your profile, preferences and other factors?  
24 responses



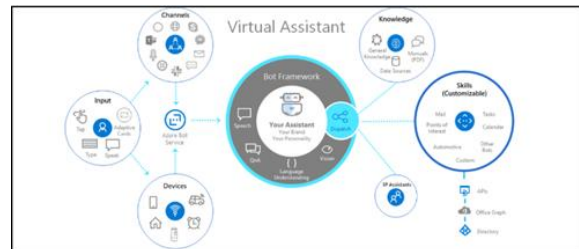
66.7% people said they will buy as per their preference and 33.3% people said not sure.

Do you think that voice assistant hears accurate when you are asking some question, or it interprets something wrong? In the scale of 1 to 5  
24 responses



### III. SYSTEM ARCHITECTURE

The system architecture of this project shows the flow of the control through the system. It also shows the hardware and the software required for the execution of the program. The architecture diagram is as follows:



### IV. HARDWARE AND SOFTWARE REQUIREMENTS

Hardware:

- A phone with a touch screen interface.

- Phone Ram should be of a minimum 512 MB.
- Internet connectivity.
- The handheld device should have USB debugging mode for development and testing purposes.

#### Software:

- Operating system should be android 4.1/win 8.1/IOS 6 or higher.
- The kernel version should be 3.0.16 or higher.
- Support of other basic applications like maps, calendar, camera, web connection etc.

### V. EXISTING AND PROPOSED MODEL

#### Existing Model

Most of the existing model use only speech recognition using neural networks. Though their systems have moderate accuracy they are not efficient worthy for practical use. There are few elementary principles used by them.

##### 1. Context Aware Computing

Context-aware Computing is a class of systems which has the ability to sense their physical environment and adapt themselves to it accordingly. It can be used for recognizing words spoken with different accents. It can also deduce words that may have been misspoken.

##### 2. MFCC:

MFCC refers to Mel-Frequency Cepstral Coefficients. MFC (Mel-Frequency Cepstrum) is a collection of these coefficients. It amounts to the short-term power spectrum of a sound. These can be used to sense variations in sound as to recognize the various variables required for voice recognition.

##### 3. Functionalities:

Some virtual assistants search for answers through a specific support database. In some situation virtual assistant doesn't understand the solution for users problem it goes in loop, misunderstanding which leads to frustration among users. As per the rise in chatbot and virtual assistance making their scope of making must be precise and accurate as per the problem. Virtual assistant doesn't understand human expressions and can't eliminate the human error completely. Virtual assistant need to be updated more

and more frequently as there is lot of problem in speech recognition so the accuracy level can never be 100%.

#### NLP:

NLP stands for Natural Language Processing is a subfield of computer science and artificial Intelligence concerned with interaction between computer and human voice in particular how to program computer to recognize the human voice and provide the result to user. The goal of Natural Language Processing is to understand the human voice and recognize what exactly the user wants and provide the result as per the human voice.

#### Proposed Model

##### Speech to Text:

It is a software which is used to recognize the human voice and converts the voice into text. It doesn't understand just anything you might say.

##### Text Analysing:

- Converted text is just letters for computer.
- A software which converts text to something that is understandable for computer.
- Computer understands the command, so Virtual Assistant like siri convert this text to computer command.
- Virtual Agent is made to understand the computer through voice command by creating set of commands and the computer will perform those functions as per the human voice. The major goal of this project is to minimize the errors caused while using speech recognition and to add more functionalities over the existing one .
- Meaning the software will convert the speech with slight modulations to increase the high level of accuracy and precision for the usability of virtual agent. The software essentially combines neural network and Machine learning to increase the clarity of human voice.
- There are lot of functionalities which virtual agents or assistant doesn't understand what exactly human queries are due to their limited functions, so

to overcome such issues the virtual agents must be made precise as per the problems.

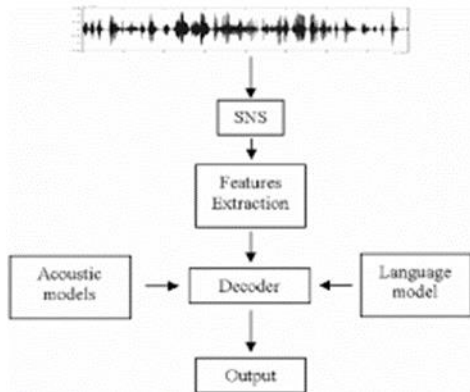
## VI. WORKING PRINCIPLES

The working of virtual Assistant using following principles:

- Natural Language Processing: Natural Language Processing (NLP) refers to AI method of communicating with intelligent devices using Natural Language using such as English.

Process of Natural Language is required when you want an intelligent device or system to perform some actions by understanding the natural language which is been given by human voice. This AI method is used to understand what exactly the human wants from the intelligent system in the absence of real time agents.

Five steps in Natural Language Processing are:



- Automatic Speech Recognition: To understand the human voice and deliver those voice to the Natural Language Processing for further functions.

- Artificial Intelligence: To learn things from humans and store information about human behaviour and relations.
- The ability of a system to understand, calculate, perceive relationships and analogies, learn from experience, store and retrieve information from memory, solve problems, use natural language fluently, and understand new situations.
- Inter Process Communication: To get important information from other software applications.

## CONCLUSION

This paper illuminates on new emerging service for computer and mobile user.

This virtual agent assistant provides an intelligent agent service to user as per the problems.

This new service is based on convergence of internet, speech recognition technology and computer/ mobile technologies.

The Virtual agent minimizes the interruption of users, improves the utilization of his time, and makes a single point of communications for all his messages, contacts, schedules and source of information.

It overcomes many of the drawbacks in the existing solutions by understanding those issues and built to make the Virtual agent more efficient and practical day by day. But the system has its own limitations. Though the efficiency may be high the time consumption can be more for each task may be higher than other virtual assistants.

## ACKNOWLEDGMENT

It is our proud privilege and duty to acknowledge the kind of help and guidance received from several people in preparation of this report. It would not have been possible to prepare this project in this form without their valuable help, cooperation and guidance. First and foremost, we thank our project coordinator Mrs.Divya Premchandran Professor Department of Computer Science and Information Technology and our Project guide Mrs. Bindy Wilson M, Asst .Professor Department of Computer Science

And Information Technology, Model College Science and Technology for their valuable guidance and all the encouragement lead towards completion of our project. We would like to thank Dr . Divya Premchandran, HOD, Department of Computer Science and Information Technology, Model College of Science and Technology for her valuable suggestions and guidance throughout the period of this project.

#### REFERENCES

- [1] Cireşan, D.; Meier, U.; Schmidhuber, J. (2012). "Multi-column deep neural networks for image classification". 2012 IEEE Conference on Computer Vision and Pattern Recognition. pp. 3642–3649. arXiv:1202.2745. doi:10.1109/cvpr.2012.6248110. ISBN 978-1-4673-1228-8. S2CID 2161592.
- [2] "Artificial intelligence: Google's AlphaGo beats Go master Lee Se-dol". BBC News. 12 March 2016. Archived from the original on 26 August 2016. Retrieved 1 October 2016.
- [3] Werbos, Paul (1982). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences" (PDF). System Modeling and Optimization. Applications of advances in nonlinear sensitivity analysis. Berlin, Heidelberg: Springer. Archived from the original (PDF) on 14 April 2016. Retrieved 16 April 2016.
- [4] Deng, L.; Yu, D. (2014). "Deep Learning: Methods and Applications" (PDF). Foundations and Trends in Signal Processing. **7** (3–4): 1–199. doi:10.1561/20000000039. Archived (PDF) from the original on 14 March 2016. Retrieved 18 October 2014.
- [5] Aletras, N.; Tsarapatsanis, D.; Preotiuc-Pietro, D.; Lamos, V. (2016). "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective". PeerJ Computer Science. **2**: e93. doi:10.7717/peerj-cs.93.
- [6] Evans, Woody (2015). "Posthuman Rights: Dimensions of Transhuman Worlds". Teknokultura. **12** (2). doi:10.5209/rev\_TK.2015.v12.n2.49072.
- [7] Wirtz, Bernd W.; Weyerer, Jan C.; Geyer, Carolin (24 July 2018). "Artificial Intelligence and the Public Sector – Applications and Challenges". International Journal of Public Administration. **42** (7): 596–615. doi:10.1080/01900692.2018.1498103. ISSN 0190-0692. S2CID 158829602. Archived from the original on 18 August 2020. Retrieved 22 August 2020.
- [8] E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," IEEE Trans. Antennas Propagat., to be published.
- [9] Ardissono, L., Boella. And Lesmo, L. (2000) "A Plan-Based AgentArchitecture for Interpreting Natural Language Dialogue", International Journal of Human-Computer Studies.