Project Report

on

VOICE ASSISTANCE FOR DESKTOP

Submitted in partial fulfillment of the
requirement for the award of the degree of

Master of Computer Applications

GALGOTIAS UNIVERSITY
(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under Supervision

Mr. Ravinder Ahuja

Assistant

Professor

Submitted By:

Krishan Mohan Yadav (18SCSE1010076)

Shrey Chaudhary (18SCSE1010342)

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING,
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA

December 2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled **"VOICE ASSISTANCE FOR DESKTOPS."** in partial fulfillment of the requirements for the award of the computer science and engineering submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of July 2021 to Dec 2021 under the supervision of Ravinder Ahuja (Professor) Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering, Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Krishan Mohan Yadav 18SCSE1010076 Shrey Chaudhary 18SCSE1010342

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Mr. Ravinder Ahuja

Assistant Professor

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of Krishan Mohan Yadav 18SCSE1010076, Shrey Chaudhary 18SCSE1010342 has been held on _____ and his/her work is recommended for the award of computer science and engineering.

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: December 2021

Place: Greater Noida

ACKNOWLEDGEMENT

We had a great experience working on this project and we got to learn a plethora of new skills through this project. However, it would not have been possible without the kind support and help of many individuals. We would like to extend our sincere thanks to all of them. We are highly indebted to the teachers and especially Mr. Ravinder Ahuja sir for their guidance and constant supervision as well as providing necessary information regarding the project and also for their support in completing the project. We would like to express our gratitude towards our parents and friends for their kind cooperation and encouragement which help us in the completion of the project.

Place - GREATER NOIDA

Krishna Mohan Yadav (18SCSE1010076**)**

Shrey Chaudhary (18SCSE1010342)

ABSTRACT

This Project Report is based on voice intelligent assistance. This tool is used for searching purposes, remainders just by using voice commands. If we want to access any document or file, we can do it by voice recognition. But there are certain limitations that the user should set some voices as a security option. If the user spells out the word it automatically types in the required field. It recognizes the speech and searches the appropriate content in the database and retrieves it. The user should select the appropriate language for the virtual assistant to understand. If any wrong or invalid communication happens it invokes some messages in the dialog box. It is like a software agent which performs tasks and events based on commands. Voice Command and speech synthesis are enhancing the level of user interaction in applications. This trend is now approaching business applications like ERP Systems. Implementing an intelligent personal assistant (IPA) will empower the application not only by navigating users through the system, but it also enables the option to navigate and explain data to the users through speech synthesis. In this paper, the architecture for a new IPA was worked out and a prototype for a web-based ERP-System resting upon this architecture has been developed. This IPA is able to listen to voice commands, to interact by opening a report and giving a brief summary via speech-to-text to the user, and explaining the most important information in the relevant context of the displayed KPIs. This prototype will be used for further researchers in the combination of IPA and data analytics.

Table of Contents

List of Tables:

Table for student data:

S.N	Name	Adm. No	Enrol. No	Branch/ Year	Semester
1.	Krishan Mohan Yadav	18SCSE1010076	18021011326	CSE/ 4th	7
2.	Shrey Chaudhary	18SCSE1010342	18021011574	CSE/ 4th	7

Table for Guide data:

S.n	Name	
1.	**Mr. Ravinder Ahuja**	

Acronyms

B.Tech.	Bachelor of Technology
M.Tech.	Master of Technology
BCA	Bachelor of Computer Applications
MCA	Master of Computer Applications
B.Sc. (CS)	Bachelor of Science in Computer Science
M.Sc. (CS)	Master of Science in Computer Science
SCSE	School of Computing Science and Engineering

CHAPTER-1

Introduction

1.1 Introduction

A virtual assistant is an application that understands voice commands and completes tasks for a user. Virtual assistants are available on most smartphones and tablets, traditional computers, and even standalone devices like the Amazon Echo and Google Home. They combine specialized computer chips, microphones, and software that listen for specific spoken commands from you and can answer back in a voice that you select.

A virtual assistant is a *technology based on artificial intelligence*. The software uses a device's microphone to receive voice requests while the voice output takes place at the speaker. But the most exciting thing happens between these two actions. It is a combination of several different technologies: voice recognition, voice analysis, and language processing. When a user asks a personal assistant to perform a task, the natural language audio signal is converted into digital data that can be analyzed by the software. Then this data is compared with a database of the software using an innovative algorithm to find a suitable answer. This database is located on distributed servers in cloud networks. For this reason, most personal assistants cannot work without a reliable Internet connection. With the increasing number of queries, the software's database gets expanded and optimized, which improves voice recognition and increases the response time of the system. Amazon succeeded in taking an important step towards the development of an entire infrastructure for virtual assistance in 2014 as the company introduced Echo, the first Smart Speaker. Later, Google and Microsoft created their intelligent loudspeakers – Google Home and Harman Kardon. Apple entered the race with its Apple Homepod only in 2018. Smart Speakers are stylish objects which are

used in the home environment and combine the functionality of a music speaker and an intelligent assistant. Equipped with an innovative voice recognition system, they are able to recognize the user's voice from a distance of more than 5 meters. Such smart speakers differ from other common gadgets, which are personal items that have to be taken along all the time. Smart speakers are designed as stationary entertainment objects for the whole family and can be used by all family members. It is not surprising that the voice recognition software, on which a smart speaker is based, feels most comfortable in its own digital ecosystem. Google Assistant is at its best with Google Apps, Google Home, and other Google devices. And Alexa, the voice software from Amazon, behaves the same on gadgets from the Amazon family.

The concept of all voice recognition systems is similar, but each system has its own unique features. Our project has the capability to play music, open different software which is installed on desktop/laptops such as PowerPoint, ms word, notepad, paint, etc. It can also play music, using this we can send mail to our friends and family.

1.2 Formulation of Problem

We are all well aware of Cortana, Siri, Google Assistant, and many other virtual assistants which are designed to aid the tasks of users in Windows, Android, and iOS platforms. But to our surprise, there's no such virtual assistant available for desktops that can do day-to-day tasks and can operate the software on voice commands.

1.2.1 Tools and Technology Used

Hardware Requirement:

- ➢ i5 Processor-Based Computer or higher
- ➢ Memory: 1 GB
- ➢ Hard Drive: 10 GB

- ➢ Monitor
- ➢ Internet Connection
- ➢ Mic
- ➢ Speaker

Software Requirement:

- ➢ Windows 7 or higher
- ➢ Visual Studio with python
- ➢ Google Chrome Browser

MERITS OF PROPOSED SYSTEM:

- ❖ Recognizes voice commands.
- ❖ Convert languages
- ❖ Controls various applications of devices.
- ❖ Send mails and WhatsApp messages
- ❖ Helps to plan your whole day.
- ❖ Reminds you of important things inaccurate situations or locations.
- ❖ Provides services regarding your location
- ❖ Store various information.
- ❖ Tell weather report
- ❖ These applications make small and smart hand-held devices to combine multiple features.

CHAPTER-2

Literature Survey

The Voice recognition technology system was developed before Apple's Siri was introduced in 2012. In 1963 at the Seattle World's Fair IBM delivered a tool named Shoebox. It can easily find and recognize 1,011 words, which is equal to the vocabulary knowledge of a 3-year-old student. It was actually like the size of a shoebox and performs some mathematical operations and can easily recognize 16 spoken words and also digits from 0-9. The rapid growth of machine learning and Artificial Intelligence made possible the growth and development of voice recognition methodology. There is a lot of human beings who exist in this society who uses speaking. An average human can be able to pronounce about 150 words in a minute, but can only type 40 words. This explains why voice technology should be taken immediately by IT companies. It takes a lot of time for humans to write a paragraph than reading it. The proposed system was based on voice recognition in commands which converts the speech to text. The user gets easy access to speech than getting typed in the keyboard. The user spells the commands in the voice so that the user should have good quality microphones. The user should spell the word incorrect pronunciation which should be understandable to the voice recognition system. But there exists some confusion in the system, for example, the words sun and son have the same pronunciation so the system finds it difficult to choose which word. So that it provides some powerful microphones to be audible to the system. The voice is a convenient and efficient mode of communication. Most people prefer to use speech rather than using text- based. The basic overview of the proposed system is it gets an input signal in the form of voice. It passes on to the feature extraction and gets into the decoder. In the decoder it consists of two main models, they are acoustic model and language model. The decoder will decode the input. After all the processing going in the decoder, it gives the specified output. The classification of the speech recognition systems is types of speech utterance, types of speaker model, types of vocabulary. It uses LGB algorithm functions which will train the VQ codebook.

Actual collaborator concentrates on center around an assortment of examination perspectives, for example, the climate language handling and recognition program procedures, virtual help interface and its client viewpoint, and the execution of the virtual help for explicit purposes, for example, e-learning or on the other hand medical care administrations. Contrasted with "truly

associated specialists", the genuine computerized specialists. However, generally,

papers simply feature modernized assistance as a programming program, e.g., a piece of a current structure be good for system typical language and taking in methodologies from the text or nearby language. different notable names for such age applications are "chatbot". Without this explanation, another expert gives morphology to virtual aides, so helps can be human, zoomorphic, on the other and utilitarian. Further, the ones VAs can have explicit sorts of pathways depending upon their morphology: for instance, human VA may in like way have the entire edge, or hands and head. Such VAs are secluded from images, tending to and reflecting human individuals, similar to images in Second Life. For the present circumstance paper, suggesting VAs, tended to by a high-level human individual . In extension to the VA key survey, concentrates on checking out human VAs base on the impression of customers of these modified expert characters. In their customer study, Shibam attempted the effect of twofold performers like individuals in the delight and interest of e-understudies. They did no more track down an incredible arrangement qualification among the trained professionals, and the understudies had been given the vague convenience without the presence of any subject matter expert. Etemad has concentrated on ten stand-out angles in light of a drawn out customized help customized adaptation made on a café site. the creator's investigation what's more accommodation to the individual study had the indistinguishable

destinations as the resulting examine, with centering at the utilization and character of aids. The impacts affirmed that being useful and fascinating immensely affected clients being equipped for return to the eating spot's site. In expansion, the magnificence element of virtual help and client pleasure has solid associations, just as trust and benefit. Different requests, which use and test aids, too focus on the internet business and showcasing highlights . Their examinations questions

point to help's impact at the extraordinary of sites. The proposed work will consideration all the more profoundly individuals virtual assistance discussion and the need for individual data, prompting questions along with 'How do customer investigate records about help and the way is insights respected with regards to computerized help?'. The review looks at the utilization of various aids, as far as not just language handling ways, yet in addition help execution what's more experience. to see according to a client's point of view. It will give an initial feeling on the individual all together VA cooperation and its effect on the conduct of client data. The main outcomes introduced in this exploration paper center around the visual norm of the program and the visual norm of the task.

Problem Formulation:-

- **Speech to Text**: User will ask the computer to run the command by giving input as speech.
- **Command Execution**: Based on the command received from the user, the system will execute the command (if available). e.g. Open Notepad, Paint, Google Chrome, etc. System accept various command such as the opening of specific applications, writing a note and saving it, opening web URL, Search for any query or details, and shutdown & Restart command.
- **Text to Speech**: Once a command is received, the application speaks the command which makes the user experience more interactive with the system.

Currently, the project aims to provide the Linux Users with a Virtual Assistant that would not only aid in their daily routine tasks like searching the web, extracting weather data, vocabulary help and many others but also help in automation of various activities. In the long run, we aim to

develop a complete server assistant, by automating the entire server management process - deployment, backups, auto-scaling, logging, monitoring and make it smart enough to act as a replacement for a 6 general server administrator.

As a personal assistant, 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.

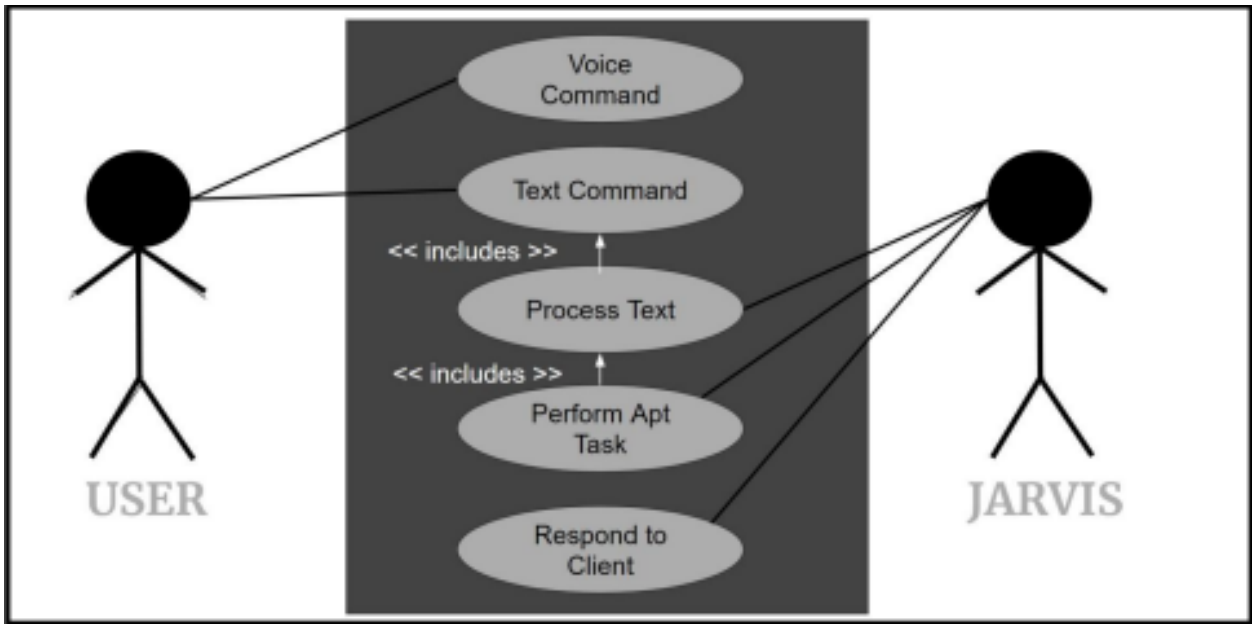
SCOPE

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

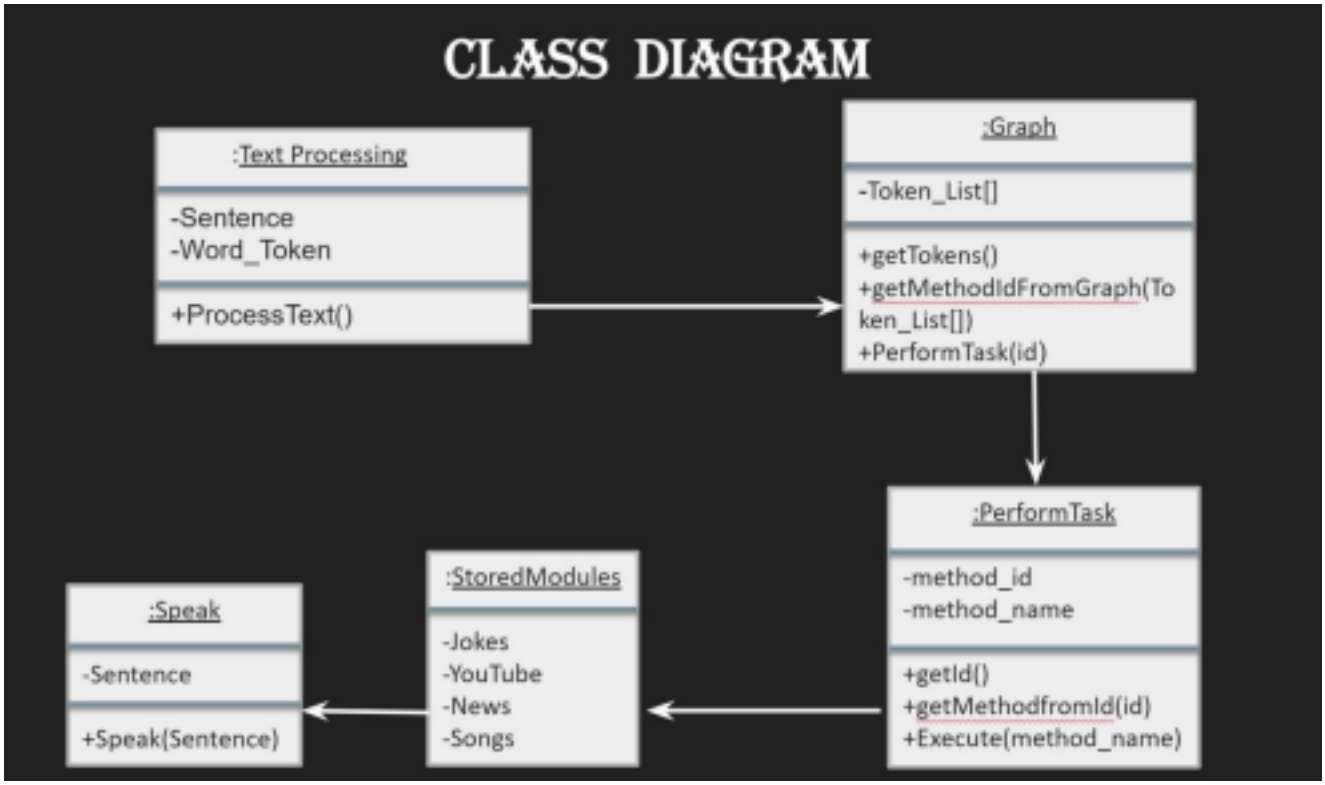
1. Search Engine with voice interactions
2. Medical diagnosis with Medicine aid.
3. Reminder and To-Do application.
4. Vocabulary App to show meanings and correct spelling errors.
5. Weather Forecasting Application.

There shall be proper Documentation available on its Official Github repository for making further development easy and we aim to release our virtual assistant as an Open Source Software where modifications and contributions by the community are warmly welcomed.

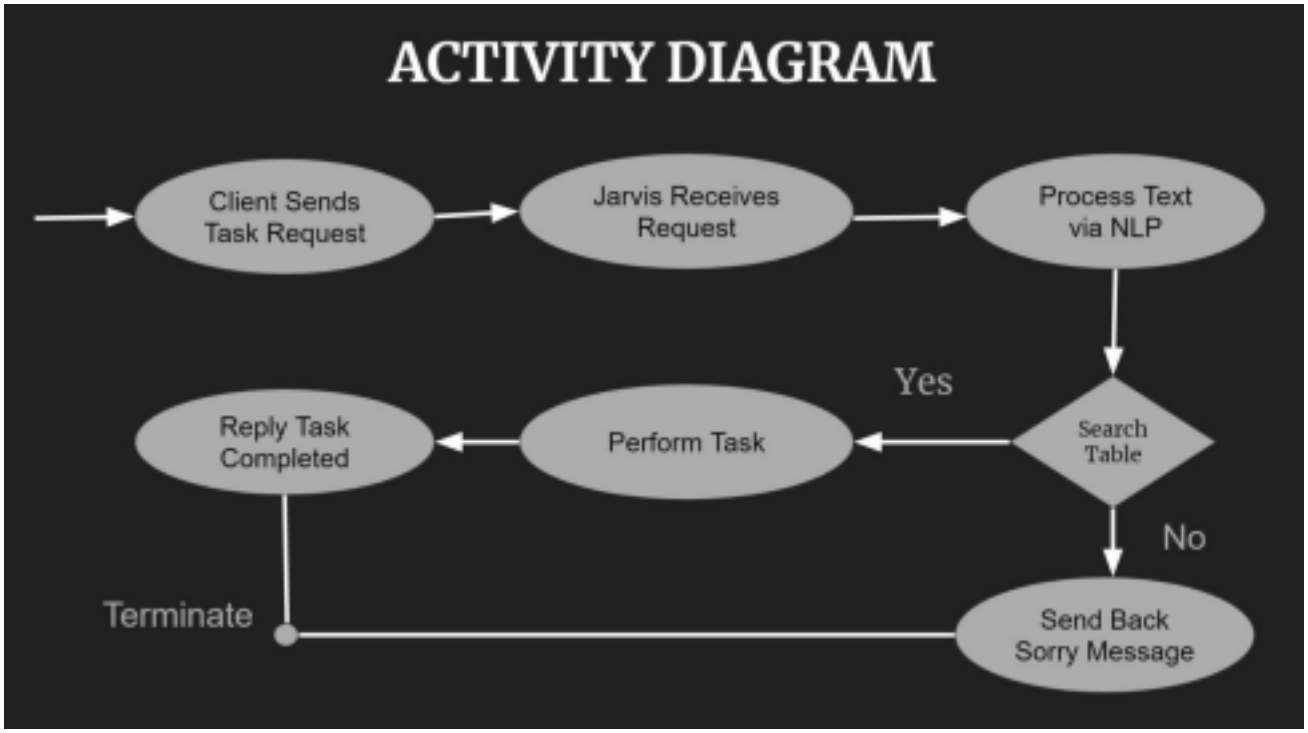
Project Diagrams:



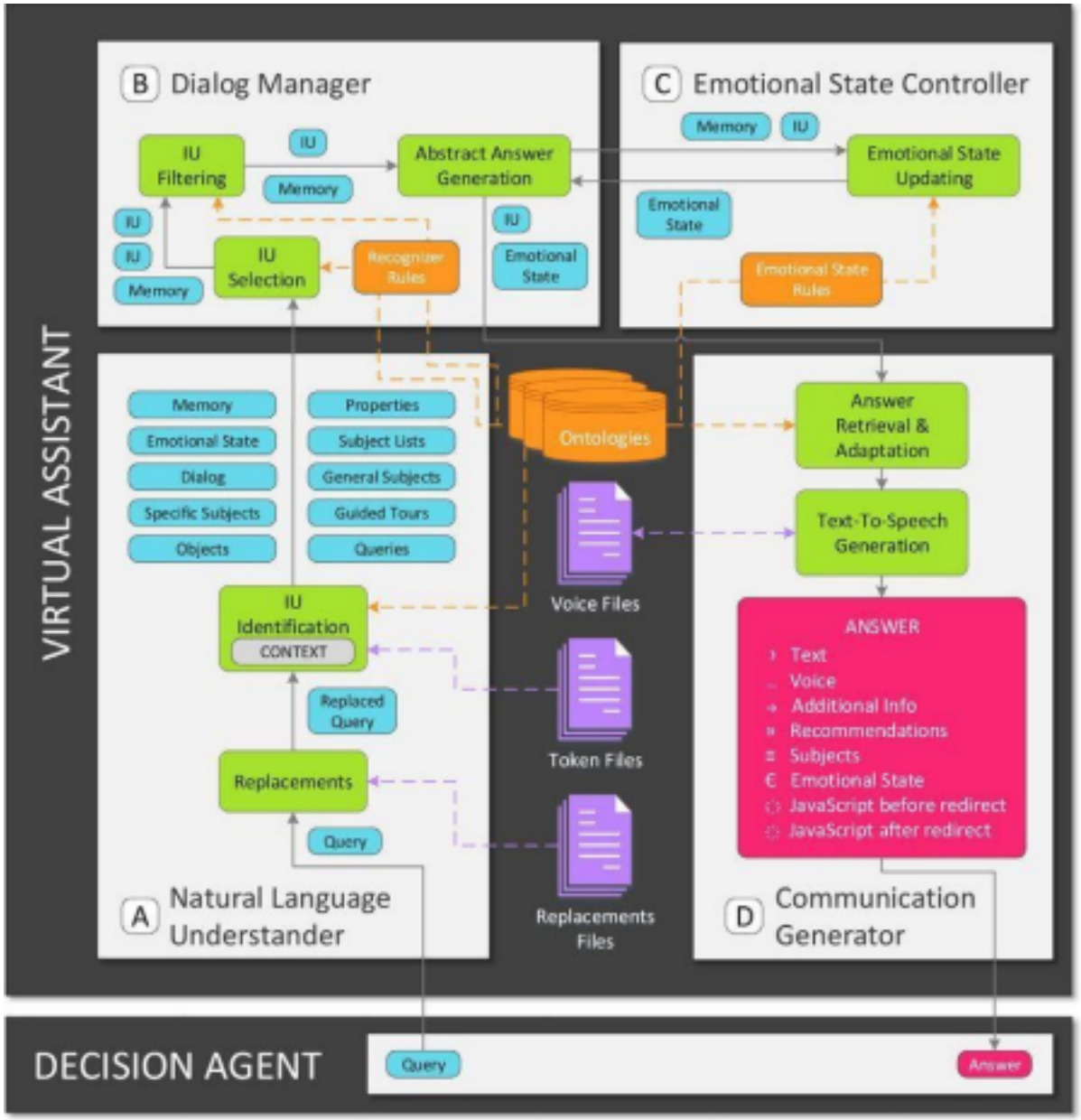
Class Diagram:



Activity Diagram:



Architecture Diagram



CHAPTER 3

FUCTIONALITY OF THE PROJECT

The program should firstly be started on the Android phone; the initial mode of the program is Voice mode since this program aims at making a voice assistant program. However, if there are users who prefer to operate in text mode by inputting the text manually, the text mode is also available. After the program has been started, the user should have correct voice input "command/request" to make those functions work properly. And this program includes the functions and services of: calling services, text message transformation, mail exchange, alarm, event handler, location services, music player service, checking weather, Google searching engine, Wikipedia searching engine, robot chat, camera, Bing translator, Bluetooth headset support, help menu. The details below explain how those functions work and different possibilities while facing different commands.

- Calling service, the calling function allows the users to give a call to the person in the contacts. By giving a correct command with the calling request to a stored person, the Android phone will check the contact list and get the phone number of the person, then successfully direct to the phone number found in the contacts.

- Text Message transformation, the text message transformation enable customers able to send the SMS to the person in the contacts. By giving a correct command contains the request keyword to send SMS together with the destination person; the program will navigate to the sending message function on the mobile phone with the phone number, message content. The message will be sent to the destination immediately if the user selects to send it with the correct content.

- Mail exchange, customers are able to send the mail to the person with mail address in the contacts. By giving a correct command contains the mail request keyword together with the destination person; the program will switch to the sending mail function on the mobile phone with the mail address and mail content. If the content is correctly detected,

the mail will be received by the recipient after the user selects to send the mail, otherwise the user can modify the mail content if the voice recognition is not well detect the mail content.

- Alarm, as a basic function on the mobile phone, the user could simply set the alarm through the command with the setting alarm keyword and a specific valid time. When the alarm request and time are detected, the program will set the alarm to the given time with dedicated hour, minute and second; when the time comes up, the alarm will be trigged with a alarm bell and an alert notification which the user can choose to stop the alarm, otherwise the alarm will keep working and the song will always be playing.

- Event handler, the application allows the user to set as many events as they want. Customers set the events with the content and title, the program switch to the event handler interface with the content and the title, and the event will be stored immediately if the user ensure the event. With the stored events, the event handler makes the events available for the user to check all events, check one event, modify the selected event and delete all events.

- Location services, location services works in two categories depending on the request. If it has been required to present the current location of the user, the location services check the GEO info by using the Google Map Service and give back the result as a map with the current location. If it has been required to provide the route trace from the current position to a specific city, the location service check the GEO info of both the origination and the destination, and provides the direction on the map with a route indicating how to get to the destination from the origination.

- Music player service, the music player offers the services to the user to play a named or random song in the pre-stored song list depending on the request. The music player service will play the specific song according to the name given by the user, the music

player check the music list and identify the song, then presenting to the user. The music player service will play a randomly picked song through the list if the user doesn't provide the song that he or she wants. The music player traces through the music list and get one from it for playing to the user. The music player could be also be stopped or paused while it is playing a song. By giving the correct commands, the working music player will be paused or stop playing.

- Checking weather, weather service provides the user the weather condition in different city on different dates. This service works in the same logic and gives back different result depending on the requested date and city. The weather service return the current date weather condition of the current location with the humidity, wind speed, temperature scope and display in a formalized entity which can be easily read by the user if the local weather for current date weather is required. The weather service return the next four days' weather condition of the current location with the date, wind speed, temperature scope and display in a formalized entity which can be easily read by the user if local weather for other dates except today's weather is required. The weather service return the current date weather condition of the given city with the humidity, wind speed, temperature scope and display in a formalized entity which can be easily read by the user if weather for current date weather for the given city is required. The weather service return the next four days' weather condition of the given city with the date, wind speed, temperature scope and display in a formalized entity which can be easily read by the user if weather for next for days of the given city is required.
- Google searching engine, the search engine enable the use to search anything on Google. By detecting the search keyword and search request, the Google search engine will returns the result list displayed on the browser on the mobile phone.

- Wikipedia searching engine, the search engine enable the use to search anything on Wikipedia. By detecting the search keyword and search request, the Wikipedia search engine will returns the Wikipedia result displayed on the browser on the mobile phone.

- Robot chat, the robot chat enables the user to chat with the Android phone to have fun. The chat mode is initially closed and will be required to activate it with the corresponding command. After entering the chat mode, a text response will given by the mobile phone whenever the user speaks to it; the response, however, were predefined and stored in the database. For each request, the program will define the request category and randomly pick a response from the response pool depending on the request category.

- Camera, the camera function enables the user to capture the current view with the camera on the mobile phone. When the camera is activated by the user, the user can selects to use the front or back camera on the mobile phone manually, and the picture will be taken by the camera if the user selects to photograph the current view, an instant picture for previously taken will be displayed in the program for viewing as a entity, and the picture will be stored in the Gallery for later checking.

- Bing translator, the translator will provide the user both the original text and the translated text depending on the objective language the user given. The user gives the original text and the object language the he wants; then the translator will give the result back of a translated text based on the original text and required language. Meanwhile, there have been 25 object languages stored in the database which the user can enjoy and the original text should be in English to use the translate function.

- Bluetooth headset support, the Bluetooth headset support makes the program well work especially the phone is playing music or the surrounding is noise which affect the voice recognition. Since it is not possible to do the voice recognition while the music player is playing, the Bluetooth will be loaded and available to the user, the user can select to turn on or turn off the Bluetooth function, and the Bluetooth headset support makes it possible to speak to the headset rather than the mobile phone if the Bluetooth is enabled.

- Help menu, the help menu provides the user a help list to each function in this program. The user can choose the help menu manually or over the voice if the user doesn't know how to work with the functions. While the help menu is opened, the help menu gives the examples and explanation of how to work with different functions, the examples clearly show how to work with the function and the user can simply imitate the example to work with different functions.

Implementation and Description of Modules:

Implementation:

```
import pyttsx3 #pip install pyttsx3
import speech_recognition as sr
engine = pyttsx3.init('sapi5')
voices =
engine.getProperty('voices')
engine.setProperty('voice', voices[1].id)

def speak(audio):
engine.say(audio)
engine.runAndWait()
def takeCommand():
#It takes microphone input from the user and returns string
output r = sr.Recognizer()
with sr.Microphone() as
source:
print("Listening...")
speak("listening ...")
r.pause_threshold = 1
audio = r.listen(source)
try:
print("Recognizing...")
query = r.recognize_google(audio, language='en-in')
print(f"User said: {query}\n")
except :
print("Say that again please...")
query= "None"
```

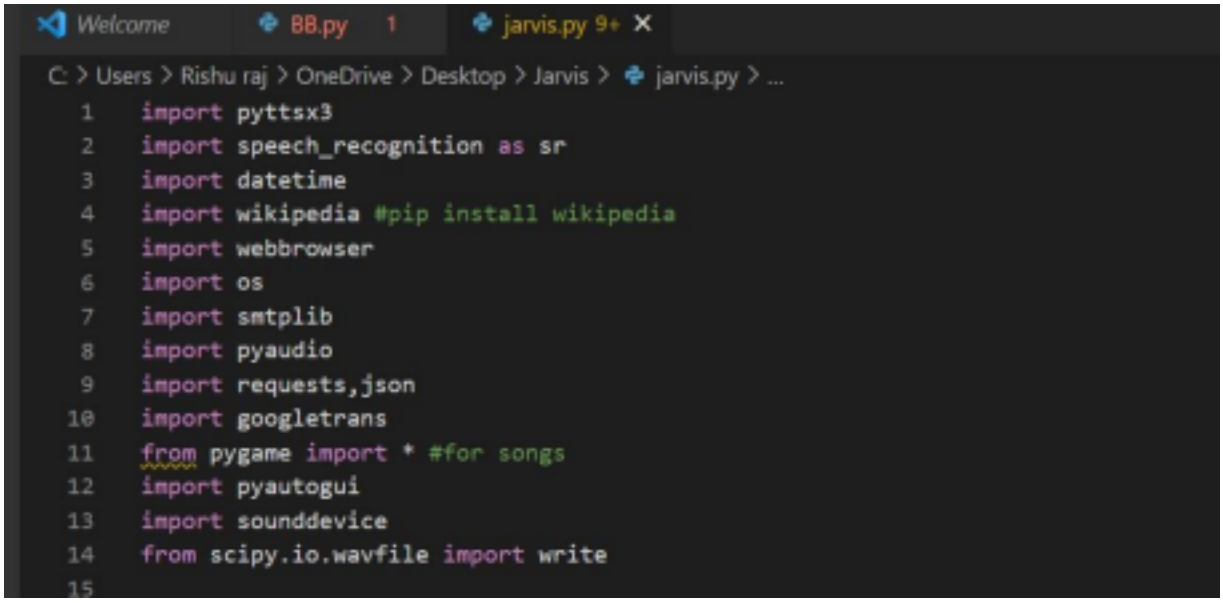
return query

Description:

pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline, and is compatible with both Python 2 and 3.

Speech_recognition :Library for performing speech recognition,with support for several enginesand APIs, online and offline.

Speech recognition engine/API support:



We are using these above-mentioned modules in our program. These modules are installed byusing “pip” command. But here in pyaudio module first we have to install “pipwin” module using“pip install pip win” then, by using “pipwininstallpyaudio” command we will be able to install “pyaudio” in our system.

```
engine = pyttsx3.init('sapi5')
voices = engine.getProperty('voices')
# print(voices[1].id)
engine.setProperty('voice', voices[1].id)


def speak(audio):
    engine.say(audio)
    engine.runAndWait()
```

This above code is responsible for converting the text into speech and make your device to deliver the speech. Microsoft Speech API (SAPI5) is the technology for voice recognition and synthesis provided by Microsoft. Starting with Windows XP, it ships as part of the Windows OS. If youare using a different OS, please consult the Microsoft Speech Technologies or the Speech SDK 5.1 sites. Here we create a function “speak” to use it this feature where we want.

```
def takeCommand():
    #It takes microphone input from the user and returns string output
    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        speak("listening ...")
        r.pause_threshold = 1
        audio = r.listen(source)

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

    except :

        print("Say that again please...")
        query= "None"
    return query
```

In this “takeCommand” function we are using speechRecognition module to recognize the command given by the user to the virtual assistance. And it will provide us desired output as per the given command.

```
def sendEmail(to, content):
    server = smtplib.SMTP('smtp.gmail.com', 587)
    server.ehlo()
    server.starttls()
    server.login('imkmyadav@gmail.com', 'kmyadav@')
    server.sendmail('imkmyadav@gmail.com', to, content)
    server.close()
    speak("Email has been sent!")

def youtube():
```

This function is used to send the Email. We send email using module name “smtplib”. **smtplib** module, which defines an SMTP client session object that can be used to send mail to any Internet machine with an SMTP or ESMTP listener daemon. Here is a simple syntax to create one SMTP object, which can later be used to send an e-mail − import **smtplib** smtpObj = **smtplib**.

```
def translator():
    speak("tell me sentence please..")
    from googletrans import Translator
    text = takeCommand()
    translator= Translator()
    speak("In which language you want to translate ...")
    m=takeCommand()
    key=m.lower()

    if 'french' in key:
        translation = translator.translate(text, dest='fr')
        print (translation.text)
        speak(translation.text)

    elif 'latin' in key:
        translation = translator.translate(text, dest='la')
        print (translation.text)
        speak(translation.text)

    elif 'italian' in key:
        translation = translator.translate(text, dest='it')
        print (translation.text)
        speak(translation.text)
```

In this function we used “googletrans” module which is used for translating one language in other language

OPERATIONS

- **Calling Service**: If the user wants to consume the calling service, he or she must have a command contains a valid name the calling keyword like “call”, ”make a phone”, then the call will be made if the person is found in the contacts. There are different ways to make a phone call, the list below shows the correct command to use the calling service.
  “Call Tom”, make a phone call to tom. The program will first capture the key words “call”, and then the program will continue to capture the person’s name “Tom” after the word “call”, then get all the contacts on the mobile and compare them one by one, if “Tom” is equal to the name that the user is give in the command, the phone call will be made to “Tom”.
  “I want to give a call to Lucy”, make a phone call to Lucy. The program will capture the command keyword “call” and the name “Lucy” and make a phone call to Lucy

- **Text Message Transfer:** If the user wants to use the application to send the text message, he or she must have a command with the SMS message keyword and a valid name, then the message will be send if the person is found in the contacts. They are different forms to send the message; the list below shows the correct command can do the message sending.
  “Send a message to LiLei Let's dinner together”, send a message to LiLei with the content “Let’s dinner together”, the program will capture the keyword “message” and the content “let’s dinner together”, then the program will check the mobile contacts and get the first phone number corresponding to “LiLei” and send the message to LiLei.

- Mail exchange: The user can send an email to the person in his contacts and with person’s email address. He or she must have a command with the email keyword like “Mail”, “Post” and a valid name; the email will be send if the person is found in the contacts. They are different forms to send the message; the list below shows the correct command can do the email sending.

  “Mail Bellis it will rain today”, send an email to Bellis the content “it will rain today”, the program will capture the keyword “Mail” and the content “it will rain today”, and then the program will check the mobile contacts and get the email address corresponding to “Bellis” and send the message to “Bellis”

  “Post Mimy a boy is waiting for you” send an email to “Mimy” with the content “a boy is waiting for you”
- Alarm: The user can use the set alarm command to set an alarm at the corresponding time. When the time is up, the alarm will be activated and play the sound; meanwhile, an alert will be presented for the user to stop the alarm.
  “"Set alarm to 10” the alarm will be set at 10 o’clock. The program will capture the setting command “Set alarm” and get the time command “10” and then the alarm will be active at 10AM.
  “Make time to 11:50” the alarm will be set at 11:50. The program will capture the command

"Make time" and the time "11:50", the alarm will be wake up at 11:50.

- Event handler: The application can allow the user set many events. He or she can set many events and be saved into the application's database by using the add event command, also he or she can view the event or delete the event by corresponding keywords like "Set up", "make up", "View one/all event(s)", "Delete". The list below shows the correct command can do the operation

  Add Event: "Set up a meeting at 10", the Program will first capture the keyword "set up" then the title of the event "a meeting" and the content "a meeting at 10", then the event activity will be start with the add event dialog, automatically fill with the title "a meeting" and content "a meeting at 10". Then the user should to choose the date time and add the event.

  View Event(s) "View all/one event(s)" / "Find event", the Program will start the event activity and present the event(s) based on the user choose to show the all events or one event. If the user chooses to show one event, the data picker will be shows up and the user can choose the event that he or she wants to present based on the date, otherwise, the application will show up all the events if the user wants to check all.

  Delete Events "Delete all events", the Program will delete all the events that in the application's database.

- Location services: The user can use this service to locate the user's position or get the routes to the destination by giving the city name. There are different ways to locate the position or navigate to a specific city. The use must use the keywords "where" and "I" or "my location" to let the application to know he or she wants to locate the current position. And the keywords "go to" and the name of the place to get the route to the destination.

- Music player service, the user can use this application to play songs, his or her command must contain keyword "play". If the user wants to play the specific song, he or she should also say the name of the song after "play", and the song should be exist in the SD-card memory. Or if the user wants to play a random song, he or she just needs to say "a song" instead of the song's name. During the playing, the user can pause or stop the song by giving the command "pause"

or “stop”.

- Checking weather: the user can use the application to check the weather for recent days in local place or specific location. He or she should say the keyword “weather”, then the user should notify the date that should be presented as “today/tomorrow/the day after tomorrow” if he or she wants to get the information about the other days otherwise the application will default set the date as today, and the user can also can choose to tell about the place name “in Malmo”, the application will check the weather belong to that place, otherwise the place will be set as locally

  Weather check today:

  “What's the weather for today”, the current weather condition for local place will be show.

  “What's the weather in Malmo”, the current weather condition for Malmo will be show.

  Weather check other days:

  “What's the weather next few days”, the forecast in next 4 days will be show.

- Google searching engine, the Google search engine is activated by the user commands which contain ‘Google’ or ‘Search’. By detecting the search keyword and search request, the Google search engine will returns the search result displayed on the browser on the mobile phone. “Google China”, the keyword ‘Google’ is detected and the result will be presented on the web browser by searching ‘China’ on Google. “Try to Google Java API”, the user can have the keyword Google in the middle of a request and the result of searching ‘Java API’ on Google will be displayed on the web browser. “Search for apple”, the user can also use the keyword ‘search’ to do the Google search, this command will have the result of searching ‘apple’ on Google

- Wikipedia searching engine, whenever the user wants to search any content in Wikipedia, it is possible to do in this program by having a command contain the keyword ‘define’. If ‘define’ is detected by the program, the program will automatically give the result by search the content after ‘define’ in Wikipedia. “Define Android”, the keyword ‘define’ is detected, and the program will return the result by searching ‘Android’ on Wikipedia. “Define true love”, the

keyword ‘define’ is detected, and the program will return the result by search the content after ‘define’, which is ‘true love’ on Wikipedia.

- Robot chat, the robot chat will work only after the chat mode is enabled which can be done with a command that contains keyword ‘chat’. After the chat mode is enabled, a response will be given every time when the user gives a request. The chat can be finished by the user commands contain the keywords of ‘finish/ disable/ end/ complete chat’.
  “Enable chat”, the keyword ‘chat’ will be detected and the chat mode will be enabled. Now the user can enjoy the chat by inputting any text he /she wants.
  “Let’s chat”, the keyword ‘chat’ will be detected and the chat mode will be enabled. Now the user can enjoy the chat by inputting any text he /she wants.
  “Finish chat”, the keyword ‘finish chat’ is detected and the chat mode will be disabled. When the user exits the chat mode, the program gets back the normal mode to receive and analyze the commands, and give correct response.

- Camera, the camera is started while the keyword ‘camera’ is detected. Therefore, the user who wants to operate with the camera will have to give a command with camera inside. After camera is started by the correct command, the camera itself will guide the user how to take photograph.
  “Open the camera”, as the keyword ‘camera’ is detected, the camera is started. And the user can work with the camera by clicking the different selection on the mobile phone.
  “Start the camera”, as the keyword ‘camera’ is detected, the camera is started. And the user can work with the camera by clicking the different selection on the mobile phone.
  “I want to use the camera”, as the keyword ‘camera’ is detected, the camera is started. And the user can work with the camera by clicking the different selection on the mobile phone.

- Bing translator, the user should have the keyword ‘translate’ / ‘how to say’ as the keywords to define this is a translate request, and ‘in’ as keyword to indicate the objective language. As the user have the command contains these keywords, the translator will return the result with the

text in the objective language.

"Translate I love you in Chinese", as 'translate' and 'in' are detected by the program, the program will call the translator with 'I love you' as the original text and Chinese as the objective language, the result will be the Chinese words of 'I love you'.

"How to say hello in Swedish", as 'how to say' and 'in' are detected by the program, the program will activate the translator with 'hello' as the original text and Swedish as the objective language, the result will be the Swedish text of 'hello'.

- Bluetooth headset support, the Bluetooth headset support will be enabled when the program is loaded. The user should firstly turn on the Bluetooth in the setting of the mobile phone, and the Bluetooth icon will be valid in the program after executing the program. The user will be required to plug in the Bluetooth headset and turn on /off it manually by clicking on the Bluetooth icon.
- Help menu, the help menu can be activated by manually select on the option menu or through the command. The commands should have 'help' as the keyword contained, then the help menu will be activated and the help menu provides the list of all functions with their explanation and examples to use it.

  "I want to check the help menu", if the users have the keyword 'help' contained in the command, it will be detected as a keyword and the help menu will be returned with a list of the functions, the functions are presented in two pages and user can scroll the pages by slipping the touch pad of the mobile phone; by selecting on each of the functions, the user can enjoy the details of the explanation and the examples of each function.

CHAPTER 4

RESULT AND DISCUSSION

CHOICE OF SOLUTION

This part discloses the real answer for build of the entire program. The capacities include:
Calling administrations, message change, mail trade, caution, occasion overseer, area administrations, music play administration, really looking at climate, looking through motor (Google, Wikipedia), camera, Bing interpreter, Bluetooth headset support, help menu and Windows sky blue distributed computing. As it has been shown in 2.3, the entire development of the program principally cover Android application improvement, the data set plan, web administration and distributed computing. The Android application, which executes and presents every one of the capacities, is built in Eclipse with Android improvement references. The program executes voice acknowledgment to catch the approaching solicitations. Making the fundamental movement and building every one of the capacities, carrying out the rationale to develop the entire program. Further by bringing the web administration on the Windows Azure Cloud, the order can be broke down with the capacity on the information base; relating reactions will be coordinated to explicit capacity in the program. Figure-7 shows the general plan of the program through UML.

The database is designed with MS SQL server. By creating different tables to store the data in different category, the data can be well stored, retrieved, updated or deleted. To well support the data process in web service, the database is uploaded on the Windows Azure Cloud. Web service, the web service is implemented in C# since it is placed on the Windows Cloud. The web service takes the incoming request as the parameter; analyze it by check the keyword contained in the request, and give correct response to the program. The same with the database, the web service is uploaded on the Windows Azure Cloud. Cloud computing, Windows Azure has been chosen as the cloud platform since it provide a three months' free use with a registered account. By establishing the database and creating the web services for intended use, the database and the web service are uploaded on the Cloud and, the data processing are going as cloud computing. The following

indicate the design for each individual function in this program.

- The programs start with the voice recognition, by implementing the RecognitionListener, it will capture the text every time the speaker speaks to it, then the generated text and send to the cloud (fig. 8)

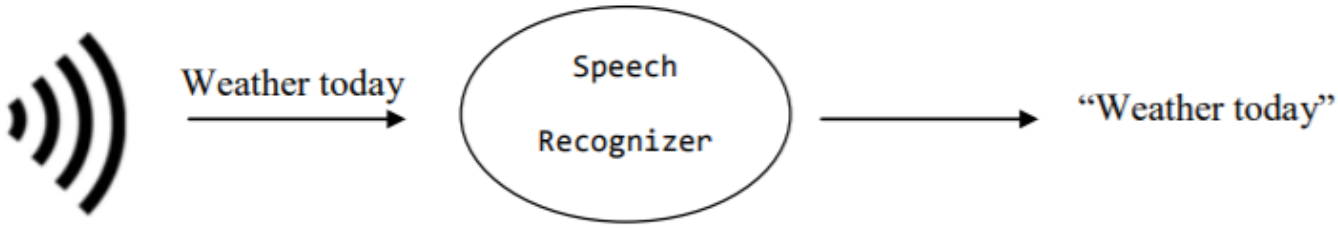
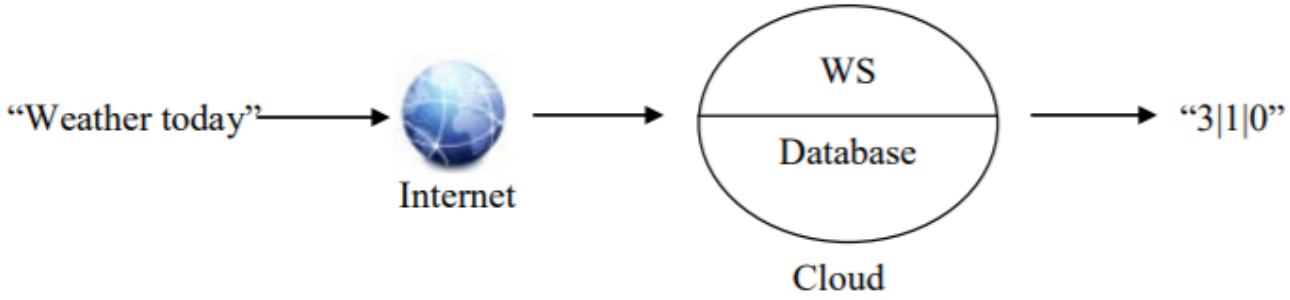


Figure-8

- The azure cloud which is an open cloud platform, where the software, database, web service can be placed there for future use. In this program, the web service and database are uploaded on the azure cloud for executing and maintenance (see Figure-9).



- The web service is written in C# and connects to the cloud database, the captured text will firstly be sent to the cloud as a parameter to call the analysis method and the method will check the keywords from the database keyword library. When the keyword is identified, it will implement different operations depending on the keyword category and give corresponding response that follows the protocol.

- The database was created in MS SQL server and uploaded on windows azure cloud through windows azure database manager, it defines the different keyword categories depending on the functions, the keywords for each category and response for different keywords category.

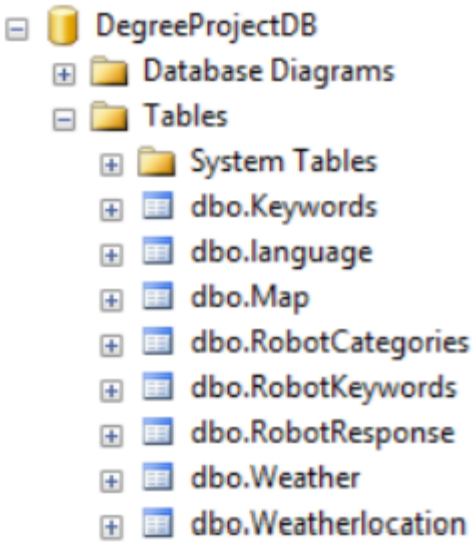


Figure-10

- The database has been designed into eight tables, each table contain different information for each category, the “Keywords”, “Language”, “Map”, “Weather”, “Weatherlocation” tables is used for the application to identify the different command, “RobotCategories”, “RobotKeywords”, “RobotResponse” table are used for the robot chat. The following chapters describe each of the table and what is intended usage. “Keywords” table: (see Figure-11)

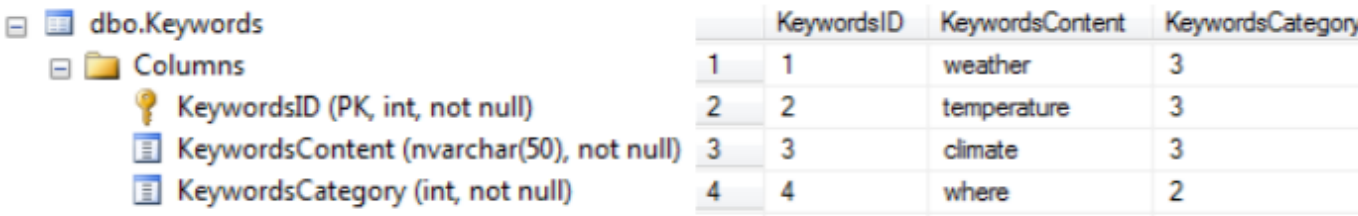
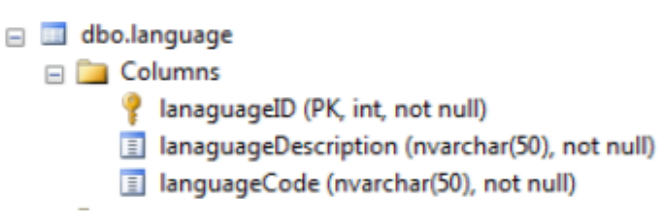


Figure-11

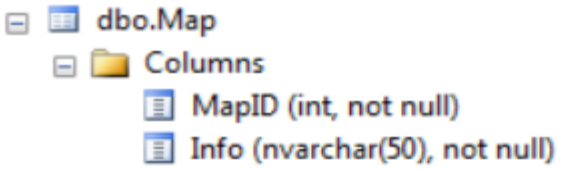
- The “keywords” table contains three columns to present the data, the “KeywordsID” column is used to specify the different keyword in its unique ID, “KeywordsContent” column is used to save the keyword info and the “KeywordsCategory” classify the content into different category. “Language” table: (see Figure-12)



	lanaguageID	lanaguageDescription	languageCode
1	1	traditional chinese	zh-cn
2	2	simplified chinese	zh-cn
3	3	japanese	ja
4	4	italian	it

Figure-12

- The “Language” table is used to discern the language and translate it to objective language code, the “languageID” column is used to specify the different language in its unique ID, “languageDescription” is used to describe the language and the “languageCode” is used to change the text-based language in to language code. “Map” table: (see Figure-13)

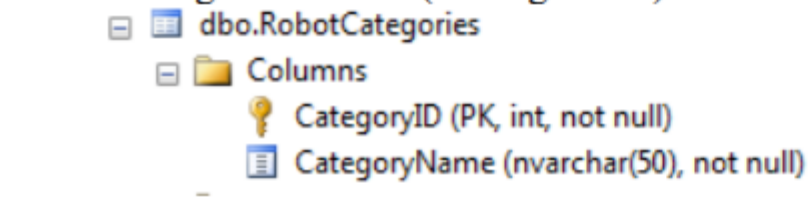


	MapID	Info
1	1	i
2	2	to

Figure-13

- The “Map” table is used to discern the user’s navigation proposes. The “MapID” is used to specify the different info in its unique ID and the “Info” table is used to specify the content.

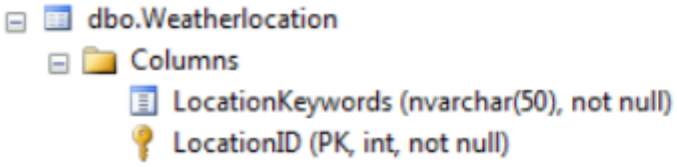
“RobotCategories” table: (see Figure-14)



	CategoryID	CategoryName
1	1	greet
2	2	age
3	3	sex

Figure-14

- The “Weather” table is used to discern the robot response category. The “KeywordCategory” is used to define the category of this content, the “TimeID” column is used to give the unique number to each “Time” content and the “Time” column is used to specify the time content.



	LocationKeywords	LocationID
1	in	1
2	around	2

Detailed solutions and implementation for each function depend on the request categories.

0. Chat Mode: The program will get the captured text and send it to the cloud web service, the cloud will loop over the robot chat keywords and identify the keyword category; the response will be randomly accessed through the response pool according to the keyword category, finally the program init the TextToSpeech engine from the Android system and generate the audio output with the response. [Code-0-1]

1. Chat Mode Switcher: the program will have a Boolean variable initiated to false. If the chat mode is enabled, the variable will be assigned as true and anything captured will be in the chat mode until the chat mode is finished. While the chat mode is exited, it gets back to the normal mode and analyzes the requested commands. [Code-1-1]

2. Location Service: The program will firstly distinguish the command in two different ways; one is to find the current location, another one is find the routes between the current location and the destination location. To find the current location, the program will check the location information from the device GPS Module and get the current Longitude and Latitude values, then start the MapActivity by assign the pair values and the mode “current”, present the maps for the user. To find the route to a specific destination, the program will also check the current location and get the GEO values, generate the target location name to an URL, read the GEO Information from the link [Code-2-1], with the GEO info for both the origination and destination, the program will start the MapActivity by assign the current location geo value and the remote location geo value with mode “Remote”. The map activity will generate that information to an URL and send to the Google map server, then get the route XML. And draw the route on the map.

3. Weather: the program will firstly check the command whether it has the specific city name, if the city name is obtained in the command, the program will send the city name to the Google map server and get the corresponding geo information with the longitude and latitude and set as a location to get the weather condition; otherwise the location will be the current location information from the mobile GPS Module, if no city name is given, the program will generate an URL by the location’s geo info, and get the corresponding weather condition XML from the Google weather Server. The program will also check the data info from the cloud response, if the user requires the weather for today, the program will present the first weather condition from the XML, otherwise, it will get the next four days conditions.[Code-3-1]

4. Wikipedia search: the program will replace the space in the search content to “+” and formalize the searching URL, and then switch to the search activity by calling ACTION_VIEW and give back the result as navigate to the previous obtained URL. [Code-4-1]

5. Calling service: the program will extract the name section from the response accessed from the cloud web service, then check through the contact list and get all the stored contacts [Code-5-1], further fetch all the details of the person with name, email, phone number [Code-5-2]. Identifying the

person and get the first phone number, and the system will make the phone call by calling the system ACTION_CALL intent and start the calling activity. [Code-5-3]

6. SMS: the program will extract the name section and the message content from the response accessed from the cloud web service, then check through the contact list and get all the stored contacts [Code-5-1], further fetch all the details of the person with name, email, phone number [Code-5-2]. Identifying the person and get the first phone number, and the system will send the message by calling the system ACTION_SENDTO intent and start the sending message activity. [Code-6-1]

7. Email: the program will extract the name section and the email content from the response obtained from the cloud web service, then check through the contact list and get all the stored contacts [Code-51], further fetch all the details of the person with name, email, phone number [Code-52]. Identifying the person and get the first email address, and the system will send the email by calling the system ACTION_SEND intent and start the sending email activity. [Code-7-1]

8. Google Search: the program will replace the space in the search content to "+" and formalize the searching URL, and then switch to the search activity by calling ACTION_VIEW and give back the result as navigate to the previously obtained URL. [Code-8-1]

9. Alarm: the program will extract the Hour and Minute parts from the response obtained from the cloud web service, set a calendar with the requested time of hour, minute and second. Then start the Alarm manager by calling the system ALARM_SERVICE with the settled calendar and broadcast. In addition, the broadcast is a trigger to activated an alert and the alarm music will by played when the alarm is activated by system action RTC_WAKEUP. [Code-9-1]

10. Music Player: When the program is loaded and initialized, it will call the system ACTION_MEDIA_SCANNER_FINISHED to scan all the media files on the SD card memory and save the file's path, id, title, and put all these attributes into a list[Code-10-1], the program will first extract the action command from the response obtained from the cloud web service, if the command

requires to playing music, it will further check whether the response contain with the song's name or not, if the request does not have a specified name of the song, the program will randomly pick a song from the list and start the music play service by given the path of the requested song, otherwise, the song's path will be obtained from the list by the song's name and start the music with start command[Code-10-2]. If the response contains the pause command, the program will set the music service at a pause state. As it is the same with pause, the stop command also will be sent in this way and the music player will stop playing the music. [Code-10-3].

11. Event handler, the program will firstly extract the command part to decide if the user wants to add or view or delete events. The event program will navigate to the event activity with the requested command. The layout of the event activity is designed through the XML file and different operations "Add/View/Delete" are set on the interface. By extending SQLiteOpenHelper and SQLiteDatabase, the events can be stored, and updated or deleted.

12. Camera: when the program receives the start camera command, it will start the Camera activity, then init the Speech Recognizer on that activity. After the user take photo by recognize the "Cheese" command and save the image into the SD card memory, a broadcast will be triggered to notify the system's gallery to refresh the photos. After the photo has been taken and stored, the camera activity is finished and give the image path back to main activity, and the main activity will present the image to the user based on the image path from the given path, the user also can touch on the preview image to view the image detail by start the ImageViewActivity.

13. Help: the program will navigate from the current activity to the help activity while the help menu is activated from the main option menu or by the detected command. The help activity contains a list of items correspond to each different function; they share the same outline with an icon, text explanation [Figure 13]. If any image button is clicked, it will switch to the help content activity with the corresponding name of the function. By getting the name of the function, the content activity will fill its content with the icon, title, and the examples to tell how to work with the function. The layout of the activity mainly been constructed with the TextView, ImageView, and ListView. [Code-13]

14. Translate: the program will get the target language code and the content text, then generate the original language code, target language code and the content text to a URL; start the URL and get the translate result from Bing, finally present the result with the original text and the translated text for user.[Code-14] 15. Bluetooth headset support: when the user plug-in the Bluetooth headset the system will send a broadcast to the program, the program will use a Bluetooth receiver to receive this broadcast then enable the button for user to select if use the Bluetooth or not.

Maintenance

After the program is completed, the program still needs long term maintenance to make it available and stable to execute. The program will be test after a certain period of time and debug each of the function and possible bugs, whenever a potential bug is detected; the program needs to be refined to a better design. Meanwhile, there will update and add more data to the database to increase the database capacity. Depending on the new keywords, responses, relevant data found that could be applied in this program; the database will always be improved and can handle more and more cases.

CHAPTER 5
CONCULSION AND FUTURE SCOPE

5.1 CONCULSION

- Project development and implementation

  As it has been previous stated, the program is mainly concerns with the techniques of Android development, Java programming, Database management, Cloud computing, different APIs for Google products, Bing translate and etc. The program is developed by two developers and follows the extreme programming model. During the eight weeks development, the developers did the same cycle in each phase of analyze requirements, construct design, implement the solutions in pair programming mode and test the result. The development is carried out as its primary planning which guide the work process of how to work with the program, how much time should the each of the developers spent in every week, the rescores needed for developing and how to handle the problems while it came up. The project was efficiently completed under the development model and the resources we found in early time were really useful when implementing the program.

- Project usage & prospect, potential

  The project is very useful and owns a large potential use in different industries. Although the program primary concerns more about how to do the personal assistant on Android phone using the voice, the concept of voice recognition can be applied in different industries as in many situations it will be more convenient, save a lot of time and helpful especially for those who have difficulty in working with manual operations. Thus, the concept is only for programming the Android application.
  For the program itself, it is a collection of 15 functions that are frequently used on a mobile phone. The user can enjoy different services within this platform. Therefore, it is easy to use

with simple operation compared with the traditional working strategies which the user should well know how to work with the mobile phone.

In addition, the program which works using the voice is helpful for those who prefer voice operation and those who have difficulty /disability with the manual operations. The primary objective of the program is to provide services using the voice, and it enables more people who can enjoy this program.

The prospect of the program can be more applications or products developed using the voice control, and it could in some sense change the working forms that is totally different from the traditional form. As people can easily operate and have a lot of fun from it, it owns an enlightened prospect as SIRI succeed in attracting people in the market.]

Here is found and researched the three Virtual The help with's terms of their client experience understanding Thus, played out a heuristic and customer assessment read. Starter results show that customers will as a rule see a single VA better, which maybe take into account the assistance's coarse discourse getting limit. Regardless, the viewpoints on the customers were through and through various. It should similarly be so it suggests that with performed tests all guides can't be attempted considering less understanding in its language taking care of advancement, which infers there are openings for future advanced testing processes. It will in general be helpful to survey those fundamental contacts with the recorded individual articulations and searches with the ultimate objective of offering express data on the buyer and private wishes working during composed exertion with the help. Following two or three years, everyone will separate Users 'viewpoint on greatness and its impact on the fun similarly as a conviction.

5.2 FUTURE SCOPE

Design Improvements

No program has a perfect design without any flaws; it is the same here in this program. Even though the program is completed with all the primary functions implemented and work properly, there are still many things that can be done with this program. As the future improvement, the potential work that can be implemented ranging from adding more functions to offering the user a more comprehensive, convenient program, refining the logic to make the program more humanized and easy to use, increase the database capacity and add more possible keywords, responses and data in this program, interface optimization and etc.

Additional Functions

Add more functions: although there have been 15 normal functions that are used really often with the mobile phone, there can be more functions which simplify our daily life and make it convenient to use. Functions as playing movies, checking stocks, exchange rate, downloading and uploading, installing APPs and etc, these can be the potential functions that make the program more comprehensive and people can enjoy more services in this program.

Database Capacity

Add database capacity and more humanized logical design; the program has a predefined logic to make it work with the corresponding commands. Thus, the user need to follow the structure of the commands, contain the dedicated keywords and well formalize the commands to work with each of the functions. In other words, the program is limited by the database capacity and no solution will be found if the user gives commands that are not readable by the program. Even if two commands have the same meaning and should get exactly same result set, the result might be that of one is working and the other one fails. Hence, the program is to some extent limited by the vocabulary and can be further optimized.

Humanized Voice Recognition

The more humanized the program is, more easier the user can use it. People should accept that even if developers constantly try to add more predefined commands, more responses to it, analyze and respond to the

command more intelligently, the program will never be completely comprehensive and contain all the possible circumstances that the users meets. Nevertheless, the program will certainly be improved and be more user-friendly if there can be more readable commands, more humanized structure and more intelligent response.

Improved Interface Interface

optimization, the interface can be further improved to make it nice to the users. Currently the interface design meets the basic requirement to present everything for this program, and the users are able to interact with the program through this interface, but the interface can always be optimized and more suitable constructed.

REFERENCES

[1] Mammone, R., Xiaoyu Zhang, and Ramachandran, R. (1996). solid speaker acknowledgment: a capacity basically based methodology. IEEE signal Processing mag,thirteen(five),fifty-eight.https://doi.org/10.1109/seventy nine.536825.

[2] Gartner. 2018. Gartner Says 25% of client care Operations Will Use virtual shopper Assistants through 2020. online article,

https://www.gartner.com/en/newsroom/pressreleases/2018-02-19-gartnersays-25-percent-ofpurchaser-service-operations-will-use-digitalcustomerassistants-by means of-2020 [3] Subtlety offers free voice notoriety on Dragon cell Assistant. (2014), 2014(5), 2. https://doi.org/10.1016/s0969-4765(14)70078-five

Nuance offers free voice reputation on Dragon cell Assistant. (2014), 2014(5), 2. https://doi.org/10.1016/s0969-4765(14)70078-five.

[4]. ColllinaSzy, J., Bundzel, M., &Zolotova, I. (2017). Implementation of wise software program using IBM Watson and Bluemix. Acta Electrotechnica Et Informatica, 17(1), fifty-eight-63. https://doi.org/10.15546/aeei-2017-0008.

[5]. Nedić, B. (2019). Gartner's pinnacle Strategy technology traits. proceedings on Engineering Sciences, 1(2), 433-442. https://doi.org/10.24874/pes01.02.04