AI DESKTOP PERSONAL ASSISTANT

G. Premraj, V. Chandu Sagar, A. Mohan Krishna Murthy, A. Rohan, Dr. L. Kiran Kumar Reddy
[1,2,3,4]UG Student, [4]Professor, [1,2,3,4]Dept. of Computer science Engineering, Visvesvaraya College of Engineering and Technology, Mangalpalle, Telangana, India

ABSTRACT

The advancement in technology over time has been unmeasurable. From the first digital computer built by Eniac having a clock speed of 100KHz to Summit developed by the US Department of Energy has a performance of 148.6 peta Flops, we have come a long way in technological advancement.

In such an era of advancement if people are still struggling to interact with their machine using various input devices then it's not worth it. For this reason, many voice assistants were developed and are still being improved for better performance and efficiency.

The main task of a voice assistant is to minimize the use of input devices like keyboard, mouse, touch pens, etc. This will reduce both the hardware cost and space taken by it.

The Most famous application is of iPhone "SIRI" which helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user. Same kind of application is also developed by the Google that is "Google Voice Search" which is used for in Android Phones.

It is named as Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user. In addition, this proposed system can change the way of interactions between end user and the mobile devices.

INTRODUCTION

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 e-commerce 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. For my project the wake word is G-one. We have so many virtual assistants, such as Apple's Siri, Amazon's Alexa and Microsoft's Cortana. For this project, wake word was chosen G-one.

This system is designed to be used efficiently on desktops. Personal assistant software improves user productivity by managing routine tasks of the user and by providing information from online sources to the user. G-one is effortless to use. Call the wake word 'G-one' followed by the command. And within seconds, it gets executed.

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.

This project was started on the premise that there is sufficient amount of openly available data and information on the web that can be utilized to build a virtual assistant that has access to making intelligent decisions for routine user activities.

Existing System and Drawbacks:

There already exist a number of desktop virtual assistants. One of the examples of current virtual assistants available in market is 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.

Main Objective

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.

Virtual assistants can tremendously save you time. We spend hours in online research and then making the report in our terms of understanding. G-one can do that for you. Provide a topic for research and continue with your tasks while G-one does the research. Another difficult task is to remember test dates, birthdates or anniversaries. It comes with a surprise when you enter the class and realize it is class test today. Just tell G-one in advance about your tests and she reminds you well in advance so you can prepare for the test.

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.

LITERATURE SURVEY

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 JARVIS, libraries used are speech recognition to recognize voice, Pyttsx3 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.

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.

Pyttsx3:

Pyttsx3 stands for Python Text to Speech. It is a cross-platform Python wrapper for text-to-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 (Structured Query Language) 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 distribution 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.

Voice Assistant The Future:

Voice search is implemented as a two-stage search procedure where string candidates generated by an automatic speech recognition (ASR) system are re-scored in order to identify the best matching entry from a potentially very large application specific database. Study provides a good example of how additional domain specific knowledge sources can be used with a domain independent ASR system to facilitate voice access to online search indices. As more data becomes available for a given speech recognition task, the natural way to improve recognition International Journal of Engineering Research and Technology. Accuracy is to train larger acoustic models.

There is a nonparametric empirical model that exploits abundant training data to directly learn pronunciation variation. Inter with a parametric model yields the best performance, with a relative improvement of 5.2% in WER over the baseline. There are a number of ways in which this work could be extended. First, closer integration with acoustic model t is likely to yield sharper distributions and a tighter fit to the data. Second, estimating word pronunciation co counts in semi-supervised fashion (e.g. through word recognition instead of forced alignment) would broaden its applicability to a wide range of speech genres and tasks.

Problem Definition

Usually, user needs to manually manage multiple sets of applications to complete one task. For example, a user trying to make a travel plan needs to check for airport codes for nearby airports and then check travel sites for tickets between combinations of airports to reach the destination. There is need of a system that can manage tasks effortlessly. We already have multiple virtual assistants. But we hardly use it. There are number of people who have issues in voice recognition. These systems can understand English phrases but they fail to recognize in our accent. Our way of pronunciation is way distinct from theirs. Also, they are easy to use on mobile devices than desktop systems. There is need of a virtual assistant that can understand English in Indian accent and work on desktop system.

When a virtual assistant is not able to answer questions accurately, it’s because it lacks the proper context or doesn’t understand the intent of the question. Its ability to answer questions relevantly only happens with rigorous optimization, involving both humans and machine learning. Continuously ensuring solid quality control strategies will also help manage the risk of the virtual assistant learning undesired bad behaviours. They require large amount of information to be fed in order for it to work efficiently.

Virtual assistant should be able to model complex task dependencies and use these models to recommend optimized plans for the user. It needs to be tested for finding optimum paths when a task has multiple sub-tasks and each sub-task can have its own sub-tasks. In such a case there can be multiple solutions to paths, and the it should be able to consider user preferences, other active tasks, priorities in order to recommend a particular plan.

Requirement Specifications

Personal assistant software is required to act as an interface into the digital world by understanding user requests or commands and then translating into actions or recommendations based on agent’s understanding of the world. JIA focuses on relieving the user of entering text input and using voice as primary means of user input. Agent then applies voice recognition algorithms to this input and records the input. It then use this input to call one of the personal information management applications such as task list or calendar to record a new entry or to search about it on search engines like Google, Bing or Yahoo etc. Focus is on capturing the user input through voice, recognizing the input and then executing the tasks if the agent understands the task. Software takes this input in natural language, and so makes it easier for the user to input what he or she desires to be done.

Voice recognition software enables hands free use of the applications, lets users to query or command the agent through voice interface. This helps users to have access to the agent while performing other tasks and thus enhances value of the system itself. JIA also have ubiquitous connectivity through Wi-Fi or LAN connection, enabling distributed applications that can leverage other APIs exposed on the web without a need to store them locally.

Virtual assistants must provide a wide variety of services. These include:

- Providing information such as weather, facts from e.g. Wikipedia etc.
- Set an alarm or make to-do lists and shopping lists.
- Remind you of birthdays and meetings.
- Play music from streaming services such as Saavan and Gaana.
- Play videos, TV shows or movies on televisions, streaming from e.g. Netflix or Hotstar.
- Book tickets for shows, travel and movies.

Feasibility Study

Feasibility study can help you determine whether or not you should proceed with your project. It is essential to evaluate cost and benefit. It is essential to evaluate cost and benefit of the proposed system. Five types of feasibility study are taken into consideration.

1. Technical feasibility: It includes finding out technologies for the project, both hardware and software. For virtual assistant, user must have microphone to convey their message and a speaker to listen when system speaks. These are very cheap now a days and everyone generally possess them. Besides, system needs internet connection. While using JIA, make sure you have a steady internet connection. It is also not an issue in this era where almost every home or office has Wi-Fi.

2. Operational feasibility: It is the ease and simplicity of operation of proposed system. System does not require any special skill set for users to operate it. In fact, it is designed to be used by almost everyone. Kids who still don't know to write can read out problems for system and get answers.

3. Economical feasibility: Here, we find the total cost and benefit of the proposed system over current system. For this project, the main cost is documentation cost. User also would have to pay for microphone and speakers. Again, they are cheap and available. As far as maintenance is concerned, JIA won't cost too much.

4. Organizational feasibility: This shows the management and organizational structure of the project. This project is not built by a team. The management tasks are all to be carried out by a single person. That won't create any management issues and will increase the feasibility of the project.

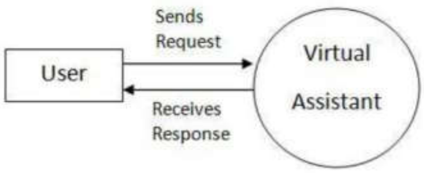
5. Cultural feasibility: It deals with compatibility of the project with cultural environment. Virtual assistant is built in accordance with the general culture. The project is named JIA so as to represent Indian culture without undermining local beliefs.

This project is technically feasible with no external hardware requirements. Also it is simple in operation and does not cost training or repairs. Overall feasibility study of the project reveals that the goals of the proposed system are achievable. Decision is taken to proceed with the project.

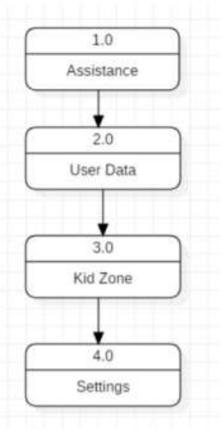
DESIGN

Data Flow Diagram:

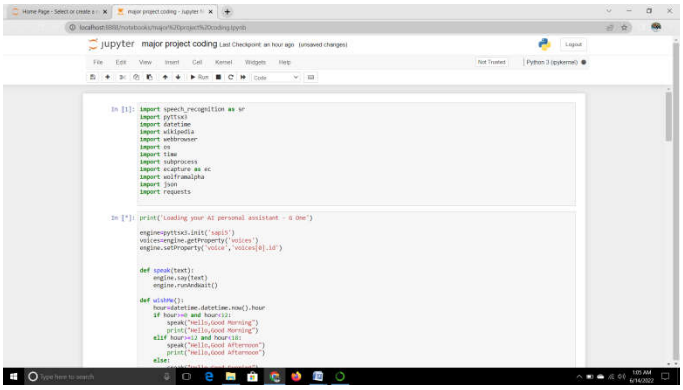
DFD Level 0 (Context Level Diagram)



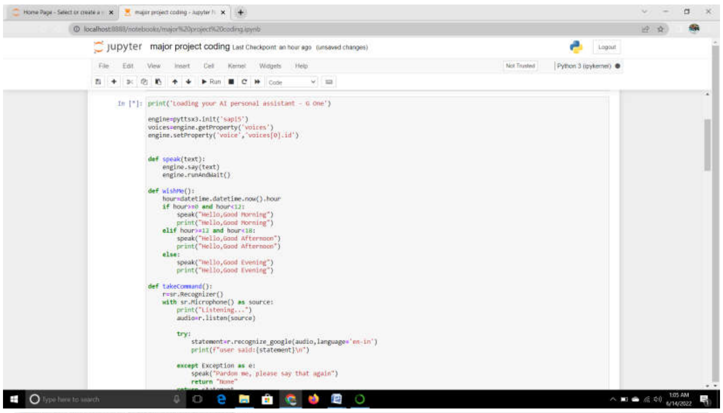
DFD Level 1



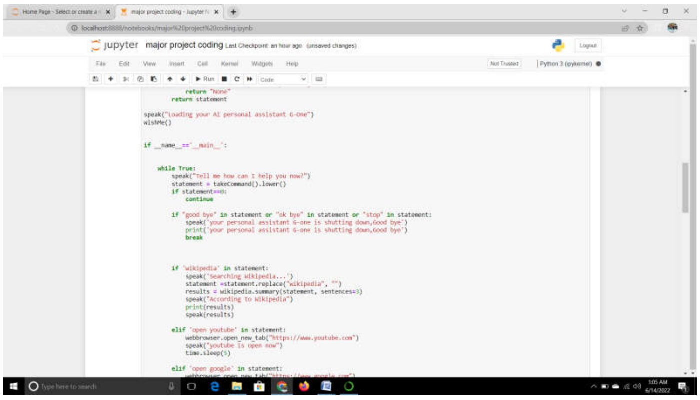
Results


jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Logout

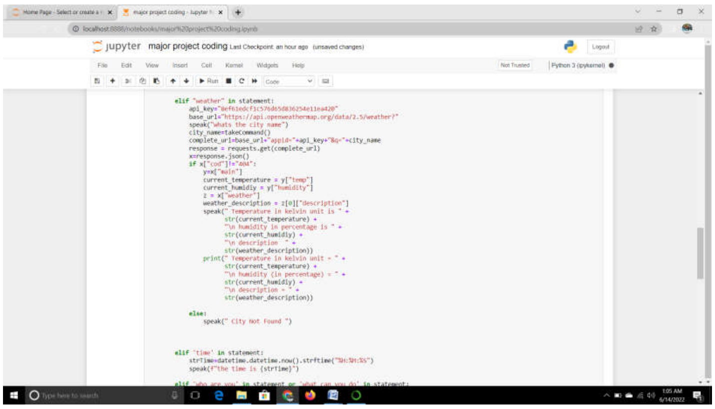
```
In [1]: import speech_recognition as sr
        import pyttsx3
        import datetime
        import wikipedia
        import webbrowser
        import os
        import time
        import subprocess
        import ecapture as ec
        import wolframalpha
        import json
        import requests

In [*]: print('Loading your AI personal assistant - G One')

        engine=pyttsx3.init('sapi5')
        voices=engine.getProperty('voices')
        engine.setProperty('voice','voices[0].id')


        def speak(text):
            engine.say(text)
            engine.runAndWait()

        def wishMe():
            hour=datetime.datetime.now().hour
            if hour>=0 and hour<12:
                speak("Hello,Good Morning")
                print("Hello,Good Morning")
            elif hour>=12 and hour<18:
                speak("Hello,Good Afternoon")
                print("Hello,Good Afternoon")
            else:
```


jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Logout

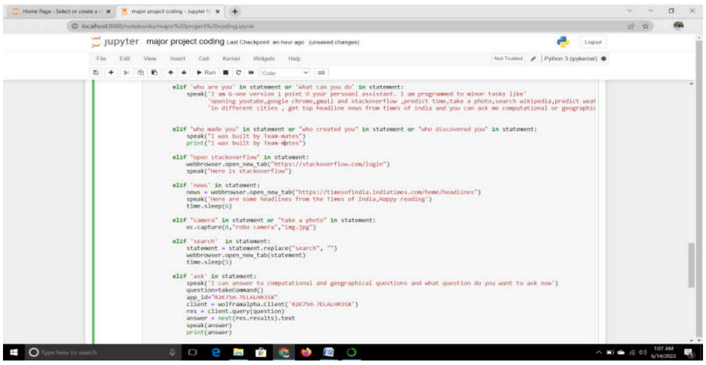
```
In [*]: print('Loading your AI personal assistant - G One')

        engine=pyttsx3.init('sapi5')
        voices=engine.getProperty('voices')
        engine.setProperty('voice','voices[0].id')


        def speak(text):
            engine.say(text)
            engine.runAndWait()

        def wishMe():
            hour=datetime.datetime.now().hour
            if hour>=0 and hour<12:
                speak("Hello,Good Morning")
                print("Hello,Good Morning")
            elif hour>=12 and hour<18:
                speak("Hello,Good Afternoon")
                print("Hello,Good Afternoon")
            else:
                speak("Hello,Good Evening")
                print("Hello,Good Evening")

        def takeCommand():
            r=sr.Recognizer()
            with sr.Microphone() as source:
                print("Listening...")
                audio=r.listen(source)

                try:
                    statement=r.recognize_google(audio,language='en-in')
                    print(f"user said:{statement}\n")

                except Exception as e:
                    speak("Pardon me, please say that again")
                    return "None"
```


jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Logout

```
                    return "None"
                return statement

        speak("Loading your AI personal assistant G-One")
        wishMe()


        if __name__=='__main__':


            while True:
                speak("Tell me how can I help you now?")
                statement = takeCommand().lower()
                if statement==0:
                    continue

                if "good bye" in statement or "ok bye" in statement or "stop" in statement:
                    speak('your personal assistant G-one is shutting down,Good bye')
                    print('your personal assistant G-one is shutting down,Good bye')
                    break



                if 'wikipedia' in statement:
                    speak('Searching Wikipedia...')
                    statement =statement.replace("wikipedia", "")
                    results = wikipedia.summary(statement, sentences=3)
                    speak("According to Wikipedia")
                    print(results)
                    speak(results)

                elif 'open youtube' in statement:
                    webbrowser.open_new_tab("https://www.youtube.com")
                    speak("youtube is open now")
                    time.sleep(5)

                elif 'open google' in statement:
```

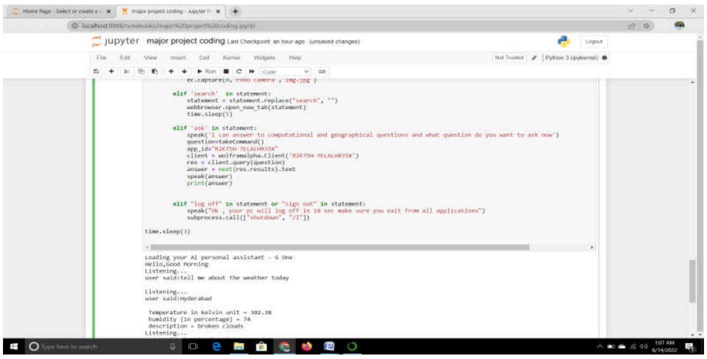

jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Logout

```
                elif "weather" in statement:
                    api_key="8ef61edcf1c576d65d836254e11ea420"
                    base_url="https://api.openweathermap.org/data/2.5/weather?"
                    speak("whats the city name")
                    city_name=takeCommand()
                    complete_url=base_url+"appid="+api_key+"&q="+city_name
                    response = requests.get(complete_url)
                    x=response.json()
                    if x["cod"]!="404":
                        y=x["main"]
                        current_temperature = y["temp"]
                        current_humidiy = y["humidity"]
                        z = x["weather"]
                        weather_description = z[0]["description"]
                        speak(" Temperature in kelvin unit is " +
                              str(current_temperature) +
                              "\n humidity in percentage is " +
                              str(current_humidiy) +
                              "\n description  " +
                              str(weather_description))
                        print(" Temperature in kelvin unit = " +
                              str(current_temperature) +
                              "\n humidity (in percentage) = " +
                              str(current_humidiy) +
                              "\n description = " +
                              str(weather_description))

                    else:
                        speak(" City Not Found ")



                elif 'time' in statement:
                    strTime=datetime.datetime.now().strftime("%H:%M:%S")
                    speak(f"the time is {strTime}")

                elif 'who are you' in statement or 'what can you do' in statement:
```



jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code

```
        elif 'who are you' in statement or 'what can you do' in statement:
            speak('I am G-one version 1 point O your persoanl assistant. I am programmed to minor tasks like'
                  'opening youtube,google chrome,gmail and stackoverflow ,predict time,take a photo,search wikipedia,predict weat
                  'in different cities , get top headline news from times of india and you can ask me computational or geographic


        elif "who made you" in statement or "who created you" in statement or "who discovered you" in statement:
            speak("I was built by Team-mates")
            print("I was built by Team-mates")

        elif "open stackoverflow" in statement:
            webbrowser.open_new_tab("https://stackoverflow.com/login")
            speak("Here is stackoverflow")

        elif 'news' in statement:
            news = webbrowser.open_new_tab("https://timesofindia.indiatimes.com/home/headlines")
            speak('Here are some headlines from the Times of India,Happy reading')
            time.sleep(6)

        elif "camera" in statement or "take a photo" in statement:
            ec.capture(0,"robo camera","img.jpg")

        elif 'search'  in statement:
            statement = statement.replace("search", "")
            webbrowser.open_new_tab(statement)
            time.sleep(5)

        elif 'ask' in statement:
            speak('I can answer to computational and geographical questions and what question do you want to ask now')
            question=takeCommand()
            app_id="R2K75H-7ELALHR35X"
            client = wolframalpha.Client('R2K75H-7ELALHR35X')
            res = client.query(question)
            answer = next(res.results).text
            speak(answer)
            print(answer)
```



jupyter major project coding Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code

```
            ec.capture(0,"robo camera","img.jpg")

        elif 'search'  in statement:
            statement = statement.replace("search", "")
            webbrowser.open_new_tab(statement)
            time.sleep(5)

        elif 'ask' in statement:
            speak('I can answer to computational and geographical questions and what question do you want to ask now')
            question=takeCommand()
            app_id="R2K75H-7ELALHR35X"
            client = wolframalpha.Client('R2K75H-7ELALHR35X')
            res = client.query(question)
            answer = next(res.results).text
            speak(answer)
            print(answer)


        elif "log off" in statement or "sign out" in statement:
            speak("Ok , your pc will log off in 10 sec make sure you exit from all applications")
            subprocess.call(["shutdown", "/l"])

time.sleep(3)
```

```
Loading your AI personal assistant - G One
Hello,Good Morning
Listening...
user said:tell me about the weather today

Listening...
user said:Hyderabad

 Temperature in kelvin unit = 302.38
 humidity (in percentage) = 74
 description = broken clouds
Listening...
```

CONCLUSION

We have discussed a Voice Activated Personal Assistant developed using python. This assistant currently works online and performs basic tasks like weather updates, stream music, search Wikipedia, open desktop applications, etc. The functionality of the current system is limited to working online only. The upcoming updates of this assistant will have machine learning incorporated in the system which will result in better suggestions with IoT to control the nearby devices similar to what Amazon's Alexa does

FUTURE 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.

References

Websites referred:

- www.stackoverflow.com
- www.pythonprogramming.net
- www.codecademy.com

- www.tutorialspoint.com
- www.google.co.in

Books referred:

- Python Programming - Kiran Gurbani
- Learning Python - Mark Lutz

YouTube Channels referred:

- CS Dojo
- edureka!