**AI Based Smart Assistant**

Dr. Suvarna Joshi Vaidya

Computer Engineering, MIT Art, Design

and Technology University, Pune

(412201), Maharashtra, India,

(+91)9011045088

[suvarna.joshi@mituniversity.edu.in](mailto:suvarna.joshi@mituniversity.edu.in)

Chaitanya Yadwad

Computer Engineering, MIT Art, Design and Technology University, Pune (412201), Maharashtra, India,

(+91) 8999246722

[chaitanyayadwad8@gmail.com](mailto:chaitanyayadwad8@gmail.com)

Vince Dsouza

Computer Engineering, MIT Art, Design

and Technology University, Pune

(412201), Maharashtra, India,

(+91)9767851524

### [vincedsouza02@gmail.com](mailto:vincedsouza02@gmail.com)

Tanuj Sinha

Computer Engineering, MIT Art, Design

and Technology University, Pune

(412201), Maharashtra, India,

(+91)8484040035

[tanujsinha2002@gmail.com](mailto:tanujsinha2002@gmail.com)

Aaditya Nair

Computer Engineering, MIT Art, Design

and Technology University, Pune

(412201), Maharashtra, India,

(+91)7757984061

### aadityanair1803@gmail.com

`

**ABSTRACT**

In the era of smart technology, voice assistants have become versatile tools, performing tasks such as answering questions and playing music. This project's second phase aims to elevate voice assistant capabilities by integrating an AI-powered smart receptionist, improving the user interface for enhanced responsiveness, and seamlessly implementing IoT hardware to boost connectivity and data exchange. Recognizing the challenge of transitioning users to a lifestyle with a computer assistant, the project emphasizes a balanced approach to provide assistance without overwhelming users. The goal is to seamlessly integrate this technology into users' lives, harnessing the potential of voice input and output. The project's advancements position it at the forefront of ushering in a new era where AI-driven assistants not only cater to personal needs but also find applications in diverse business scenarios. As voice assistants continue to evolve, the utility of this project is expected to grow, making a significant contribution to the ongoing transformation of human-computer interaction and the integration of smart technologies in daily life.

# INTRODUCTION:

Voice assistants have undergone a transformative evolution, evolving into compact and versatile tools with the ability to execute a diverse array of tasks upon recognition of specific trigger words or commands. Distinguished from their human-operated counterparts, these digital assistants operate on technology-based principles, presenting immense potential as they advance in both personal and business domains. This project is dedicated to the realm of speech input, where verbal commands are translated into voice output, concurrently displayed as text on a screen. A significant challenge addressed by this project revolves around facilitating a seamless transition for individuals adapting to a lifestyle where a computerized assistant takes center stage in daily activities, eliminating the need for traditional input devices like keyboards and mice. The primary objective is to achieve this transition without overwhelming or intimidating users. Termed 'JARVIS,' the speech assistant in this project harnesses the capabilities of the Python programming language, leveraging a myriad of libraries to enhance functionality. While the intricacies of voice assistants might seem enigmatic to some, their operation parallels that of well-known counterparts such as Siri and Google Assistant.

This introduction delves into the key features of our Desktop-based AI voice assistant, building on the initial foundation. With the addition of new objectives, the project extends its capabilities:

- AI Smart Receptionist: The introduction of an AI-powered smart receptionist expands the voice assistant's repertoire, enabling advanced user interaction and a broader range of task execution.

- Improved UI Responsiveness: A concerted effort has been made to enhance the user interface, ensuring responsiveness and effectiveness. The UI now facilitates a more seamless and enjoyable interaction experience.

- IOT Hardware Integration: The project embraces the integration of Internet of Things (IoT) hardware, fostering improved connectivity and data exchange within the system.

This introduction aims to provide a comprehensive overview of the project's evolution, emphasizing its commitment to advancing user-friendly, technology-driven interactions through speech input and output.

# Literature Review:

1. **Personal Voice Assistant** –

**Mr. K. Vikram Reddy, Assistant Professor, Department of Computer Science and Engineering, Matrusri Engineering College, Saidabad, Hyderabad, Telangana,**

Voice assistants have been around for a while and have important innovations in waves. for voice assistant Dictation, voice commands, and searches have all become a typical component of wearables and smartphone devices. The study is the result of an oversight examination of the literature to provide a general understanding of voice (theory and concepts) control, virtual helpers, application areas, and more. When examining a number of the current automated systems that speak natural language processing power, numerous examples include found performing a number of functions in daily life. The Audrey was the first speech recognition system. was developed in 1952 by Bell Laboratories. Audrey Being relatively primitive and technologically constrained, recognizing only ten numbers when stated by certain persons (Pieraccini, 2012). (Pieraccini, 2012). A decade or so later, IBM created and presented its Shoebox. Machine. The gadget acknowledged and responded to 16 different spoken words, all ten digits including "0" to "9" and calculations commands like (IBM, 2018) Shoebox Machine Plus or Minus 16 spoken words were identified and replied to. Only the ten numbers "0" through "9" are included. by a designated speaker, in English. When limits become problematic, amplifying the voice of opposition's skepticism acknowledgment. The HMM significantly changed the creation of workable software for speech recognition. Through the aid of Initially, HMM speech recognition used a statistical method for calculating the likelihood of unknown words made of sounds.

1. **Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa –**

**Amrita S. Tulshan and Sudhir Namdeorao**

A method based on computers that executes a command on a subset of objects using a voice user interface. The subset is chosen from a given number of items, each of which has an object type and at least one taggable field with a corresponding value. The group of items is stored in the memory of the laptop. An utterance is a command, an object type selection, a taggable field selection, and a price for the taggable discipline that is obtained from the person. When the command is given, at least one item from the group of gadgets is retrieved in response to the user's selection of the sort, and the item's price in the taggable area selection matches the user's input on the taggable field charge. The object contains text content that will be converted to speech output. They envisioned computers one-day understanding natural language, considering what we need, when and where we need it, and actively taking full responsibility for us. However, speech recognition and machine learning continue to become more sophisticated, with the emergence of base records provided by packages and content providers. We agree as computing systems are proving to be smaller and more ubiquitous [e.g. B. wearables and the Internet of Things (IoT). Recognizers are designed to transform a person's verbal representation into another data method, such as text. A portable personal correa with a speech recognizer and a distinctive dialect processor is disclosed. This snippet of data can be your daily schedule, personal journal data, or personal address book data. B. Telephone number. The most well-known iPhone application is "SIRI". This allows the end customer to deliver a variety of voices to the end customer and also respond to the customer's voice charges. It's called a personal assistant with speech recognition intelligence that takes customer input regarding speech or content types, processes it, and returns it in various structures such as B. An activity performed or an element directed to an end customer. In addition, this proposed framework can change the communication method between end customers and mobile phones. Open data is currently garnering attention for its imaginative governance, especially in the areas of government, life science, and smart entrepreneurship. Either way, his Open Data web crawler helps him find out what kind of information is

out there to facilitate his application to Buyer Management.

**3. Artificial Intelligence Based a commutative Virtual Voice Assistant Using python and visual studio code technology** **–**

**Raj Kum Jain, Vikas Sharma, Mangilal, Rakesh Kardam, Mamta Rani**

There have been some significant advances and innovations in this area of virtual assistants (VA) with voice recognition. This is due to the huge demand for the device. This helps devices with voice recognition only, such as smartwatches or fitness bands, smart speakers, Bluetooth headphones, intelligent mobile phones, AI-based laptops or desktops, TV. A new set of AI technologies are continually being developed to improve his performance in automatic voice search. As the amount of knowledge in has grown exponentially and is now called big data, the easiest way to improve the results of the virtual assistant is to incorporate the assistant into machine learning and depending on usage train the device. Other emerging technologies of equal importance are artificial intelligence, the Internet of Things, big data access, etc. Use voice assistants to easily automate tasks and convert spoken audio into text form It performs all the tasks, from converting it to extracting keywords from that text and running queries to provide the results to the user. Machine learning programming is only a subset of artificial intelligence technology. This is one of the most important and useful advances in this technology. Before the advent of AI, we constantly upgraded our technology to accomplish tasks, but today, machines themselves can take on new tasks and solve them without human involvement in development. can do. Useful for today's daily life. From smart mobile phones to personal desktops to automation and engineering industries, these types of his assistants are in high demand to automate tasks and increase efficiency.

# OBJECTIVES:

Our voice assistant system, designed to function as a versatile tool encompassing note-taking, searching, music playback, and entertainment, reflects the evolving landscape of voice assistants permeating various aspects of our daily lives. This research aims to delve into the existing features, expanding upon them to provide an enriched user experience by addressing the following objectives:

1. Leverage Computing Power:

- Objective: Harness the significant increase in computing power to optimize the voice assistant's efficiency, responsiveness, and capability to handle a broad spectrum of tasks.

- Implementation (New): Integrate an AI-powered smart receptionist, leveraging enhanced computing power for advanced user interactions.

2. Enhance User Productivity and Convenience:

- Objective: Create a multifaceted voice assistant, serving as a noted writer, search tool, music player, and entertainer, thereby significantly improving user productivity and convenience.

- Implementation (New): Enhance the user interface to be more responsive and effective, ensuring a seamless interaction experience.

3. Utilize Linguistic Data:

- Objective: Effectively utilize the abundance of linguistic data to enhance the voice assistant's understanding and processing of natural language commands, thereby improving overall performance and user experience.

- Implementation (New): Implement IoT hardware for improved connectivity, leveraging linguistic data for efficient data exchange within the system.

4. Apply Machine Learning:

- Objective: Integrate highly successful machine learning methods to continually enhance the voice assistant's ability to recognize and interpret user commands, refining speech recognition and natural language understanding.

- Implementation (New): Implement machine learning techniques for continuous improvement, focusing on refining the AI-powered smart receptionist's capabilities.

5. Achieve Rich Language Understanding:

- Objective: Develop a voice assistant that not only understands language but also comprehends the nuances of human language and its usage in various social contexts, facilitating more meaningful and context-aware interactions.

- Implementation (New): Focus on IoT hardware integration to achieve richer language understanding, enabling the voice assistant to engage in contextually aware interactions

**FLOWCHART**

Other

Services

Waiting

Voice

Generated

Text

Voice Recognition

Speech Queue

Recognized

Sentence

Check Device State

Natural Language

Understanding

Speak ing?

[Yes]

Waiting

Intent &

Entities

[No]

Voice Synthetization

# METHODOLOGY:

The methodology for developing a voice assistant encompasses a structured approach to ensure the successful creation and implementation of the project. This methodology involves several key phases which include:

1. Python 3.6 or Higher: - Python is the programming language used for developing your voice assistant. It's essential to have Python 3.6 or a higher version.

2. pyttsx3 Package - Offline Text-to-Speech: - The project relies on the pyttsx3 package, a Python library for text-to-speech conversion. One of its key advantages is that it functions offline,

3. Python's Speech Recognition Library: - Another fundamental requirement is Python's Speech Recognition library. This library enables your voice assistant to understand and interpret spoken language, converting speech into text..

4. Installation of Python Libraries: -

- pyttsx3

- Speech Recognition

- Other dependencies and libraries necessary for specific functionalities or modules in your project.

5. Operating System Compatibility: - Ensure that your project is compatible with the operating system you're using.

1. Creating and designing if else statements that search for the command and intent given in each query submitted and providing a suitable response or action for all such recognized commands. This is implemented by storing the query in a string text format and checking if a particular sequence of words exists within it that can be used to identify a command.

For example - generally if a query starts with “What is”, it’s a query that entails searching the world wide web for answers so when “what is” is detected the query can be used as the input in a web search for satisfactory results.

1. Designing and incorporating a simple and user-friendly User Interface (UI) with several functions such as

- Display time

- Provide Subtitles

- Interact and guide the user with helpful tips and tutorials

The user starts the Assistant Application

The assistant greets the user via a speak() function based on the time by calculating the system time to the category of morning, afternoon, and evening

The assistant then asks the user for a voice command and suggests the command "open help guide" to see the list of all possible commands.

A "ding" sound will be played after which the Assistant will wait for a command.

Once the command is spoken into a microphone the system will utilize the Google speech recognition API to convert the audio input into a string format ( a sentence )

This sentence will be subjected to minor transformations like converting all alphabetical characters to lowercase and then sent as input to a large if-else block of code which will look for sequences of words with the sentence which will be called the Query from now

Accumulating all the software into a hardware case to have a tangible Assistant with various uses and functions which is the final goal of creating an AI Assistant needs to be carried out by 3D printing a durable housing for the internal components which will include the following-

-Computer (raspberry pi)

-Microphone

-Speakers

-Ethernet Port

-Power supply`

-Bluetooth Adapter

-Wi-Fi Adapter

The method of communication from the assistant to the user is Voice based on a speak() function which converts a string ( a sentence ) to its audio format ( in English ) and plays this audio through the audio output in use/

# FUTURE SCOPE:

Based on technological advancements, we can create android applications that meet the requirements of a large number of consumers. We could also go the route and integrate this technology as the new standard for home security and contact-less interaction with entities outside a smart home. A user would like to use ‘JARVIS’ in Modern Trends in Science and Technology Assistants that make life easier through inclusion Users are supported by the following features:

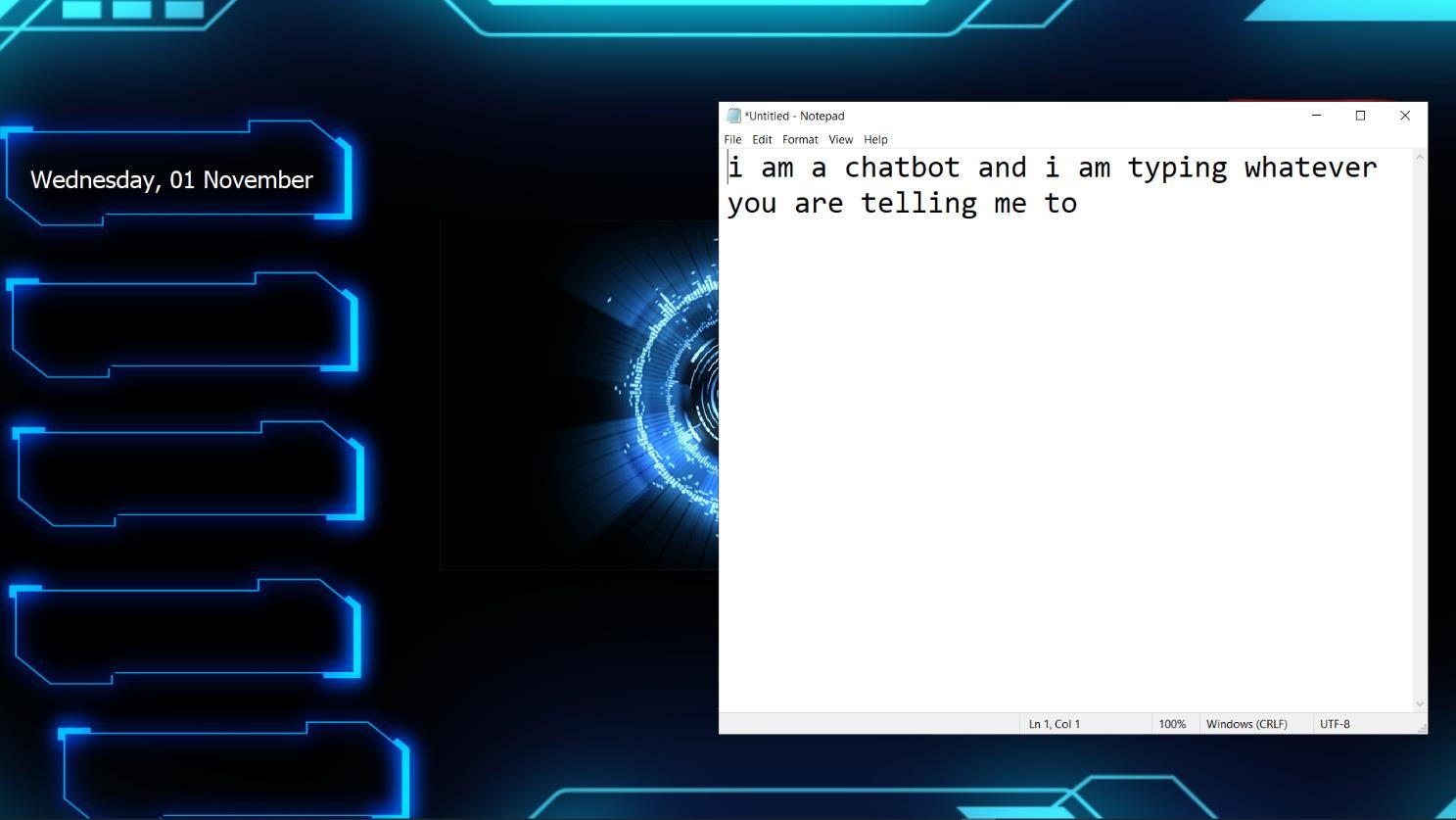
1. Make ‘JARVIS’ more independent and make it learn some new skills.
2. ‘JARVIS’ Android and IOS apps can also be created.
3. Build more ‘JARVIS’ voice terminals.
4. Voice commands are widely used to ensure safety Encrypted.
5. Jarvis could also be used in bots
6. It could also serve in assisting military operations and military mechanics.
7. It could be used as the main terminal for controlling a smart home
8. It could be used as a tool for education providing visual aid and information from the internet

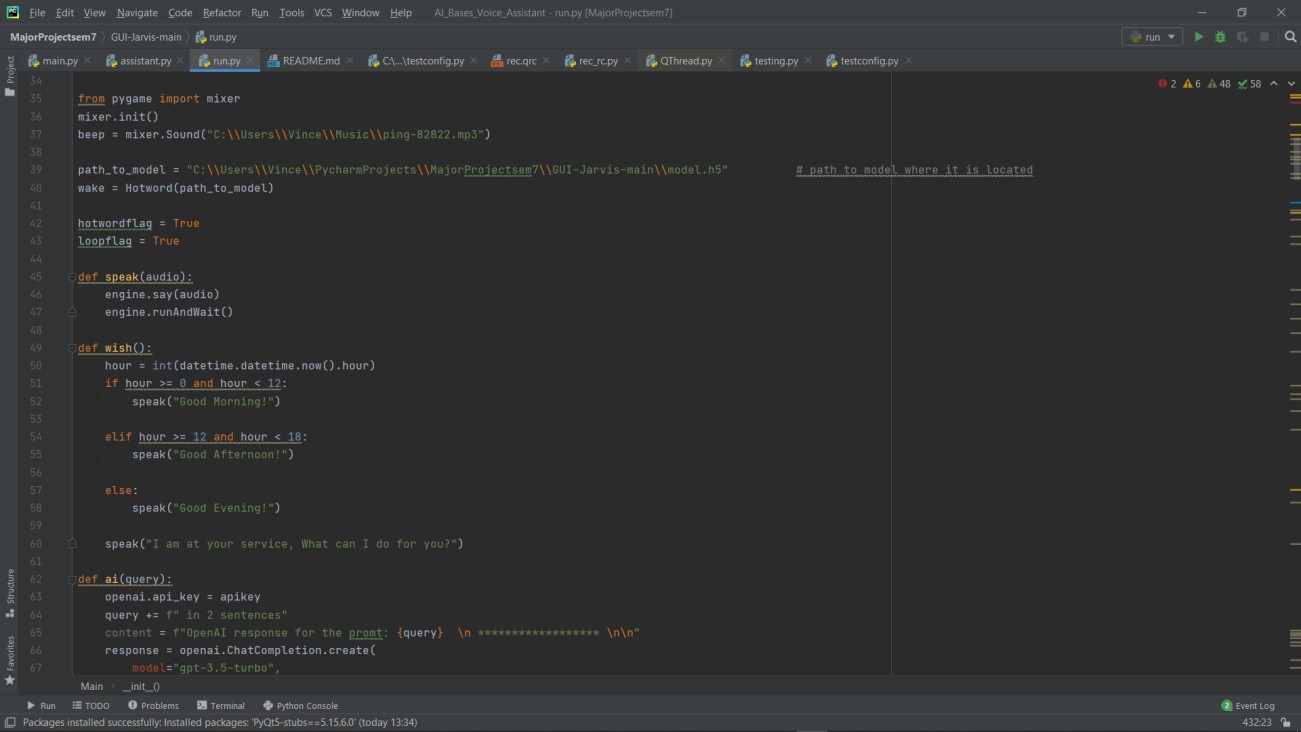
# 6. ACKNOWLEDGEMENT:

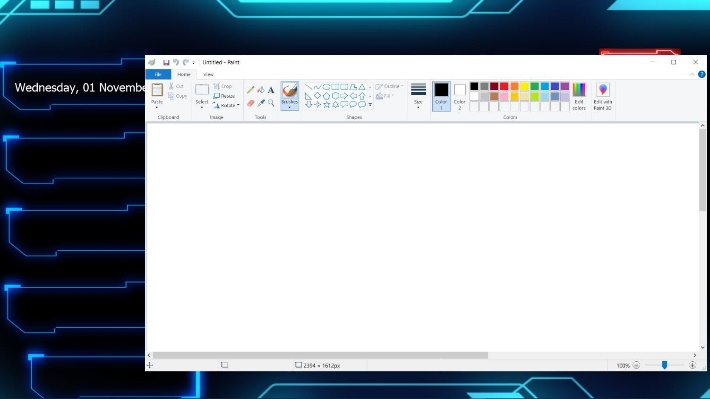
We would like to take this opportunity to thank all those guideposts who paved the way for us throughout the project leading to the successful and satisfactory completion of this study.

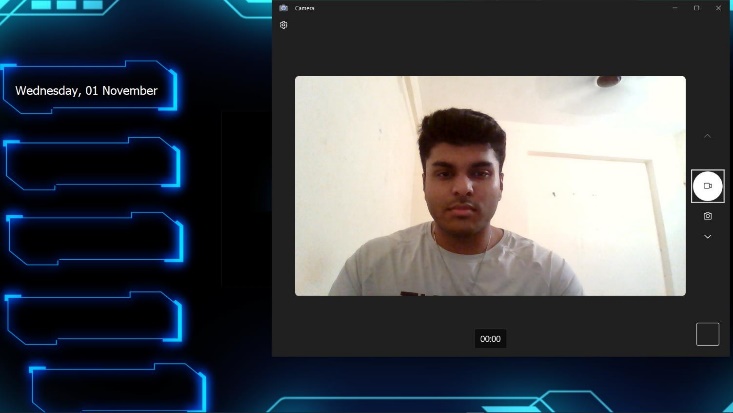
We are very grateful to Dr. Suvarna Joshi for giving us the opportunity to carry out this project and providing us with all the facilities. We are very grateful to Dr. Joshi for her active support, valuable time and advice, sincere cooperation and dedication.

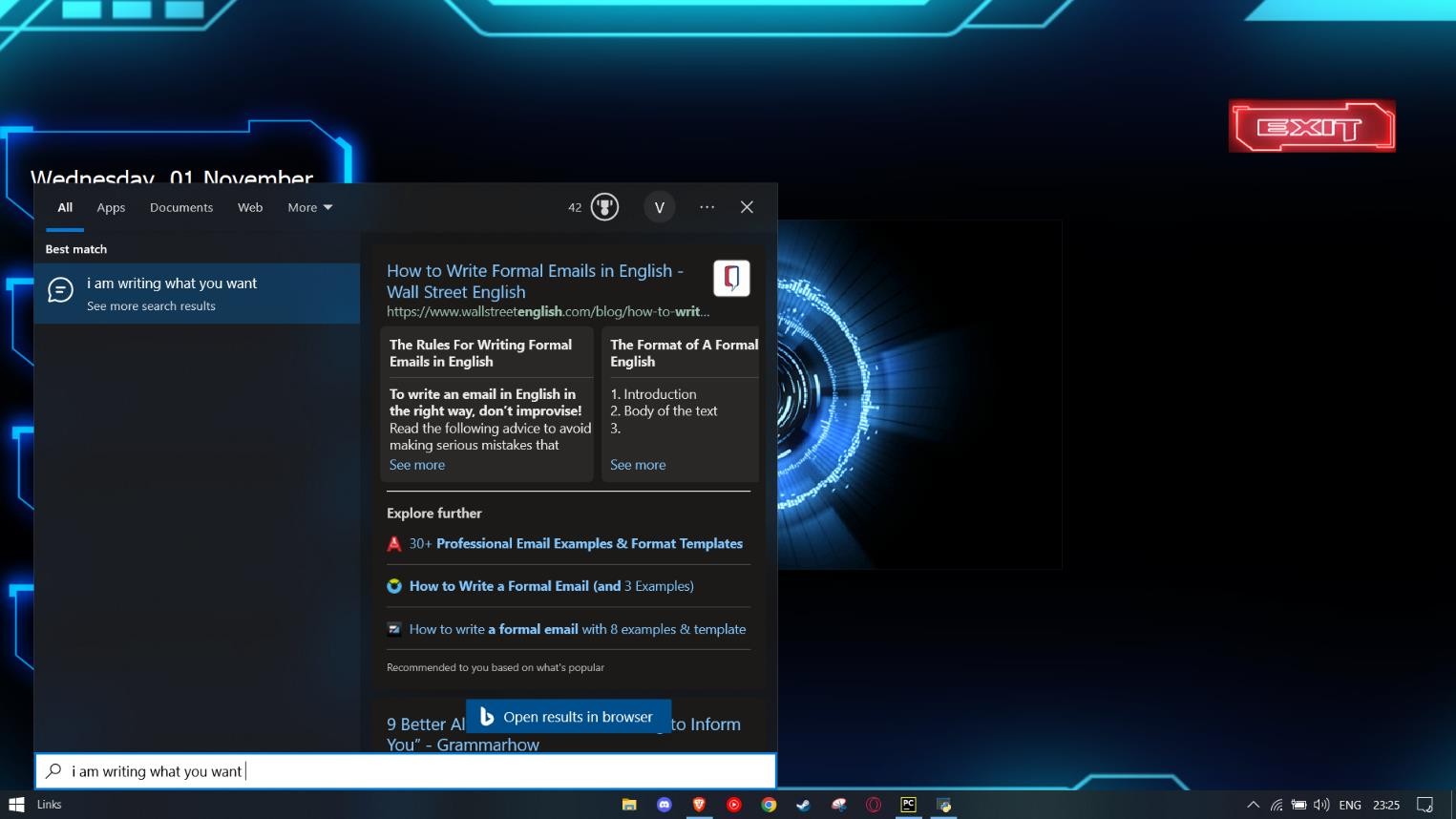
# 7. Results











**8. REFERENCES**

1. Voice Assistants and Smart Speakers in Everyday Life and in Education George TERZOPOULOS, Maya SATRATZEMI Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece Email: [gterzopoulos@uom.edu.gr,](mailto:gterzopoulos@uom.edu.gr) [maya@uom.edu.gr](mailto:maya@uom.edu.gr)
2. PERSONAL VOICE ASSISTANT Mr. K. Vikram Reddy, Assistant Professor, Department of Computer Science and Engineering, Matrusri Engineering College, Saidabad, Hyderabad, Telangana, S. Lahari, A. Naveen, G. Sarath Chandra, B.E Scholars, Department of Computer Science and Engineering, Matrusri Engineering College, Saidabad, Hyderabad, Telangana
3. Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa: 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers
4. Links
   * [**https://www.youtube.com/**](https://www.youtube.com/)
   * [**https://stackoverflow.com/**](https://stackoverflow.com/)
   * [**https://scholar.google.com/**](https://scholar.google.com/)
   * [**https://www.geeksforgeeks.org/**](https://www.geeksforgeeks.org/)

[5] Youtube Channels referred

* + Product Alliance
  + Soundhound

**9. Performance Review**

Jarvis can handle the following things, many more implementations can take place as technological advancements take place slowly and gradually

What can Jarvis do:

* Speech to text, Text to speech
* Taking down notes
* Controlling audio settings
* Weather/temperature details
* Can interact with user and provide human feel.
* It runs in the background and needs no

wake up call

* Smart home integration that incorporates

Voice control.