*A Project Report On*

# “POLYMOD VIRTUAL COMPANION”

*Submitted in the Partial Fulfillment for the Award of degree for*

## Master Of Computer Applications

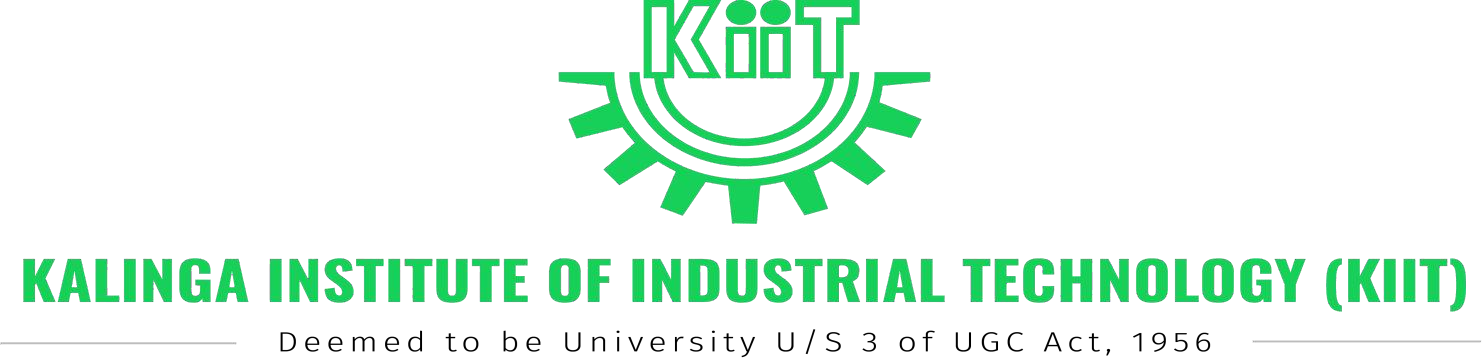
*Submitted By*

# A D A R S H A R O U T R A Y 2 2 7 0 0 1 8

UNDER THE GUIDANCE OF

**Prof. BIBHUTI BHUSAN DASH**

# SCHOOL OF COMPUTER APPLICATIONS



BHUBANESWAR, ODISHA – 751024 APRIL 2024

**CERTIFICATE OF ORIGINALITY**

This is to certify that the project report entitled “ Poly mod Virtual Companion” submitted to **School of Computer Applications, KIIT Deemed To Be University** in partial fulfillment of the requirement for the award of the degree of **MASTER OF COMPUTER APPLICATION(MCA) ,** is an authentic and original work carried out by Mr.ADARSHA ROUTRAY bearing Roll no. 2270018 under my guidance.

The matter embodied in this project is genuine work done by the student and has not been submitted whether to this University or to any other University / Institute for the fulfillment of the requirements of any course of study.

ADARSHA ROUTRAY Signature of the Guide Date: Date:



### KIIT DEEMED TO BE UNIVERSITY

School of Computer Application Bhubaneswar, ODISHA 751024

**CERTIFICATE**

This certifies that the project entitled

“POLYMOD VIRTUAL COMPANION“

submitted by

**ADARSHA ROUTRAY**

### 2270018

This is to certify that the project work entitled POLYMOD VIRTUAL COMPANION by ADARSHA ROUTRAY of

MASTER OF COMPUTER APPLICATION bearing roll no. 2270018 is an authentic and original work.

Signature Signature

(Internal Examiner) (External Examiner)

Date.................. Date..................

**DECLARATION**

I, Adarsha Routray ,bearing roll - 2270018 do hereby declare that the project report entitled **POLYMOD VIRTUAL COMPANION** submitted to **School of Computer Applications, KIIT University, Bhubaneswar** for the award of the degree of **MASTER OF COMPUTER APPLICATION (MCA) ,** is an authentic and original work carried out by me under the guidance of **Prof. BIBHUTI BHUSAN DASH.**

### ADARSHA ROUTRAY

Date:

**Acknowledgement**

I am profoundly grateful to **Prof. BIBHUTI BHUSAN DASH,** my project mentor for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion.

I am also thankful to my team member for guiding and helping me all the way during the working period.

ADSRSHA ROUTRAY (2270018)

**ABSTRACT**

The "Poly-mod Virtual Companion" project aims to create a sophisticated personal assistant for Windows systems, drawing inspiration from established virtual assistants like Cortana and Siri. Users interact with the assistant through voice commands or keyboard input, enabling seamless execution of various tasks. The assistant assists with day-to-day activities such as conversation, web searches, video playback, music streaming, weather updates, word definitions, medication details, and reminders. Its Poly mod capabilities allow it to process voice input and provide accurate responses. Leveraging internet connectivity, it retrieves relevant information and presents it to users. The project implements a voice- concentric input/output system for enhanced user experience. Poly mod Virtual Companion aims to boost user productivity and intelligence through instant and accurate assistance. It is designed as a comprehensive tool for modern digital interactions, catering specifically to Windows system users.

**TABLE OF CONTENTS pg.no**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1. **INTRODUCTION 1-9**    1. Introduction-----------------------------------------1    2. Objective of the project--------------------------- 3-4    3. Motivation ------------------------------------------ 5    4. Scope Of Project---------------------------------- 6-7    5. Purpose---------------------------------------------- 7    6. Applicability -------------------------------------- 8-9 2. **SURVEY OF TECHNOLOGY 9-32**    1. Sample Implementation ------------------------ 9-15    2. System Architecture----------------------------- 16-17    3. Interaction & Related Work -------------------- 17-24    4. Implementation Details --------------------------24-30 3. **PRELIMINARIES 30-33**     1. Existing System------------------------------------30-32    2. Purposed System-----------------------------------32-33 4. **LITRATURE SURVEY 34-36**    1. “ Voice Assistant Using Python”--------------- 34    2. “Desktop Voice Assistant “--------------------34-35    3. “Smart Pyton Through Voice Recognition”-----36 |  |  |

|  |  |
| --- | --- |
|  | 4.4 “Survey On Virtual Assistant”---------------------36   1. **SYSTEM SPECIFICATION 37-38**    1. Hardware Requirements----------------------------37    2. Software Requirements-----------------------------38 2. **SYSTEM DESIGN 38-45**   6.1Architecture Design------------------------------- 38   * 1. UML Design----------------------------------------39   2. Activity Diagram-----------------------------------40   3. Sequence Diagram---------------------------------41   4. Deployment Diagram-----------------------------42-44   5. Data Dictionary-------------------------------------45  1. **SYSTEM STUDY 46-48**    1. Feasibility Study---------------------------------------46    2. Economically Feasibility------------------------------47    3. Social Feasibility---------------------------------------48 2. **SYSTEM TESTING 48-51**    1. Testing --------------------------------------------------48-49    2. Unit Testing---------------------------------------------49-50 |

|  |  |
| --- | --- |
|  | 8.4System Testing --------------------------------------51   * 1. White and Black-box Testing---------------------52   2. Test Results------------------------------------------52  1. **CONCLUSION 52-53** 2. **REFERENCES 53-54** |

* 1. **INTRODUCTION**

In today’s era almost all tasks are digitized. We have Smartphone in hands and it is nothing less than having world at your fingertips. 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 JIA. 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 as JIA. 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. JIA is effortless to use. Call the wake word ‘JIA’ 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 2025.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.Additionally ,the integration of artificial intelligence and machine learning algorithm enables virtual assistants to continuously improve their performance and adopt to user preferences overtime .By leveraging the wealth of information available on the internet,Virtual Assistants can provide valuable insights and personalized assistance to user across various tasks and applications.

### OBJECTIVE OF THE PROJECT

The development of technology allows introducing more advanced solutions in everyday life. This makes work less exhausting for employees, and also increases the work safety. As the technology is developing day by day people are becoming more dependent on it, one of the mostly used platform is computer. We all want to make the use of these computers more comfortable, traditional way to give a command to the computer is through keyboard but a more convenient way is to input the command through voice. Giving input through voice is not only beneficial for the normal people but also for those who are visually impaired who are not able to give the input by using a keyboard. For this purpose, there is a need of a virtual assistant which can not only take command through voice but also execute the desired instructions and give output either in the form of voice or any other means.

A Poly-mod Virtual Assistant is the software that can perform task and provide different services to the individual as per the individual’s dictated commands. This is done through a synchronous process involving recognition of speech patterns and then, responding via synthetic speech. Through these assistants a user can automate tasks ranging from but not limited to mailing, tasks management and

media playback. It understands natural language voice commands and complete the tasks for the user. It is typically a cloud-based program that requires internet connected devices and/or applications to work. The technologies that power virtual assistants are machine learning, natural language processing and speech recognition platforms. It uses sophisticated algorithms to learn from data input and become better at predicting the end user's needs.

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.

Polymod Virtual Assistants can tremendously save you time. We spend hours in online research and then making the report in our terms of understanding. JIA can do that for you. Provide a topic for research and continue with your tasks while JIA does the research. Another difficult task is to remember test dates, birth date or anniversaries. It comes with a surprise when you enter the class and realize it is class test today..

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.

# MOTIVATION

The main purpose of this project is to build a program that will be able to service to humans like a personal assistant. This is an interesting concept and many people around the globe are working it. Today, time and security are the two main things to which people are more sensitive, no one has the time to spoil; nobody would like their security breach, and this project is mainly for those kinds of people.

This system is designed to be used efficiently on desktops. Virtual Assistants software improves user productivity by managing routine tasks of the user and by providing information from an online source to the user. This project was started on the premise that there is a 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.

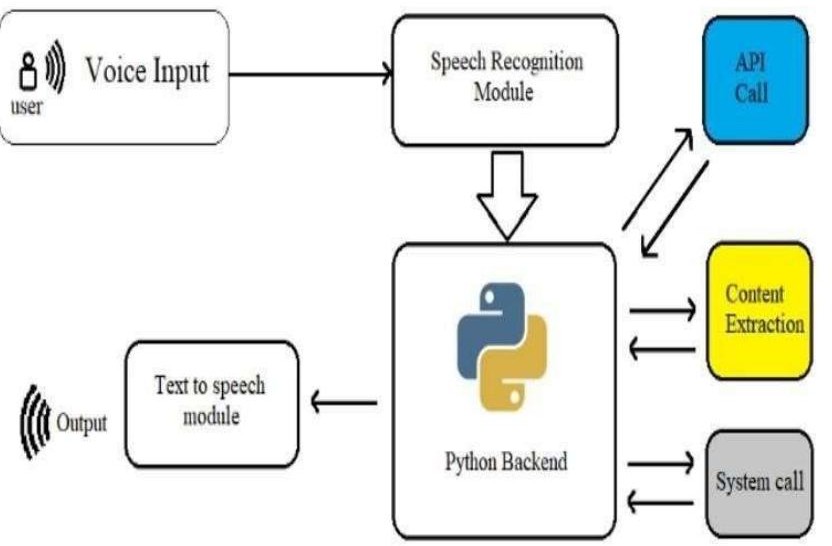
## 1.4 SCOPE OF THE PROJECT

Virtual 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 virtual assistants is missing. Users simply cannot see or touch a voice interface. Virtual Assistants are software programs that help you ease your day- to-day tasks, such as showing weather report, playing music etc. They can take commands via text (online chat bots) or by voice.

The main purpose of this project is to build a program that will be able to service to humans like a personal assistant. This is an interesting concept and many people around the globe are working it. Today, time and security are the two main things to which people are more sensitive, no one has the time to spoil; nobody would like their security breach, and this project is mainly for those kinds of people.

This system is designed to be used efficiently on desktops. Virtual Assistants software improves user productivity by managing routine tasks of the user and by providing information from an online source

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.



VIRTUAL ASSISTANT

### PURPOSE :

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

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

# 1.5 APPLICABILITY



The mass adoption of artificial intelligence in users' everyday lives is also fueling the shift towards voice. Chamber of Io T 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 Io T (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.

# 1.6 PURPOSE

Purpose of virtual assistant is to being capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audio books, and providing weather, traffic, sports, and other real-time information, such as news. Virtual assistants enable users to speak natural language voice commands in order to operate the device and its apps. There is an increased overall awareness and a higher level of comfort demonstrated specifically by millennial consumers. In this ever-evolving digital world where speed, efficiency, and convenience are constantly being optimized, it’s clear that we are moving towards less screen interaction.

## 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 JIA, libraries used are speech recognition to recognize voice, Pyttsx for text to speech, selenium for web automation etc.

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

### DBPedia:

Knowledge bases are playing an increasingly important role in enhancing the intelligence of Web and enterprise search and in supporting information integration. The DBpedin leverages this gigantic source of knowledge by extracting structured information from

Wikipedia and by making this information accessible on the Web. The DBpedia knowledge base has several advantages over existing knowledge bases: it covers many domains; it represents real community agreement it automatically evolves as Wikipedia changes, and it is truly multilingual.

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

**QUEPY :**

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

**PYTTSX:**

Pyttsx stands for Python Text to Speech. It is a cross-platform Python wrapper for 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 API s, online and offline. It supports API s like Google Cloud Speech API, IBM Speech to Text, Microsoft bing Voice Recognition etc.

**SQLITE :**

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

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

### SYSTEM IMPLEMENTATION

**MODULES**

* Pyttsx3
* Sapi5
* Speech recognition
* Pyaudio
* Wikipedia
* Webbrowser
* Cv2
* Stmp lib
* Image
* Numpy
* Os

**Pyttsx3( Python Text to Speech)**

Pyttsx 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.

**Sapis (Speech Application Programming Interface)**

The Speech Application Programming Interface or SAPI is an API developed by Microsoft to allow the use of speech recognition and

speech synthesis within Windows applications. To date, a number of versions of the API have been released, which have shipped either as part of a Speech SDK, or as part of the Windows OS itself.

Applications that use SAPI include Microsoft Office, Microsoft Agent and Microsoft Speech Server. Many versions (although not all) of the speech recognition and synthesis engines are also freely re distributive. SAPI 5 however was completely new interface, released in 2000, Since, then several sub-versions of this API have been released.

**Speech recognition**

Speech Recognition is a technology that enables the conversion of spoken language into written text, powered by AI and machine learning. It has a wide range of applications, including aerospace, automatic subtitling, emotion recognition, translation, hands-free computing, home automation, and education. In practice, it is used in products like Windows Speech Recognition and online tools such as Speech notes and Speech logger. With ongoing advancements, speech recognition technology continues to evolve and improve, demonstrating its value as a versatile and impact tool in modern technology.

**PyAudio**

Py Audio is a Python library that provides bindings for Port Audio, a cross-platform audio I/O library. It allows for easy audio playback and

recording in Python on various platforms. Audio's functionality includes checking if a specified device configuration is supported, obtaining the size for a specified sample format, and more. Users can utilize Py Audio for tasks such as recording sound from the microphone, playing and recording sound files in different formats, and working with Numpy and Python arrays containing sound. The library's capabilities make it a valuable tool for audio-related tasks in Python, offering a range of functionality with straightforward methods for implementation.

**Webbrowser and Cv2**

In My project, webbrowser and cv2 (Open-CV) play crucial roles in enabling diverse functionalities. The webbrowser module in Python provides an easy interface to open, search, and navigate web pages, making it an essential tool for implementing web-related features within your virtual assistant. Whether it's opening a specific website, conducting web searches, or automating web interactions, the webbrowser module simplifies web integration. On the other hand, cv2 (Open CV) is powerful . With Open CV, your assistant can perform tasks like face detection, object recognition, image manipulation, and video analysis, enhancing its visual perception capabilities. By leveraging the capabilities of webbrowser and cv2, your virtual assistant project gains the ability to seamlessly interact with web content and process visual data..

## System Architecture

The underlying multi modal assistant architecture supporting the

MVA app. The user interacts with a native iOS client. When the user taps the microphone icon, this initiates the flow of audio interleaved with gesture and context information streamed over a Web Socket connection to the platform.

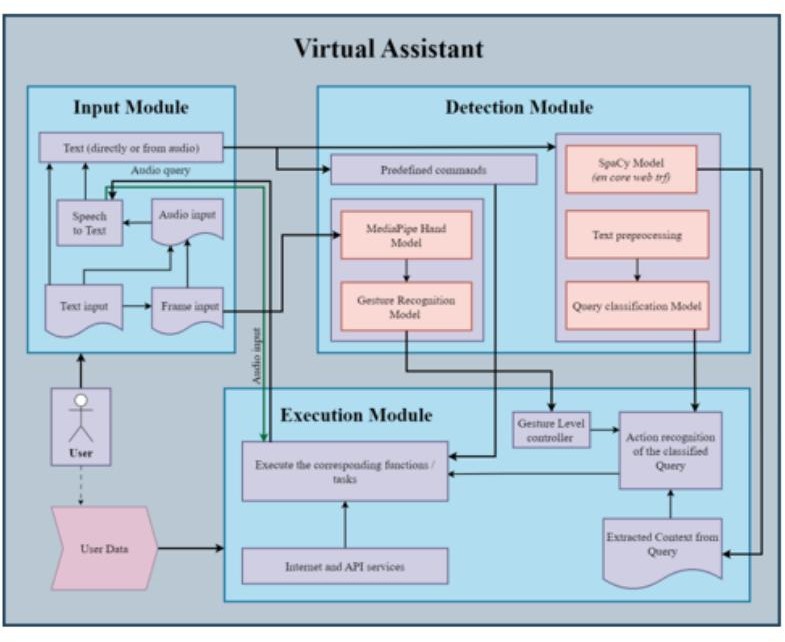
This stream of interleaved data is handled at the server by a multi modal natural language processing pipeline. This fields incoming packets of data from the client, demureness the incoming data stream, and sends audio, ink traces, and context information to three modules that operate in parallel. The audio is processed using the AT&T Watson speech recognition engine (Boffin Al., 2005). Recognition is performed using a dynamic hierarchical language model (Gilbert el., 2011) that combines a statistical N-gram language model with weighted sub- grammars. Ink traces are classified into gestures using a linear classifier. Speech recognition results serve as input to two NLU modules. A discriminate stochastic sequence tagger assigns tags to phrases within the input, and then the overall string with tags is as- signed by a statistical intent classifier to one of a number of intents handled by the system e.g. search(music event), refine(location).

The NLU results are passed along with gesture recognition results and the GUI and device context to a multi modal dialog manager. The contextual resolution component determines if the input is a query refinement or correction. In either case, it retrieves the previous command from a user con- text store and combines the new content with the context through destructive unification (Ehlen and Johnston, 2012). A location salience component then applies to handle cases where a location is not specified verbally. This component uses a su- per-vised classifier to select from among a series of candidate locations, including the gesture (if present), the current device location, or the current map .

## Sample Interaction

In Figure we present a sample of interaction from MVA that illustrates some of its capabilities. The user starts with a spoken natural language query where they specify some constraints: the type of music (jazz), location (San Francisco), and time (tomorrow). The system gets low confidence on the location, so it constructs a targeted clarifi- cation for clarifying only that constraint. The user repeats the location, and then the system searches for events meeting the user's constraints. The user then reviews the results, and follows on with a refinement: "What about blues?". Even though many parameters in this query are under specified, the system applies contextually-aware natural language understanding and interprets this as "Blues concerts near San Francisco tomorrow".

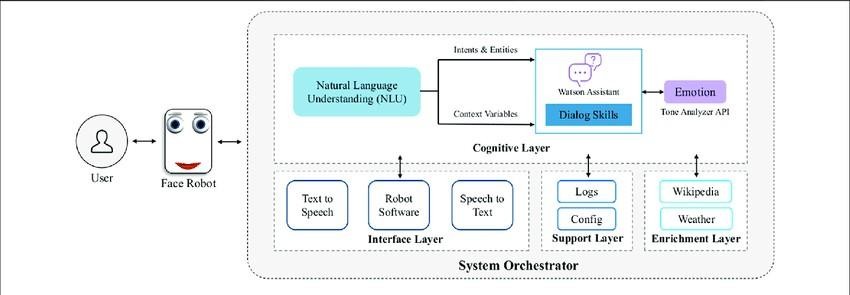
After selecting a concert, the user then searches for a restaurant nearby. The location of the concert re- mains salient. The user follows up with a mul- timodal query combining speech and gesture to search for similar restaurants in an adjoining area.



**Architecture 3.1**

Proposed architecture design of the intelligent virtual assistant for engagement in multi modal HRI. Multiple cloud-based AI services were combined and integrated with the robot's affective capabilities

in the back-end system by an orchestrate coded in Python. The orchestrate manages the flow of conversation and adapts the robot's facial representation of emotion in response to users' speech in real-time. The orchestrate coordinates the interface, cognitive, support and enrichment layers.



Architecture 3.2

**2.3RELATED WORK ON VOICE ONLY**

* + 1. **Modeling quality of experience**

Researchers have explored modeling the QoE in a multimedia system. One of the earliest works on modeling QoE defines it as a measure of the impact of content on a specific user in a specific context . QoE can be measured through subjective assessment or estimated through a model based on parameters of content, specific user, and specific context.

Nowadays there are three approaches for measuring QoE: subjective, performance based, physiological measures, and hybrid approaches . Each method enables the collection of specific type of information regarding the user's experience. For instance, subjective measures evaluate the user's satisfaction, fatigue, intuitiveness, preferences, etc., and they are collected typically via surveys. Performance measures evaluate the user's behavior when performing a specific task (such as task completion time, accuracy, error rate, etc.). Finally, the physiological measures evaluate non-voluntary responses of the human body during and immediately after the test session.

Parameters such as presence, satisfaction, and expectations are extremely difficult to measure with performance-based or physiological-based approaches . Therefore, this study adopts the

subjective approach to establish a conceptual framework of QoE for interactive multimedia environments al. proposed a framework that divided user experience into cognitive perception and behavioral consequences based on subjective measures. Utilizing the same framework, we define QoE as a composition of 8 dependent variables: presence, involvement, attention, reliability, dependency, easiness, satisfaction and expectations.

### From voice-only to immersive VPA

There is a growing international interest, both in academia and in industry, in developing VPA applications aimed at improving the user performance and the overall quality of user experience. Early implementations of VPA systems provided audio- only modality for interaction. Subsequent researches investigated the use of audio- visual VPA system. Lately, 3D and immersive displays are considered in an effort to provide higher level of immersion .While dictation technology has not advanced to the state of reliability as depicted in popular science fiction movies, it has been extremely useful in number of contexts . Most of the existing voice recognition systems like Siri and Google Now are mobile . The applications that use natural language user interfaces to answer questions, make recommendations, and perform actions by delegating requests to a set of web services .

They are claimed to be very smart, they adapt to the users individual

preferences over time in order to personalize results, and perform tasks such as finding recommendations for nearby restaurants or getting directions .

Visual interaction is one of the essential components to develop a VPA into something that can be compared to a human assistant. It is question worth asking if we should even explore the visual aspect in order to improve the contemporary voice assistants. Nonetheless, there are several other components of the contemporary V Pas that are known to be the limitation. Firstly, developing an optimal tone and pitch to deliver that desired impact of the voice is a hot research topic . Secondly, the accurate detection of input commands given in various accents limits the dependency and usage of a VPA system. No one likes to repent the commands multiple times in order to execute a desired outcome.

Advancements in 3D multimedia paved the way towards more realistic visual representation of the VPA using 3D displays. Examples of existing 3D displays include the Void Box , Mus ion Eyeliner], Peppers Ghost, 3D Fog Projection , inform and the 2.5D Shape Display . These 3D displays have raised questions whether such displays can be used to create immersive and realistic VPA systems. The ultimate goal is the digital recreation of real-world presence.

Entertainment, and social media search has demonstrated that enhancing virtual reality with digital characters will fundamentally allow us to interact with computers in a human way and on a personalized basis . For example, the authors in examined the roles that virtual humans can play in six areas: performance, physiology, leaming, connection, and security. Results showed that virtual humans enhance the quality of experience when multimodal interactions are incorporated. Jun et al. highlighted the significance of not only appearance and behavior but also nonverbal communication and affective components towards the quality of user experience. The relationship between rendering/display fidelity of virtual human models and emotional reactions is investigated in.

The evaluation of VPA systems is important to guarantee user satisfaction . Multimodal interfaces combine visual and auditory cues to enrich interactions but they raise new challenges concerning the usability and acceptability of such interfaces. As part of the European Project FASIL (Flexible and Adaptive Spoken Language and Multimodal Interfaces), a VPA system was developed to understand which factors affect the user experience and the acceptance of multimodal services . Results show that a conversational and multimodal approach was very well accepted and supported by the users. Furthermore, the quality and speed of the system feedback as

well as the recognition accuracy of the spoken components are key f actors to a better user experience. A recent study evaluated VPA to provide the elderly with a wide range of online services such as weather information and social networking [29]. Results demonstrated the need for several input/output modalities, distribution of modalities across different devices (PCs, Tablets, etc.), and adhering to international standards and avoiding closed solutions. A subsequent work evaluated the use of smart phones for VPA in immersive virtual reality

## Implementation details

* + 1. **Objectives and experimental variants**

The purpose of this research is to evaluate if adding a visual component to the voice- only virtual assistants would actually enhance the user experience. We hypothesized that providing a higher level of visual/auditory immersion would enhance the quality of user experience. In order to test this hypothesis, first stage involved the development of four different variants (sometimes referred to as treatments) of virtual assistant, each with a different audio/visual level of immersion. Developed VPA systems were the following; audio only, audio and 2D visual display, audio and 3D visual display, and audio immersive 3D visual display. The detailed explanation of each of the aforementioned variants is given below:

Variant 1: Voice-only display The first version of the prototype was an audio only version. We used an existing AI and voice recognition system as the

foundation for the functionality of the virtual assistant. The primary objective was to test the interaction modality, not the development of a better AI. The development of the first version was actually the development of a benchmark with which we compared our next versions that involved different output techniques. The audio only version acted similar to Siri or Google-Now where it took an audio input and gave an audio output. The user was able to perform a certain task, using Dragon Naturally Speaking.

Variant 2:

Audio and 2D visual display The second version of the prototype included adding a 2D visual display to the audio only virtual assistant. At this stage, we gave the virtual assistant some visual identity using the 2D avatar developed by Site Pal . The user was able to interact with a virtual assistant that had some kind of physical manifestation. The 2D display of the

virtual assistant was provided in the users laptop or computer screen. Variant 3:

Audio and 3D visual display The third version of the prototype brought about a paradigm shift from 2D visual display to a 3D visual display. At this stage we developed a visual representation of the virtual assistant and projected it using a 3D projection technique called peppers ghost . A snapshot of the 3D virtual assistant for variant 3 is shown in Fig..

Variant 4:

Audio and immersive 3D display The fourth version of the prototype further moved the virtual assistant from 3D (with peppers ghost) to an immersive 3D display with the Locus Rift.

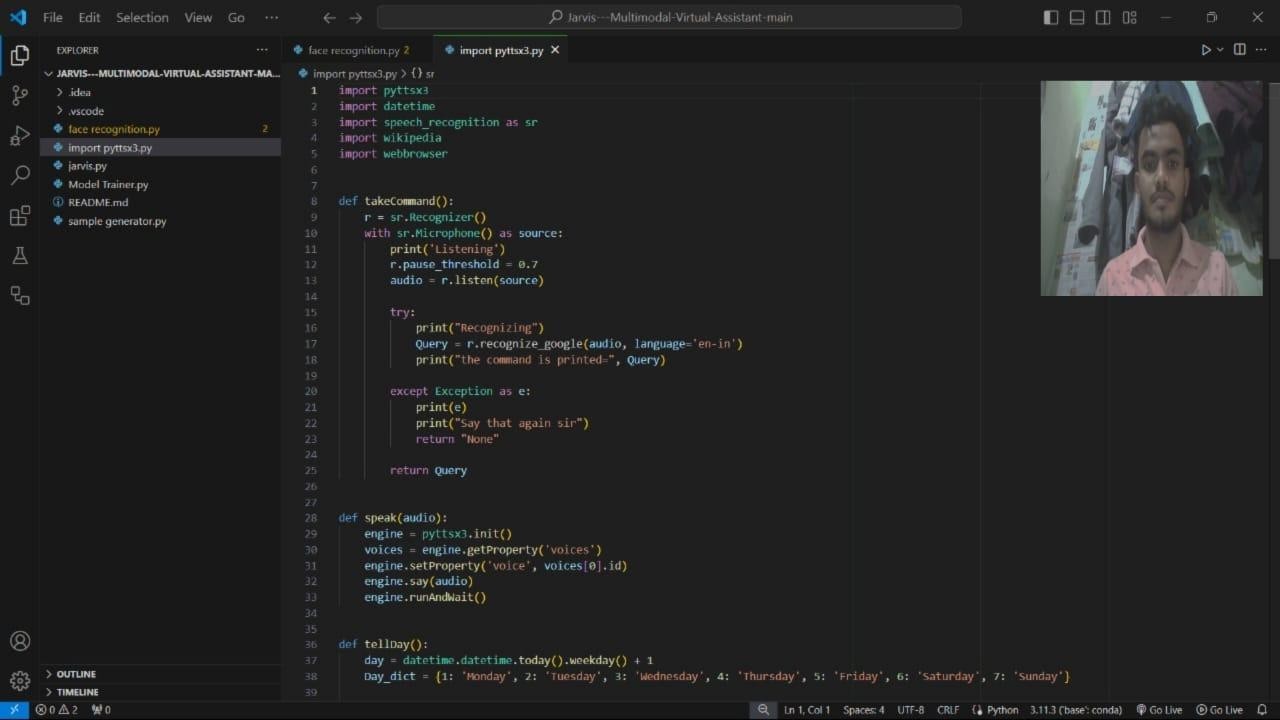
The virtual assistant was projected to the users eyes through the Locus Rift virtual reality glasses. Both the third and the fourth variants of the system

enabled us to dig deeper into understanding the questions posed in our hypothesis by using different mediums of displays for the virtual assistant.

* + 1. **Overview of system implementation**

Figure shows the system implementation architecture overview. The following subsections describe representative pieces of the system implementation. The architecture consists of three overalls components: the input systems (Email, Calendar, Dragon Naturally Speaking), the intermediate processing system (Processing, Web server, Web front-end, Site Pal API, Microsoft Windows, Locus HMD), and the output system (one of the four variants).

Fig2.4(I)



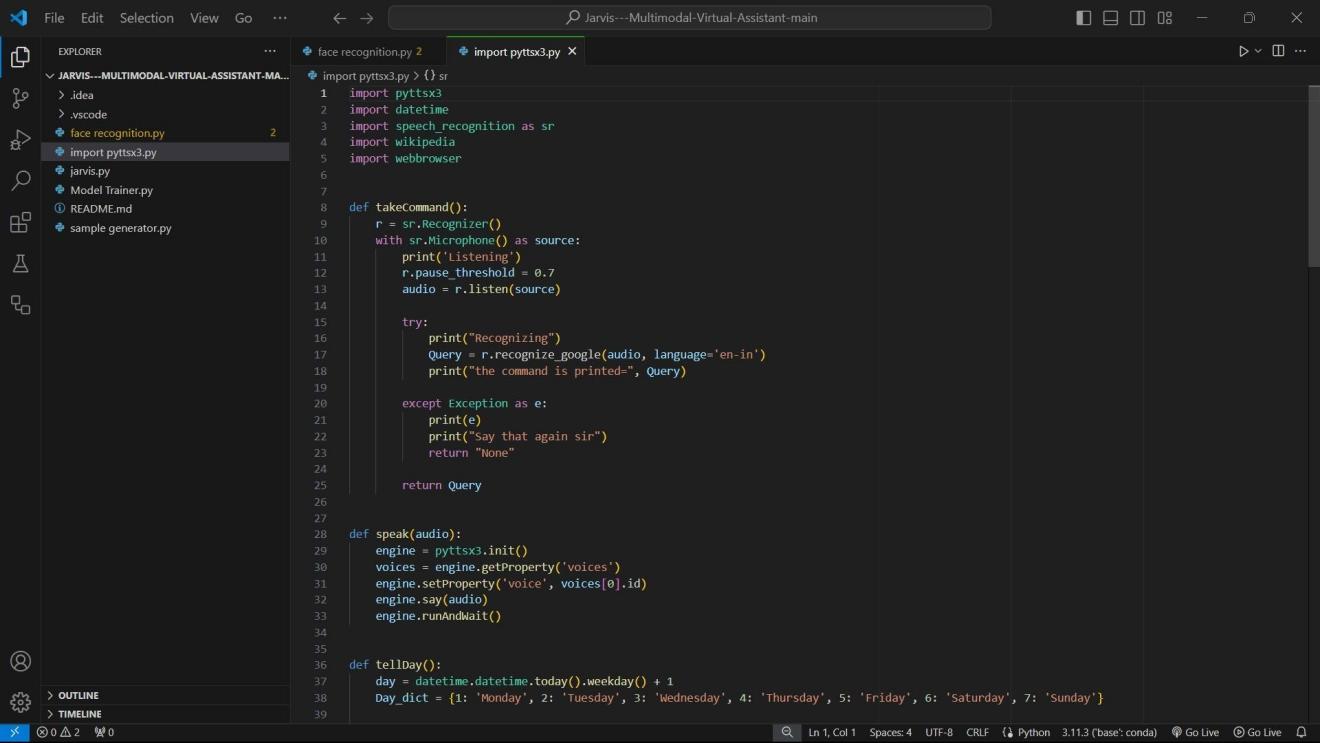


Fig 2.4.(ii)

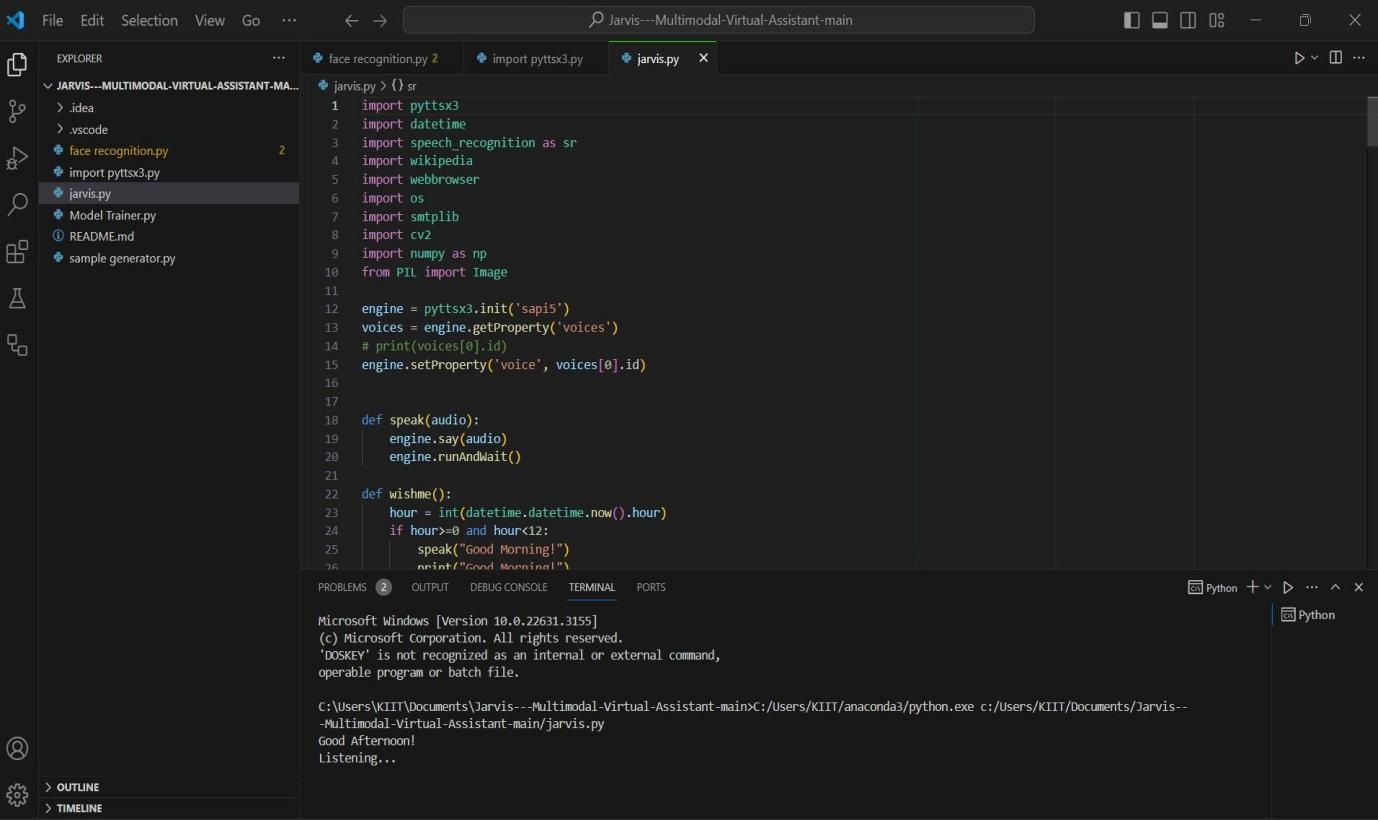


Fig 2.4(iii)

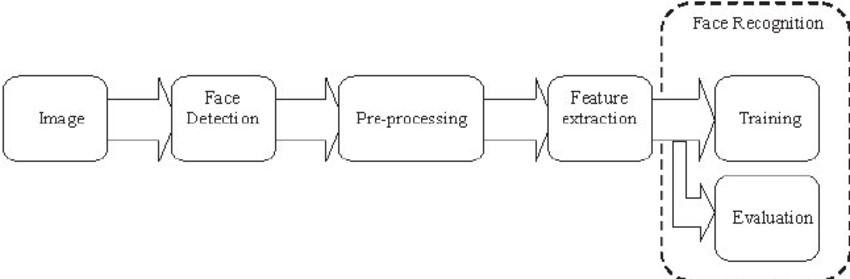


Fig 2.5 **Implementation Architecture**

* + 1. **Retrieving emails**

After establishing a connection, the user's inbox is checked for unread messages, and unread messages are stored in a file named email subjects.txt". The file is transferred from the local machine (that is running the code on Processing IDE) to the server where the rest of the VPA system is running on, using the FTP protocol. Now the updated version of the file email subjects. txt has been uploaded to the server, and is made available to the web server running the entire front-end of the VPA system, as well as part of the back-end. This entire process is continually repeated in order to keep updated the list of unread messages in a user's inbox.

**2.4.4 Reading emails**

After writing the newest set of unread emails to the file B email subjects.txt" and uploading it to the server, the front-end system must read the emails. In order to do this, the file Be email subjects.txt is read and the VPA responds as appropriate to the contents of this file. The function say Text() from the Site Pal API is utilized to instruct the virtual assistant to speaking out the text. Now the virtual assistant has speaks out loud the number of new emails if any. Next a loop continues to execute and accordingly reads out metadata, such as the subject and sender, of the user's new emails.

**2.4.5 Site Pal application programming interface (API) Implementation**

The Site Pal API is used to output the avatar of the front-end of the virtual assistant via a built- in text-to-speech engine . In synchronization with JavaScript, the Site Pal API drives the visual and audio output of the VPA to the browser window. A number of API function calls are used throughout the front- end system - such as say Text() and say Silent(), as demonstrated in the previous section.

**2.4.6 Google calendar authorization**

Similarly to reading a user's emails and notifying when a new email is received, it is important for the VPA system to have access to the user's calendar information. When there is an upcoming calendar event, the VPA system notifies the user. In order to obtain access to the user's calendar, it is necessary to obtain authorization from both Google and from the user.

**2.4.7 Querying google calendar**

Once the Google Calendar API is loaded and initialized, the VPA system is

able to query for information from the user's calendar. The query is called through request.execute() function that defines what to do with the response. This is defined in function(resp). In the case of calendar events, the response iterates through upcoming events and uses the say Text() function from the Site Pal API to speak out information about the user's events. For example, say Text; will read out the title, of the third calendar event in the response. Note that because the output of Site Pal text-to-speech engine is queued, the function to read upcoming events would simply add the information of each calendar event to the queue for the text-to-speech engine to dictate.

**2.4.8 Performance Evaluation**

The performance evaluation of the Poly-mod virtual assistant encompasses various key metrics to assess its effectiveness, efficiency, and overall user satisfaction. These metrics include conversational accuracy, task completion rate, response time, task completion time, user satisfaction, engagement, technical troubleshooting, and system metrics. Effectiveness is measured by the accuracy of conversational exchanges and the assistant's ability to complete tasks. Efficiency is evaluated through response and task completion times. User satisfaction is gauged through feedback and the assistant's ability to provide accurate information. Additionally, ongoing technical troubleshooting and system metrics contribute to the evaluation process. Tailoring the evaluation to align with specific implementation objectives and gathering continuous feedback from users are essential components of the evaluation process.

### ReQall

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

***Supported Tasks***

* Reminders
* Email
* Calendar, Google Calendar
* Outlook
* Facebook, LinkedIn
* News Feeds
* Ever note

*Drawback :*

One potential drawback of the Polymod virtual assistant is the potential

for a lack of personal interaction. As virtual assistants operate remotely, there is

no opportunity for face-to-face communication, which may impact the depth and

nuance of interactions. While virtual assistants offer numerous benefits, including flexibility and the ability to handle various tasks. This drawback underscores the importance of considering the specific communication needs and preferences of users when implementing virtual assistant systems.

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

## PRELIMINARIES

* 1. **EXISTING SYSTEM**

This project describes one of the most efficient ways for voice recognition. It overcomes many of the drawbacks in the existing solutions to make the Virtual Assistant more efficient. It uses natural language processing to carry out the specified tasks. It has various functionalities like network connection and managing activities by just voice commands. It reduces the utilization of input devices like keyboard. This project describes the method to implement a virtual assistant for desktop using the APIs. In this module, the voice commands are converted to text through Google Speech API. Text input is just stored in the database for further process. It is recognized and matched with the commands available in database. Once the command is found, its respective task is executed as voice, text or through user interface as output.

**3.1 PROPOSED SYSTEM**

1. QUERIES FROM THE WEB:

Making queries is an essential part of one’s life. We have addressed the essential part of a Citizen’s life by enabling our voice assistant to search the web. Virtual Assistant supports a plethora of search engine like Google displays the result by scraping the searched queries.

1. ACCESSING NEWS:

Being up-to-date in this modern world is very much important. In that way news plays a big crucial role in keeping ourselves updated. News keeps you informed and also helps in spreading knowledge.

3.TO SEARCH SOMETHING ON WIKIPEDIA:

Wikipedia's purpose is to benefit readers by acting as a widely accessible and free encyclopedia; a comprehensive written compendium that contains information on all branches of knowledge.

1. ACCESSING MUSIC PLAYLIST:

Music have remained as a main source of entertainment, one of the most prioritized tasks of virtual assistants. you can play any song of your

choice. However, you can also play a random song with the help of a random module. Every time you command to play music, the Virtual

Assistant will play any random song from the song directory.

1. OPENING CODE EDITOR:

Virtual Assistant is capable of opening your code editor or IDE with a single voice command.

## LITERATURE SURVEY

* 1. **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

* 1. **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

that will offer the services of a full-time employee usually reserved for the rich.

* 1. "VOICE ASSISTANT USING PYTHON" YEAR: 2020

AUTHORS: Subhash Mani Kaushal, Megha Mishra CONCEPT: Natural Language Processing.

* 1. "DESKTOP VOICE ASSISTANT" YEAR: 2020

AUTHORS: Gaurav Agarwal, Harsha Gupta, Chinmay Jain CONCEPT: Prerequisite APIs For Virtual Assistant.

* 1. "SMART PYTHON CODING THROUGH VOICE RECOGNITION YEAR: 2019

AUTHORS: M. A. Jawale, A. B. Pawar, D. N. Kyatanavar

CONCEPT: User experience field for better programming Integrated Development Environment Development (IDE).

* 1. "VPA: VIRTUAL PERSONAL ASSISTANT" YEAR: 2018

AUTHORS: Nikita Saibewar, Yash Shah, Monika Das

CONCEPT: Implementation of Functionalities of VPA. Q4.5 "SURVEY ON VIRTUAL ASSISTANT"

YEAR: 2018

AUTHORS: Amrita Sunil Tulshan, Sudhir Namdeorao CONCEPT: Functionalities of Existing IVAs.

4.6 "PERSONAL ASSISTANT WITII VOICE RECOGNITION INTELLIGENCE”

YEAR: 2017

AUTHORS: Dr. Kshama, V Kulhalli, Dr. Kotrappa Sirbi, Mr. Abhijit J. Patank

CONCEPT: Developing a Personal Assistant which has capability to work with without Internet Connectivity.

## 5.SYSTEM SPECIFICATIONS

* 1. HARDWARE REQUIREMENTS
* Processor Intel Pentium 4
* RAM 512 MB
* Hardware capacity: 80GB
* Monitor type-15inch colour monitor
* CD-Drive type-52xmax
* Mouse
* Microphone
* Personal Computer/Laptop
  1. SOFTWARE REQUIREMENTS
     1. Operating System - Windows
     2. Simulation Tools - Visual Studio Code
     3. Python- Version 3.9.6
     4. Packages 1.Pyttsx3

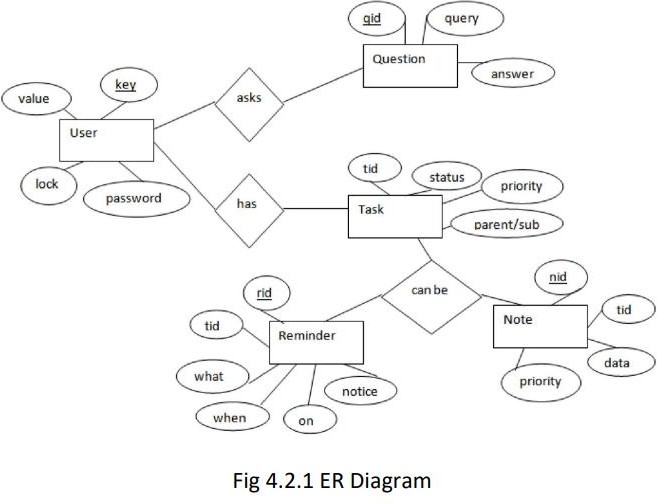
1. Speech Recognition
2. Wikipedia
3. Pyaudio
4. Webbrowser
5. Face\_Recognition
6. Smtplib
7. Cv2

9.Date time

10.OS

### 5.2 .5 System Design & Block Diagrams

**4.2.1 ER DIAGRAM**



The above diagram shows entities and their relationship for a virtual assistant system. We have a user of a system who can have their keys and values. It can be used to store any information about the user. Say, for key “name” value can be “Jim”. For some keys user might like to keep secure. There he can enable lock and set a password (voice clip). Single user can ask multiple questions. Each question will be given ID to get recognized along with the query and its corresponding answer. User can also be having n number of tasks. These should

have their own unique id and status i.e. their current state. A task should also have a priority value and its category whether it is a parent task or child task of an older task.

* + 1. **UML DIAGRAM**

The Unified Modeling Language is a general-purpose, developmental, modeling language in the field of software engineering that is intended to provide a standard way to visualize the design of a system. A UML diagram is a diagram based on the UML (Unified Modeling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system.

**UML defines several models for representing systems**

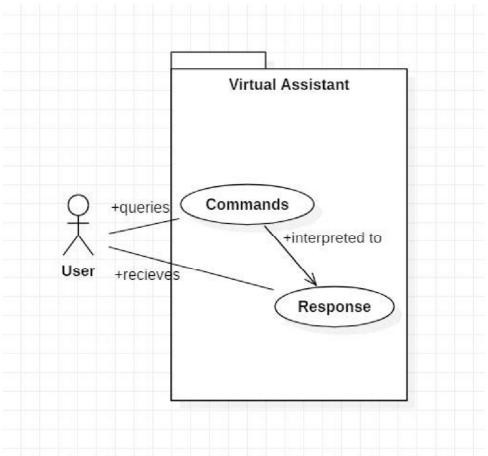
1. The class model captures the static structure
2. The state model expresses the dynamic behavior of objects
3. The use case model describes the requirements the requirements of the user
4. The interaction model represents the scenarios and messages flows
5. The implementation model shows the work unit

**ADVANTAGES:**

Most used and flexible Development time is reduced

Provides standard for software development

### USE CASE DIAGRAM

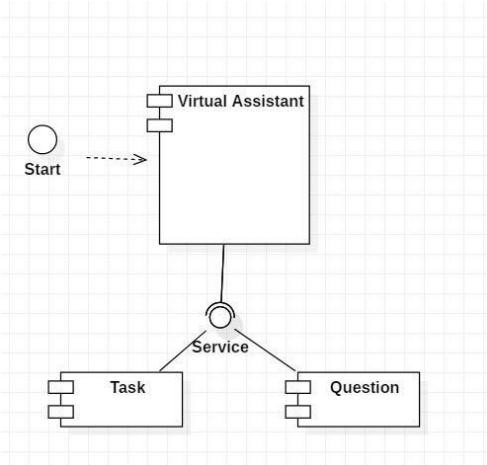


Use Case Diagram

In UML, use-case diagrams model the behavior of a system and help to capture the requirements of the system. Use-case diagrams describe the high-level

functions and scope of a system. These diagrams also identify the interactions between the system and its actors. In this project there is only one user. The user queries command to the system. System then interprets it and fetches answer. The response is sent back to the user..

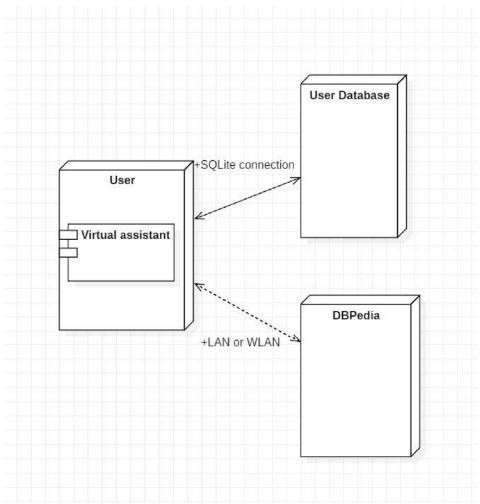
### COMPONENT DIAGRAM:



The main component here is the Virtual Assistant. It provides two specific service, executing Task or Answering Wei question.

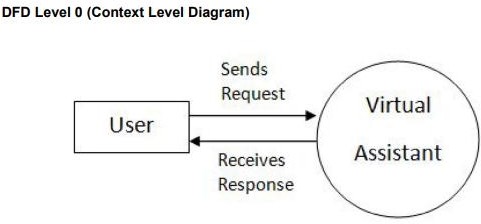
### DEPLOYMENT DIAGRAM

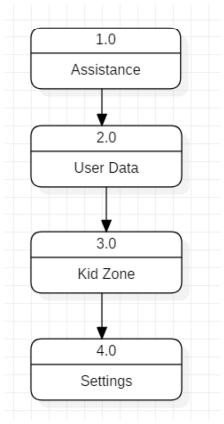
The main component here is the Virtual Assistant. It provides two specific service, executing Task or Answering Ourr question.



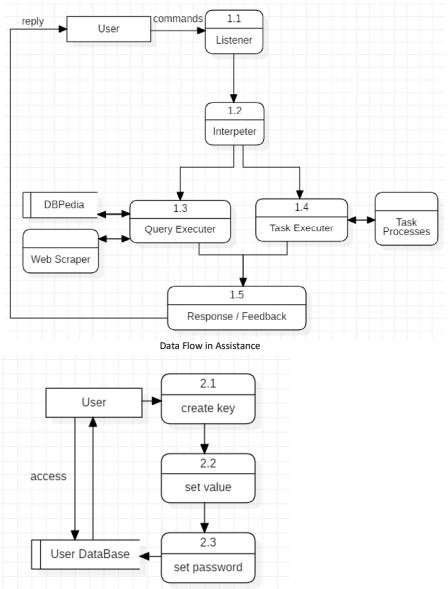
The user interacts with SQLite database using SQLite connection in Python code. The knowledge database DBPedia must be accessed via internet connection. This requires LAN or WLAN / Ethernet network.

### DATAFLOW DIAGRAM

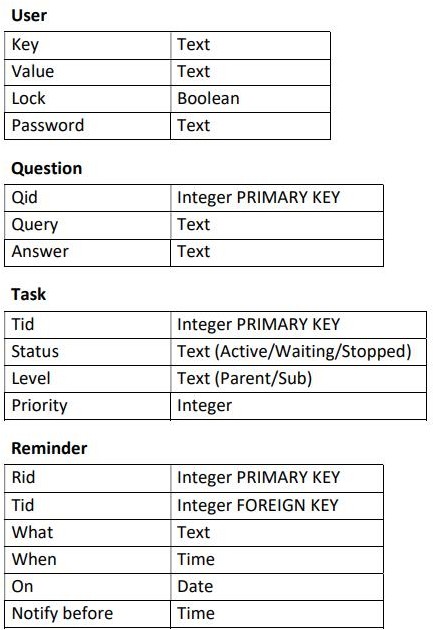




DFD LABEL- 1



### DATA DICTIONARY



* 1. **SYSTEM STUDY**
  2. **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 of the proposed system.

Three key considerations involved in the feasibility analysis are:

* Economical feasibility
* Technical feasibility
* Social feasibility
  1. **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 microphones and speakers. Again, they are cheap and available

* 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 what system speaks. These are very cheap now a days and everyone generally possess them. Besides, system needs internet connection. It is also not an issue in this era where almost every home or office has Wi-Fi.

**7.2 SOCIAL FEASIBILITY**

The social feasibility of the Polymod virtual assistant is evident in its alignment with the rapidly evolving landscape of remote work and digital communication. As remote work continues to gain prominence in the global workforce, virtual assistance has emerged as a vital component of this trend. The societal acceptance and demand for remote career categories, including virtual administration, indicate a growing recognition of the practical utility and relevance of virtual assistant services. This acceptance is further underscored by the availability of tailored virtual assistant services designed to assist small business owners, busy professionals, and individuals with day-to-day tasks, thereby demonstrating the adaptability and social relevance of virtual assistant technology. Moreover, the diverse range of tasks that virtual assistants can handle, such as social media management, lead generation, administrative support, and customer service, reflects the broad applicability of these systems to meet various social and business needs. Additionally, studies highlighting the feasibility and potential benefits of virtual assistants in addressing patient FAQs and streamlining administrative processes in the healthcare system showcase the broader social applicability and acceptance of virtual assistant technology in critical domains. In essence, the social feasibility of the Polymod virtual assistant is supported by its alignment with evolving work trends, adaptability to diverse tasks, and its potential to address societal needs in various domains, positioning it as a relevant and viable solution in the virtual assistant landscape.

### SYSTEM TESTING

* 1. **TESTING**

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

* 1. **TYPES OF TESTING UNIT TESTING**
     1. **UNIT TESTING**

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. It is the testing of individual software units of the application. It is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process.When considering what to add in the unit testing of a project, it's important to focus on several key aspects to ensure comprehensive coverage and reliability. Firstly, it's crucial to write unit tests that verify specific units of code in isolation, rather than testing integration, as this approach can help in identifying and addressing issues at a granular level, contributing to the overall robustness of the codebase. Additionally, emphasizing code completeness through unit tests can demonstrate the readiness of the code for integration and further testing, potentially saving development time by reducing the occurrence of bugs in subsequent testing phases. Furthermore, reducing coupling .

* + 1. **INTEGRATION TESTING**

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

* + 1. **FUNCTIONAL TESTING**

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical 29 requirements, system documentation, and user manuals. Functional testing is centered on the following items:

-- Valid input: identified classes of valid input must be accepted identified classes of valid input must be accepted.

-- Invalid output: identified classes of valid input must be accepted. Functions: Identified functions must be exercised Output: identified classes of application outputs must be exercised.

* + 1. **SYSTEM TESTING**

System testing ensures that the entire integrated software system. It tests a configuration to ensure known and predictable results. System testing is based on the process descriptions, flows, emphasizing pre-driven process links and integration points.

* + 1. **WHITE BOX TESTING**

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the 30 software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

* + 1. **BLACK BOX TESTING**

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document.

* 1. **TEST RESULTS**

All the test cases mentioned above have passed successfully. No defects encountered.

### CONCLUSION

**9.1. CONCLUSION**

Through this virtual assistant, we have automated various services using a single line command. It eases most of the tasks of the user like searching the web,

playing music and doing Wikipedia searches. We aim to make this project a complete.server assistant and make it smart enough to act as a replacement for a general server administration. The project is built using available open-source software modules with visual studio code community backing which can accommodate any updates in future. The modular nature of this project makes it more flexible and easier to add additional features without disturbing current system functionalities. It not only works on human commands but also give responses to the user based on the query being asked or the words spoken by the user such as opening tasks and operations. The application should also eliminate any kind of unnecessary manual work required in the user life of performing every task.

**9.2 FUTURE WORK**

The virtual assistants which are currently available are fast and responsive but we still have to go a long way. The understanding and reliability of the current systems need to be improved a lot. The assistants available nowadays are still not reliable in critical scenarios. The future plans include integrating our virtual assistant with mobile using React Native to provide a synchronized experience between the two connected devices.

REFERENCES

1. Abhay Dekate, Chaitanya Kulkarni, Rohan Killedar, "Study of Voice Controlled Personal Assistant Device", International Journal of Computer Trends and Technology (IJCTT) - Volume 42 Number 1- December 2016.
2. Deny Nancy, Sumithra Praveen, Anushria Sai, M.Ganga, R.S.Abisree, "Voice Assistant Application for a college Website", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 22773878, Volume-7,April 2019.
3. Deepak Shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhange, "Al Based Voice Assistant Using Python", Journal of Emerging Technologies and Innovative Research (JETIR), February 2019, Volume 6.
4. Dr.Kshama V.Kulhalli, Dr.Kotrappa Sirbi, Mr.Abhijit J. Patankar, "Personal Assistant with Voice Recognition Intelligence", International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017).
5. Isha S. Dubey, Jyotsna S. Verma, Ms. Arundhati Mehendale, "An Assistive System for Visually Impaired using Raspberry Pi", International Journal of Engineering Research & Technology (IJERT), Volume 8, May-2019.
6. Kishore Kumar R, Ms. J. Jayalakshmi, Karthik Prasanna, "A Python based Virtual Assistant using Raspberry Pi for Home Automation", International Journal of Electronics and Communication Engineering (IJECE), Volume 5, July 2018.

**Websites referred** [www.stackoverflow.com](http://www.stackoverflow.com/)

[www.pythonprogramming.net](http://www.pythonprogramming.net/)

[www.codecademy.com](http://www.codecademy.com/)

[www.tutorialspoint.com](http://www.tutorialspoint.com/)

[www.google.co.in](http://www.google.co.in/)

**Books referred**

Python Programming - WSCube Learning Python - Dr R. N Rao