CHARUSAT

CSPIT

K.D.PATEL DepARTment of

InformATIon Technology

VOICE ASSISTANT

Chandubhai S Patel institute of Technology,

CHARUSAT, Changa, Anand.

Anujkumar Mistry

18it064@charusat.edu.in

Prof. Amit Parmar

amitparmar.it@charusat.ac.in

ABSTRACT:

In today's era almost all tasks are digitalized. We have Smartphone in hands and it is nothing less than having world at your finger tips. These days we aren't even using fingers. We just speak of the task and it is done. There exist systems where we can say Text Dad, "I'll be late today." And the text is sent. That is the task of a Virtual Assistant. It also supports specialized task such as booking a flight, or finding cheapest book online from various ecommerce sites and then providing an interface to book an order are helping automate search, discovery and online order operations.

Virtual Assistants are software programs that help you ease your day to day tasks, such as showing weather report, creating reminders, making shopping lists etc. They can take commands via text (online chat bots) or by voice. Voice based intelligent assistants need an invoking word or wake word to activate the listener, followed by the command. We have so many virtual assistants, such as Apple's Siri, Amazon's Alexa and Microsoft's Cortana.

Voice searches have dominated over text search. Web searches conducted via mobile devices have only just overtaken those carried out using a computer and the analysts are already predicting that 50% of searches will be via voice by 2020.Virtual assistants are turning out to be smarter than ever. Allow your intelligent assistant to make email work for you. Detect intent, pick out important information, automate processes, and deliver personalized responses.

BACKGROUND:

There already exist a number of desktop virtual assistants. A few examples of current virtual assistants available in market are discussed in this section along with the tasks they can provide and their drawbacks.

SIRI from Apple

SIRI is personal assistant software that interfaces with the user thru voice interface, recognizes commands and acts on them. It learns to adapt to user's speech and thus improves voice recognition over time. It also tries to converse with the user when it does not identify the user request.

It integrates with calendar, contacts and music library applications on the device and also integrates with GPS and camera on the device. It uses location, temporal, social and task based contexts, to personalize the agent behavior specifically to the user at a given point of time

Supported Tasks

- Call someone from my contacts list
- Launch an application on my iPhone
- Send a text message to someone
- Set up a meeting on my calendar for 9am tomorrow
- Set an alarm for 5am tomorrow morning
- Play a specific song in my iTunes library
- Enter a new note

Drawback

SIRI does not maintain a knowledge database of its own and its understanding comes from the information captured in domain models and data models.

ReQall

ReQall is personal assistant software that runs on smartphones running Apple iOS or Google Android operating system. It helps user to recall notes as well as tasks within a location and time context. It records user inputs and converts them into commands, and monitors current stack of user tasks to proactively suggest actions while considering any changes in the environment. It also presents information based on the context of the user, as well as filter information to the user based on its learned understanding of the priority of that information.

Supported Tasks

- Reminders
- Email
- Calendar, Google Calendar
- Outlook
- Evernote
- Facebook, LinkedIn
- News Feeds

Drawback

Will take some time to put all of the to-do items in – you could spend more time putting the entries in than actually doing the revision.

OBJECTIVES

Main objective of building personal assistant software (a virtual assistant) is using semantic data sources available on the web, user generated content and providing knowledge from knowledge databases. The main purpose of an intelligent virtual assistant is to answer questions that users may have. This may be done in a business environment, for example, on the business website, with a chat interface. On the mobile platform, the intelligent virtual assistant is available as a call-button operated service where a voice asks the user "What can I do for you?" and then responds to verbal input.

One of the main advantages of voice searches is their rapidity. In fact, voice is reputed to be four times faster than a written search: whereas we can write about 40 words per minute, we are capable of speaking around 150 during the same period of time15. In this respect, the ability of personal assistants to accurately recognize spoken words is a prerequisite for them to be adopted by consumers.

PRODUCT DESCRIPTION

As a personal assistant, V.A. assists the end-user with day-to-day activities like general human conversation, searching queries in various search engines like Google, Bing or Yahoo, searching for videos, retrieving images, live weather conditions, word meanings, searching for medicine details, health recommendations based on symptoms and reminding the user about the scheduled events and tasks. The user statements/commands are analysed with the help of machine learning to give an optimal solution.

SCOPE:

Presently, V.A. is being developed as an automation tool and virtual assistant. Among the Various roles played by V.A. are:

1. Search Engine with voice interactions
2. Medical diagnosis with Medicine aid.
3. Reminder and To-Do application.
4. Vocabulary App to show meanings and correct spelling errors.
5. Weather Forecasting Application.

RELATED WORK [HISTORY]

The development of voice assistants was started in 1962 at the Seattle world's fair where IBM presented a device called shoebox IBM that could recognize spoken digits and then return them back through igniting lamps labelled next to the digits 0-9. It had the ability to perceive a total of 16 words. Currently most of the voice assistances are developed for the mobile phones like google make voice assistance support for android mobiles, apple use Siri and amazon have Alexa these assistants used language processing to perform its task. another voice assistant is Cortana which is been developed by Microsoft and used on desktop. All of these voice assistants perform the same intended function – that is, voice initialized processing, and all of these developments have been a result of the same new age technology- Artificial intelligence.

At the core of all these assistants is a simple synchronous cycle – Voice commands and hear responses. Sutar Shekhar, and various researchers have jointly come up with an application which implements most of the system functionalities through voice and they also included the feature of sending a message with their voice command to help those people who are visually impaired.

They aim to continuously develop their application so as to a finally have an engine which can also recognise different local languages like Bengali, and a number of dialects of prevalent Hindi. Miss. Priyanka V. Mhamunkar and others proposed a system which will let the individual fetch meanings of words through synthesised sounds. Omyonga Kevin and Kasamani Bernard Shibwabo claim to have come up with an application that is able to implement spoken commands even without an internet connection thus, giving us flexibility over data costs. As there is no need of internet connection this feature makes their solution faster than many of the high-profile engines like Alexa and Cortana.

PROPOSED WORK

Methodology

1. Speech recognition

The proposed system used the google API to convert input speech into text. The speech is given as an input to google cloud for processing, As an output, the system then receives the resulting text.

2. Backend work

At backend the python gets the output from speech recognition and after that it identifies whether the command is a system command or a browser command. The output is send back to python backend to give desired output to user.

3. Text to speech

Text to speech, or TTS, is a new wave technique of for transforming voice commands into readable text. Not to mix it up with VR Systems that instead, generate speech by joining strings gathered in an exhaustive DB of preinstalled text and have been developed for different goals which form full-fledged sentences, clauses or meaningful phrases through a dialect's graphemes and phonemes. Such systems have their limits as they can only determine text on the basis of pre-determined text in their databases TTS systems, on the other hand, are practically to "read" strings of characters and dole out resulting sentences, clauses and phrases.

Proposed Architecture:

The system design consists of

1. Taking the input as speech patterns through microphone.

2. Audio data recognition and conversion into text.

3. Comparing the input with predefined commands

4. Giving the desired output

The initial phase includes the data being taken in as speech patterns from the microphone.in the second phase the collected data is worked over and transformed into textual data using NLP. Inthe next step, this resulting stringified data is manipulated through Python Script to finalise the required output process. In the last phase, the produced output is presented either in the form of text or converted from text to speech using TTS

PURPOSE, SCOPE AND APPILCABILITY

Purpose:

Purpose of virtual assistant is to being capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audiobooks, and providing weather, traffic, sports, and other real-time information, such as news. Virtual assistants enable users to speak natural language voice commands in order to operate the device and its apps.

There is an increased overall awareness and a higher level of comfort demonstrated specifically by millennial consumers. In this ever-evolving digital world where speed, efficiency, and convenience are constantly being optimized, it's clear that we are moving towards less screen interaction.

Scope:

Voice assistants will continue to offer more individualized experiences as they get better at differentiating between voices. However, it's not just developers that need to address the complexity of developing for voice as brands also need to understand the capabilities of each device and integration and if it makes sense for their specific brand. They will also need to focus on maintaining a user experience that is consistent within the coming years as complexity becomes more of a concern. This is because the visual interface with voice assistants is missing. Users simply cannot see or touch a voice interface.

Applicability

The mass adoption of artificial intelligence in users' everyday lives is also fueling the shift towards voice. The number of IoT devices such as smart thermostats and speakers are giving voice assistants more utility in a connected user's life. Smart speakers are the number one way we are seeing voice being used. Many industry experts even predict that nearly every application will integrate voice technology in some way in the next 5 years.

The use of virtual assistants can also enhance the system of IoT (Internet of Things). Twenty years from now, Microsoft and its competitors will be offering personal digital assistants that will offer the services of a full-time employee usually reserved for the rich and famous.

SURVEY OF TECHNOLOGY

Python:

Python is an OOPs (Object Oriented Programming) based, high level, interpreted programming language. It is a robust, highly useful language focused on rapid application development (RAD). Python helps in easy writing and execution of codes. Python can implement the same logic with as much as 1/5th code as compared to other OOPs languages.

Python provides a huge list of benefits to all. The usage of Python is such that it cannot be limited to only one activity. Its growing popularity has allowed it to enter into some of the most popular and complex processes like Artificial Intelligence (AI), Machine Learning (ML), natural language processing, data science etc. Python has a lot of libraries for every need of this project. For V.A., libraries used are speechrecognition to recognize voice, Pyttsx for text to speech, selenium for web automation etc.

Python is reasonably efficient. Efficiency is usually not a problem for small examples. If your Python code is not efficient enough, a general procedure to improve it is to find out what is taking most the time, and implement just that part more efficiently in some lower-level language. This will result in much less programming and more efficient code (because you will have more time to optimize) than writing everything in a low-level language.

DBpedia:

Knowledge bases are playing an increasingly important role in enhancing the intelligence of Web and enterprise search and in supporting information integration. The DBpedia leverages this gigantic source of knowledge by extracting structured information from Wikipedia and by making this information accessible on the Web. The DBpedia knowledge base has several advantages over existing knowledge bases: it covers many domains; it represents real community agreement; it automatically evolves as Wikipedia changes, and it is truly multilingual.

The DBpedia knowledge base allows you to ask quite surprising queries against Wikipedia for instance "Give me all cities in New Jersey with more than 10,000 inhabitants" or "Give me all Italian musicians from the 18th century".

Quepy:

Quepy is a python framework to transform natural language questions to queries in a database query language. It can be easily customized to different kinds of questions in natural language and database queries. So, with little coding you can build your own system for natural language access to your database.

Pyttsx:

Pyttsx stands for Python Text to Speech. It is a cross-platform Python wrapper for textto-speech synthesis. It is a Python package supporting common text-to-speech engines on Mac OS X, Windows, and Linux. It works for both Python2.x and 3.x versions. Its main advantage is that it works offline.

Speech Recognition:

This is a library for performing speech recognition, with support for several engines and APIs, online and offline. It supports APIs like Google Cloud Speech API, IBM Speech to Text, Microsoft Bing Voice Recognition etc.

SQLite:

SQLite is a capable library, providing an in-process relational database for efficient storage of small-to-medium-sized data sets. It supports most of the common features of SQL with few exceptions. Best of all, most Python users do not need to install anything to get started working with SQLite, as the standard library in most distributions ships with the sqlite3 module.

SQLite runs embedded in memory alongside your application, allowing you to easily extend SQLite with your own Python code. SQLite provides quite a few hooks, a reasonable subset of which are implemented by the standard library database driver.

HARDWARE AND SOFTWARE REQUIREMENTS:

The software is designed to be light-weighted so that it doesn't be a burden on the machine running it. This system is being build keeping in mind the generally available hardware and software compatibility. Here are the minimum hardware and software requirement for virtual assistant.

Hardware:

- Pentium-pro processor or later.
- RAM 512MB or more.

Software:

- Windows 7(32-bit) or above.
- Python 2.7 or later
- Chrome Driver
- Selenium Web Automation
- SQLite

Features:

The System shall be developed to offer the following features:

1) It keeps listening continuously in inaction and wakes up into action when called with a particular predetermined functionality.

2) Browsing through the web based on the individual's spoken parameters and then issuing a desired output through audio and at the same time it will print the output on the screen.

Other useful services such as playing any kind of media, browsing weather forecasts, setting, reminders, shut down computer, sending an Email etc. Are provided as a result of spoken commands.

VOICE ASSISTANT PRIVACY CONCERNS:

A user may have a privacy concern as personal assistant require a huge amount of data and are always listening to take the command This passive data is then, retained and sifted through humans that are employed by almost all of the major companies- Amazon, Apple etc. In addition to this discovery over the Ai being able to record our audio interactions, there have been concerns over the type of data such employees and contractors were hearing. So, in the cloud-based voice assistant, privacy policy must be there which will protect the user information

Features in Voice Assistant:

1. **Accessing youtube videos**

   Videos have remained as a main source of entertainment, one of the most prioritized tasks of virtual assistants. They are equally important for entertainment as well as educational purposes as most teaching and research activities in present times are done through Youtube. This helps in making the learning process more practical and out of the four walls of the classroom.

2. **Get weather for a location**

   Getting live weather conditions about a place remains an important task of virtual assistants. It helps the user charter the course of their action. V.A. addresses this issue with the help of Python.

3. **Dictionary meaning**

   One of the usages of the web is to find word meaning and its usage in our day to day life. Instead of going through the bulky books, our users can simply search for it using the voice assistant and get the meaning within a fraction of seconds.

4. **Set Reminders**

   One of the main features of a voice assistant is to set a reminder for the user accordingly. V.A. is no different when it comes to this. The user can set reminders to be notified about a task at a particular time. This will help users, especially developers to schedule their time and resources easily. All the user have to do is to input Set reminder to the assistant. A form will be displayed. Fill the form with the required details and click on set reminder button.

5. **Sending Emails**

   Integrating mailing features to V.A. eases the job of mailing, which otherwise would have to be done by opening the concerned email address. With V.A., you do not need to go for another tab to do one of the major task of your day to day affairs. The user can send emails to the desired receiver. He should input Send mail, after which a form will be displayed. Fill the form with the require details and click on the send mail button.

CONCLUSION:

Voice assistants have had a huge change in user' s interaction with technologies embedded in their devices. Like any other technology of such magnitude, they have altered the basic genome of the sphere in which they operate. While this has largely created a better world with drastic benefits for communities, which were before kept in dark with reference to technological innovations, they have posed new kind of threats with respect to user's privacy and security.

FUTURE SCOPE:

The future of voice assistants can be parameterised on an array of dimensions. With respect to compatibility and integration with other devices and systems, there i sill a lot to be achieved, Another dimension would be with respect to the redundant use of wake words at the beginning of each command. The individuality of results also poses a big problems.

But for all intents and purposes, the future of these technology is a bright one. With advances in it and in technologies related to it (search processes, for example) Voice assistants can carry out even more complex tasks like booking tickets, etc. At its core, this technology might have its own trials and tribulations, but it is still a boon for many who might have been kept in the dark with all spheres of technological developments. Apart from this, it is just too beneficial a technology to not go through continuous research and development.