

INTELLIGENT VOICE ASSISTANT

Akshay Pramod Patil

Department of Electronics and Telecommunications Engineering

K. J. Somaiya Institute of Engineering and IT

Mumbai, Maharashtra, India.

akshay.patil2@somaiya.edu

Mohit Raju Ramtekkar

Department of Electronics and Telecommunications Engineering

K. J. Somaiya Institute of Engineering and IT

Mumbai, Maharashtra, India.

mohit.ramtekkar@somaiya.edu

Suyash Subhash Samant

Department of Electronics and Telecommunications Engineering

K. J. Somaiya Institute of Engineering and IT

Mumbai, Maharashtra, India.

suyash.samant@somaiya.edu

Shubham Gurunath Ragaji

Department of Electronics and Telecommunications Engineering

K. J. Somaiya Institute of Engineering and IT

Mumbai, Maharashtra, India.

shubham.ragaji@somaiya.edu

Dr. Jayashree Khanapuri

Department of Electronics and Telecommunications Engineering

K. J. Somaiya Institute of Engineering and IT

Mumbai, Maharashtra, India.

jayashreek@somaiya.edu

Abstract— Now-a-days voice assistants are making people's day-to-day routine simpler and it saves time as well. But, even after voice assistants are easy to operate, there are many people who cannot use voice assistant even if they wanted to. The main reason behind this issue is language barrier. People living in urban India are using voice assistants on daily basis. But, people living in rural areas are not fluent in English even to start a conversation with any voice assistant. Hence, paper proposes of the unique system based on the initiative to improve connectivity between the people living in rural India and the world. The core idea behind this initiative is to create an intelligent voice assistant system that can interact with people in their local language so that the system can help them to make their day to day life simpler. The system is designed while taking the rural life in consideration as well as their basic needs. This paper also tends to explore some new possibilities.

Keywords: wake-word, voice assistant, voice recognition, Alexa, API – Application Program Interface, localization

I. INTRODUCTION

A company in Japan has made the primary big steps toward a robot companion—one that can understand and feel emotions. Introduced in 2014, the companion robot named “Pepper” went on sale in 2015, with all 1,000 initial units selling out within a moment. The robot was programmed to read and understand human emotions as well as it can develop its own emotions and help its human friends stay happy. Pepper goes on sale in the U.S. in 2016, and more sophisticated friendly robots are bound to follow. In a new report of eMarketer, they clearly declare the use of voice assistants has reached a “critical mass”. As of 2019, there are an estimated 3.25 billion digital voice assistants getting used in devices round the world. Forecasts suggest that by 2023 the number of digital voice assistants will reach around eight billion users – a number higher than the current world's

population. But still a high number of populations cannot use the voice assistant. The most iconic application of iPhone is “SIRI”. It can communicate with user by voice commands [2]. There are still a large number of people who cannot understand English and can't speak in it fluently. This person mainly includes farmers, labors, people having small businesses, elders and many more.

Therefore, to overcome this issue we started to create a voice assistant which will communicate people in their language [8]. The proposed software will respond only after user speaks the wake-word [3].

II. LITERATURE SURVEY

- As per Google, 52% people keep their voice assistant speakers in their living room, 25% keep it in their bedroom while 22% use it in the kitchen [5].
- As per Call To Action, 1 out of 4 shoppers use voice assistant's software in their holiday shopping during the 2017 season [5].
- 25% of individual's age 16-24 use voice searches on mobile [6].

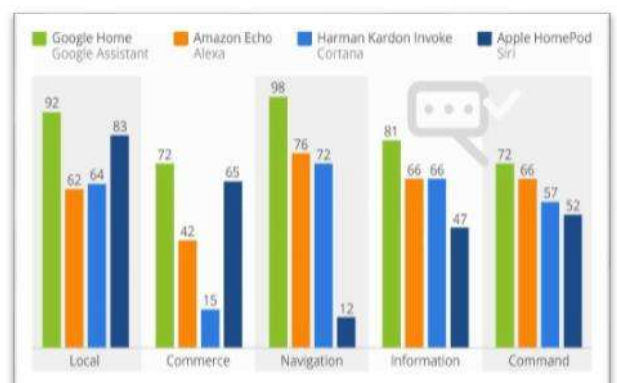


FIG. 1: PERCENTAGE OF QUERIES ANSWERED CORRECTLY BY DIFFERENT LEADING VOICE ASSISTANTS.

- From Google, 41% people who use intelligent voice assistant software feel like they are talking with a friend [6].
- As per survey done by GeoMarketing, 65 percent of people who own an Amazon Echo or Google Home or any other software can't imagine going back to the days before they had a smart speaker [6].
- By research done by Technavio, the voice assistant industry will be a \$601 million by 2019 [6].

III. PROPOSED METHODOLOGY

A. Structure

A voice assistant performs action at the mention of a keyword set by programmer. This is usually a greeting that has the assistant's name, like "Computer" or just "Jarvis". Once it gets picked up by its virtual ears, say a particular command, or ask it any question. The voice assistant uses natural language processing to decipher the command given to it, and then either answer your question or obey your command. The proposed system will mainly consist of a mic, a speaker and a raspberry pi. At the backend, a database, voice recognition software, a voice sample and the main virtual assistant software will control and run all the components related to system [7].

B. How it works

- The intelligent voice assistant responds to the commands given by the user. To accept the command, at first, the voice recognition tool of system has to be awakened to accept and execute the request [4].
- To awake the system, user has to speak a wake-word like "Computer" or "Jarvis" or by any other name that programmer associates with the software. User needs to speak its command/request/question just after he/she has said the wake-word.
- The wake-word initiates the voice recognition tool and all the words spoken after the wake-word gets accepted by the voice recognition tool. This tool then analyzes the input and bifurcate it as a request or question or command.
- After analyzing and understanding of the input, the software gathers the related information about the query from the internet. As soon as the tool gathers the necessary and relevant information, system triggers the language convertor tool to translate that information from English to the desired language.
- The language convertor tool automatically triggers the voice system. The voice tool speaks the translated

information by use of speaker.

IV. WORKING PRINCIPLE

The proposed system connects to the Internet by using any Wi-Fi network. It's always on and listening for the wake word to wake the device up. Once it hears that, the device gathers the voice commands that follow after the wake-word and sends a signal to a natural voice recognition service in the cloud called The Voice Service, which interprets them and sends back the appropriate response. The system may consist of an array of microphones that can pick up your voice from across the room, even over music and other environmental noise.

V. SKILLS

In this system, various skills have been created in Hindi and Marathi languages such as **facts**, **good thoughts**, **weather and time** and **nearby hospitals** and **a city guide**. These skills are created by using an Amazon Developer account.

Amazon has released all its software components for developers to create their voice assistant as well as it will help to develop Alexa so that it can speak multiple languages. The information regarding these skills is as follows:

- **Facts:** Here, a skill using Alexa API's and its voice model which will inform us about historical as well as generalized facts. This skill can be used for educational purpose. In rural areas, where the frequency of schools is less, this skill can be used as a knowledge provider.
- **Good Thoughts:** This skill in Hindi and Marathi language can be pronounced as सुविचार. This skill will provide positive thoughts from the list of thoughts entered into the skill. This skill will narrate good thoughts randomly at the specified time which the user decides.
- **Weather and Time:** Amazon has created intents which gathers the required information from the internet and converts it in the required language using the Language translator tool. In this skill, among many languages, Hindi language is selected. The weather and time skill can inform about climate conditions and the time of almost every major city around the globe. This skill can be used for touring, the traveler can get the information about his/her destination's live weather conditions and the skill can inform about the maximum and minimum temperature as well.
- **Nearby Hospitals:** As everyone knows, the number of hospitals and dispensaries are very less in rural areas as compared to urban areas. If any medical emergency arises, the patient needs to be moved to the nearest hospital as soon as possible.

The Nearby Hospital skill can inform the nearest hospital in the local area and the skill provides its contact information, ambulance number as well as the toll free number for emergency services.

- City Guide:** If any person has moved to the new location which he/she isn't familiar with, then this skill can be used to get familiar with the surroundings. The City guide skill can provide information about schools, colleges, restaurants, bar, hospitals, hotels, etc.

Some of these skills are already created by Amazon software developers and some of them are not. The skills created by amazon developers and only communicate to the user in major global languages like English, Spanish, Italian, etc. But most of the times, people living in rural parts of India cannot even understand English. Hence, all the skills proposed in the Intelligent Voice Assistant Software system supports Marathi and Hindi language. So that people who are familiar with these two languages can interact with the system more easily.

VI. RESULTS

- Facts skill:** This skill narrates a random fact whenever it is invoked. If user says Computer...अंतरिक्ष facts खोलें. और मझु अंतरिक्ष के बारेमें एक fact बताओ, then the software randomly selects any fact and shows it as output. It is as shown below :



FIG. 2: SPACE FACTS

- Good thoughts:** Unlike facts skill, the good thought i.e. सुविचार skill narrates one good thought randomly every time whenever this skill is invoked. Users just have to say आजका सुविचार. The result is as shown below:

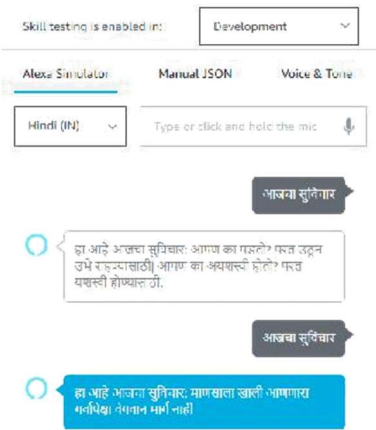


FIG. 3: GOOD THOUGHTS

- Weather and Time:** The Weather and time skill not only informs about the home location's time and weather conditions but also it can inform about the weather conditions and that location's actual time as well. Users just have to say समय और तापमान to invoke the skill. After invocation, the skill responses as follows :



FIG.3: WEATHER AND TIME

- Nearby Hospitals:** During the emergency medical situations, the nearby hospital skill gives details about the nearest hospital from its location.

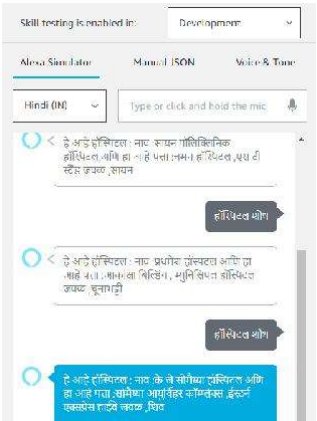


FIG.5: NEARBY HOSPITALS

- City Guide:** If any person is new to a particular location, then by using this skill, user can gather information about the stores, schools, etc.

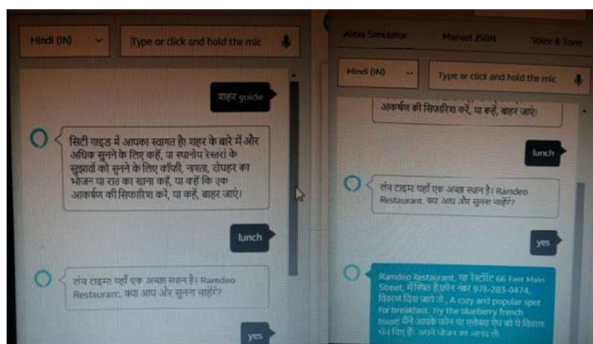


FIG. 6: CITY GUIDE

VII. APPLICATIONS

Voice assistants can provide a wide variety of services. These services can be used to save time as much as possible

- Provide information like weather, facts from e.g. Wikipedia, set an alarm, and make to-do lists and shopping lists.
- Play music from streaming services like Spotify and Pandora; play radio stations; read audio books.
- It can acquire information about holidays and birthdays from your calendars and other personal information as well.
- It can search for any information online, from restaurant bookings to directions to any place, weather updates and news. During any emergency condition, the proposed system can brief about the nearest hospital location as well.
- It can set or run timers and reminders [3].
- It can read your notifications to you.
- The proposed system can do real-time language translation.
- Automated Transportation.

Nowadays, in many auto shows, we have seen the cars which can drive automatically without any driver. But, for safety, driver has to be at the driver's seat behind the wheel. Despite all the technological developments, today, roads and people are not fully ready to accept the unmanned vehicle. Nowadays, big automotive brands like BMW, MG and Mercedes-Benz are now creating a voice assistant for their cars. This step is taken to provide more comfort to the customer.

- **Military:**
Intelligent voice assistants can be a step forward to create the Artificial Intelligence [4]. Military of Superpower Countries are now developing voice assistants that can update military about the targeted area such as its geographical conditions, climate, population living nearby and many more things. Soldiers can update about any casualty happened during war. Voice assistants can act as a hands-free translator. Hence, it is used to decrease the communication gap as well. Navy and air force are now making their artilleries voice activated.
- **Improving Private Sector:**

Many Private industries are now creating their own voice assistant to enhance the communication between boss and their employees. By using a voice assistant, people working in corporate sector can have their day scheduled and the system can inform user about his/her appointments and it can update user by reading its mails, etc.

- **Robot as friends:**

A company in Japan has made the primary big steps toward a robot companion—one that can understand and feel emotions. Introduced in 2014, “Pepper” the companion robot went on sale in 2015, with all 1,000 initial units selling out within a moment. The robot was programmed to read and understand human emotions as well as it can develop its own emotions, and help its human friends stay happy. Pepper goes on sale in the U.S. in 2016, and more sophisticated friendly robots are bound to follow.

- **Improve Connectivity:**

The given system will interact with people living in rural areas in their local language. So that they can blend in their surrounding more easily and they can have all the necessary information like schools, hospitals, etc. more easily by use of voice assistants. The voice assistant can update about the weather on any location and it can acknowledge user with the current news as well.

Thereby, it can be concluded that if this method is implemented at a full scale, better connectivity as well as ease in day to day life can be achieved by the intelligent voice system.

VIII. CONCLUSION

The proposed system is designed solely by taking rural life into consideration and it will surely help to make rural lifestyle more comfortable and less time consuming. The implemented skills will surely be very much helpful to the user as well as others. As the virtual assistant industry is rapidly progressing day by day, there will be fewer hurdles to the application of the system in upcoming days and it will surely help creating rural lifestyle more and easier and rural people will surely have more connectivity and interaction with the outside world. Hence, urbanization of rural region can take place at the faster rate.

IX. FUTURE SCOPE

By considering rural lifestyle, there are many more skills that can be made to make their lifestyle simpler. Some important skills that can be implemented by using the given software are as follows:

- **Traffic Updates:** The traffic updates skill can be used to check traffic conditions nearby. The given skill may be able to inform the distance as well as time required to reach from point A to point B. Also, it may provide information about the shortest route between source and destination. If ambulance services are unavailable in certain areas then,

by using the Traffic Updates skill as well as Find Hospital skill, the patient can reach nearest hospital in minimum time.

[8] https://en.wikipedia.org/wiki/Virtual_assistant

- **Real-Time Voice Translator:** The proposed software can be used as a real-time language translator. There are several language translators by google and other companies, but unlike google this system can be used on an instant basis[3].
- **Hindi Gaane:** In rural areas, the tools of entertainment are very less as compared to the urban locations. This skill can be used as an entertainment purpose. The Hindi Gaane skill will mostly contain old as well as new songs. But, mostly it will contain popular songs. This skill will gather the required data from internet and plays it on speaker.
- **Reminder:** The reminder skill can be the most used skill in the villages. It can be used to set medicinal reminder, shopping reminder, etc. The Reminder skill can be used to keep a weekly schedule containing all the tasks which are important for the day to day life as well as it can be used for occupational purpose.
- **Current Affairs:** The Current Affairs skill can be mostly used in the workplace like in shops, farming field, etc. The Current Affairs skill can inform user about the current news broadcasting in its location. Also, it can inform about the rates of basic necessary items such as vegetables, fruits, petrol, etc.

X. REFERENCES

- [1] Abhay D, Chaitanya K and Rohan K, "Study of Voice Controlled Personal Assistant Device", International Journal of Computer Trends and Technology (IJCTT), December, 2016, India
- [2] Dr. Kshama Kulhalli, Dr. Kotrappa Sirbi and Mr. Abhijit Patankar, "Personal assistant with Voice Recognition Intelligence", International Journal of Engineering Research and Technology (IJERT), 2017, India
- [3] Matthew B. Hoy, "Alexa, SIRI, Cortana, and more: An Introduction to Voice Assistants", Medical Reference Services Quarterly Journal, Jan 12, 2018
- [4] <https://www.voicesummit.ai/blog/how-the-military-is-using-voice-technology>
- [5] <https://www.wordstream.com/blog/ws/2018/04/10/voice-search-statistics-2018>
- [6] <https://www.searchenginejournal.com/33-of-people-are-now-using-voice-assistants-regularly/321413/#close>
- [7] <https://gearpatrol.com/2017/07/19/how-does-a-virtual-assistant-work/>