

# **SCREEN INTERPRETING ASSISTANT**

**A PROJECT REPORT**

**Submitted in partial fulfilment of requirements to**

**RGUKT - SRIKAKULAM**

**For the award of the  
degree**

**B. Tech in CSE**

**By**

**UDAY SANKAR GOTTIPALLI  
(S180443)**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING  
RGUKT SRIKAKULAM, ETCHERLA  
March 2024**



**Rajiv Gandhi University of Knowledge  
Technologies**

## **SCREEN INTERPRETING ASSISTANT**

Rajiv Gandhi University of Knowledge Technologies Srikakulam –  
532402

**Project done by:**

S180443 –

UDAY SANKAR GOTTIPALLI

**Under the Guidance of**

Mr. CH SATISH KUMAR

Asst. Professor, Dept of CSE

IIIT RGUKT, Srikakulam



### **DECLARATION**

I certify that

- a. The work contained in this report is original and has been done by us under the guidance of my supervisor(s).
- b. The work has not been submitted to any other Institute for any degree or diploma.
- c. We have followed the guidelines provided by the University in preparing the report.
- d. We have confirmed the norms and guidelines given in the Ethical Code of Conduct of the University.
- e. Whenever we have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the report and giving their details in the references.
- f. The work was done in the academic semesters period i.e., B.Tech 4<sup>th</sup> year.



## CERTIFICATE

Certified that this project work titled “**SCREEN INTERPRETETING ASSISTANT**” is the bona fide work of **Mr. Uday Sankar Gottipalli** bearing ID. No. **S180443** who carried out the work under my supervision and submitted in partial fulfilment of the requirements for the award of the degree, BACHELOR OF TECHNOLOGY, during the year 2023 – 2024.

Mr. Ch Satish Kumar,  
Project Guide,  
Department of CSE,  
RGUKT, SRIKAKULAM.

Mrs. Ch Lakshmi Bala ,  
Head of the Department,  
Department of CSE,  
RGUKT, SRIKAKULAM

## **Abstract:**

Websites have huge and valuable amounts of information in a transactional format. Developing websites which are accessible to all the users has become a priority. Thus, everyone can use the internet and experience their true potential. Despite visually elegant websites, accessible for visually impaired people seems complicated.

**Screen readers (SRs)** made much of this information available to visually impaired users. However, since their development, screen readers have offered only limited access to all information on the web page. Often these screen readers either be verbose or skim through information which is biased by accessibility of the website. The handling of SRs comes with numerous and complicated keyboard commands and shortcut keys.

**Voice assistants (VAs)** work comparatively better than screen readers in terms of interaction. They interact vocally using voice commands and perform tasks based on the keywords identified in the conversation. Existing systems of voice assistants have limited modular features as reading out the contents of the webpages.

Webpages differ from each other in different ways such as in structure, design, content, and HTML elements. This project is about making a voice assistant that can analyze any type of website and understand its contents. This project helps in obtaining the information from the webpages and perform required actions on webpage through user voice command. Through this assistant, users can interact with websites with only voice commands, which helps visually impaired users effectively.

**Keywords:** Screen Readers, Voice Assistant, Natural Language Processing, visually impaired web Users.



RAJIV GANDHI UNIVERSITY OF KNOWLEDGE TECHNOLOGIES  
RGUKT-SRIKAKULAM, Srikakulam Dist. - 532402

---

### ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people whose perpetual cooperation made it viable, whose constant guidance and encouragement crown all efforts with success.

We would like to express our profound gratitude and deep regards to our guide **Mr. Ch Satish kumar** for his exemplary guidance, inspiration, monitoring, and constant encouragement with constructive suggestions throughout the course of this project.

We are extremely grateful for the confidence bestowed in us and for entrusting our project entitled “**Live Lecture Visual Generation**”. At this juncture, we feel deeply honoured in expressing our sincere thanks to her for providing valuable insights leading to the successful completion of our project.

S180443 – UDAY SANKAR GOTTIPALLI

## List of Figures

Figure 1: Use Case Diagram.....	7
Figure 2: SIA State Diagram .....	8
Figure 3: Sequence Diagram.....	9
Figure 4: Flow Chart .....	11
Figure 5: output 1.....	12
Figure 6: output 2.....	12
Figure 7: output 3.....	13
Figure 8: output 4.....	13
Figure 9: Raspberry pi .....	14
Figure 10: Speaker.....	14
Figure 11: Microphone.....	14

## Table Of Contents

CHAPTER 1: INTRODUCTION .....	1
1.1. Purpose:.....	1
1.2. Scope: .....	1
CHAPTER 2: OVERALL DESCRIPTION .....	2
2.1. Product Perspective:.....	2
2.2. Product Features.....	2
2.3. User Classes and Characteristics.....	3
2.4. Operating Environment: .....	4
2.5. Technologies Used .....	4
2.6. Requirements.....	4
CHAPTER 3: FEASIBILITY STUDY .....	5
Introduction: .....	5
3.1. Technical Feasibility: .....	5
3.2. Economic Feasibility:.....	5
3.3. Operational Feasibility: .....	6
3.4. Legal and Ethical Considerations: .....	6
CHAPTER 4: DESIGN .....	7
4.1. Use Case Diagram: .....	7
4.2. SIA State Diagram:.....	8
4.3. Sequence Diagram .....	9
CHAPTER 5: IMPLEMENTATION.....	10
5.1. Algorithm: .....	10
5.2. Flow Chart.....	11
CHAPTER 6: OUTPUT .....	12
CHAPTER 7: HARDWARE .....	14
CHAPTER 8: NON-FUNCTIONAL REQUIREMENTS .....	15
8.1. Performance: .....	15
8.2. Usability: .....	15
8.3. Reliability: .....	15
CHAPTER 9: RESULT .....	16
9.1. Progress.....	16
9.2. Drawbacks:.....	16



CHAPTER 10: CONCLUSION.....	18
CHAPTER 11: FUTURE SCOPE.....	19
CHAPTER 12: REFERENCES.....	20



# CHAPTER 1: INTRODUCTION

## 1.1. Purpose:

This project aims to enhance web accessibility for visually impaired users by developing a Voice Assistant capable of comprehensively analysing and understanding the content of diverse websites. Unlike traditional screen readers, the Voice Assistant utilizes voice commands to facilitate effortless interaction, allowing visually impaired individuals to navigate and engage with web pages effectively. The goal is to overcome existing limitations in accessibility tools, providing a more inclusive and seamless internet experience for users with visual impairments.

## 1.2. Scope:

### 1. Visually Impaired Users:

Targeting those with visual impairments, the Voice Assistant ensures easy, voice-controlled interaction with websites.

### 2. Universal Web Accessibility:

Benefits extend beyond visual impairments, aiding users with motor disabilities and diverse accessibility needs.

### 3. E-commerce:

Applicable in e-commerce for an inclusive user experience.

### 4. Adaptability to Web Technologies:

Ensures ongoing relevance by adapting to evolving web technologies for sustained effectiveness.

## CHAPTER 2: OVERALL DESCRIPTION

### 2.1. Product Perspective:

The assistant is to be created as a portable hardware device that contains mic and speaker along with the abilities such as voice assistant, voice recognition, cloud connection. This can be used to control the webpage using only voice recognition.

### 2.2. Product Features

#### 1. **Website Analysis:**

Capable of comprehensively analysing diverse websites, considering variations in structure, design, content, and HTML elements.

#### 2. **Content Understanding:**

Advanced algorithms enable the Voice Assistant to accurately interpret and comprehend the content present on web pages.

#### 3. **Voice Interaction:**

Users can interact solely through voice commands, offering a natural and intuitive interface for individuals, particularly those with visual impairments.

#### 4. **Task Execution:**

Performs actions on web pages based on vocal instructions, including tasks such as retrieving information, navigating through pages, and interacting with website elements.

#### 5. **Adaptive Learning:**

Incorporates adaptive learning mechanisms to improve accuracy and efficiency in understanding user preferences and nuances over time.

**6. Contextual Awareness:**

Enhances user experience by maintaining contextual awareness, allowing for more meaningful interactions and responses.

**7. Accessibility Across Platforms:**

Ensures compatibility and accessibility across various web platforms, including educational portals, e-commerce websites, social media platforms, and more.

## 2.3. User Classes and Characteristics

**1. Accessibility-Dependent Users:**

*Includes:* Visually impaired users, individuals with motor disabilities.

*Characteristics:* Rely on voice interactions for web accessibility, seeking ease of navigation and task execution.

**2. Transactional and Service Users:**

*Includes:* E-commerce shoppers, government service users.

*Characteristics:* Seek independent and accessible online shopping, efficient interaction with government portals for information and transactions.

**3. Information and Social Interaction Users:**

*Includes:* News and information seekers, healthcare information seekers, social media enthusiasts, entertainment consumers.

*Characteristics:* Navigate news websites, access information, engage with social networks, and discover entertainment through voice interactions.

#### 4. Diverse Language and Cultural Users:

*Includes:* Multilingual users.

*Characteristics:* Prefer interaction in languages other than the default, benefiting from multilingual support for a personalized experience.

### 2.4. Operating Environment:

The hardware can be portable and be connect to cloud from anywhere.

The required modules are:

1. Raspberry pi.
2. Mic.
3. Speaker.

### 2.5. Technologies Used

1. Python
2. Natural Language Processing
3. Selenium

### 2.6. Requirements

All the required elements will be inbuild in the hardware device like:

1. Raspberry pi ( min 2GB ram) processor
2. Mic.
3. Speaker.
4. WIFI network

## CHAPTER 3: FEASIBILITY STUDY

### Introduction:

The feasibility study aims to assess the practicality and viability of developing an adaptive voice assistant for enhancing web accessibility, particularly for visually impaired users. This study examines various aspects including technical feasibility, economic feasibility, and operational feasibility.

### 3.1. Technical Feasibility:

- **Data Availability:** Adequate datasets of diverse webpages(Amazon) are available for training models and natural language processing algorithms.
- **Technological Infrastructure:** Required technologies such as natural language processing and voice recognition are mature and readily accessible.
- **Integration Challenges:** Integrating various technologies and algorithms into a cohesive voice assistant system may pose technical challenges, but feasible with proper expertise and resources.

### 3.2. Economic Feasibility:

- **Cost Analysis:** The development of the voice assistant may involve expenses related to hardware devices.
- **Return on Investment (ROI):** The potential benefits of improved web accessibility, including expanded user base and positive societal impact, may outweigh development costs.
- **Revenue Generation:** Potential revenue streams include licensing the technology to web developers, offering premium features, or partnering with assistive technology providers.

### 3.3. Operational Feasibility:

- User Acceptance: Conducting user testing and feedback sessions will be crucial to ensuring that the voice assistant meets the needs and expectations of visually impaired users.
- Training and Support: Providing adequate training and support to users and developers will be essential for successful adoption and utilization of the voice assistant.
- Scalability: The system should be designed to scale effectively to accommodate a growing user base and evolving web technologies.

### 3.4. Legal and Ethical Considerations:

- Accessibility Regulations: Compliance with accessibility standards and regulations, such as the Web Content Accessibility Guidelines (WCAG), is essential to ensure legal compliance and user satisfaction.
- Ethical Usage: The voice assistant should be designed and used in a manner that respects user privacy, autonomy, and dignity.



## CHAPTER 4: DESIGN

### 4.1. Use Case Diagram:

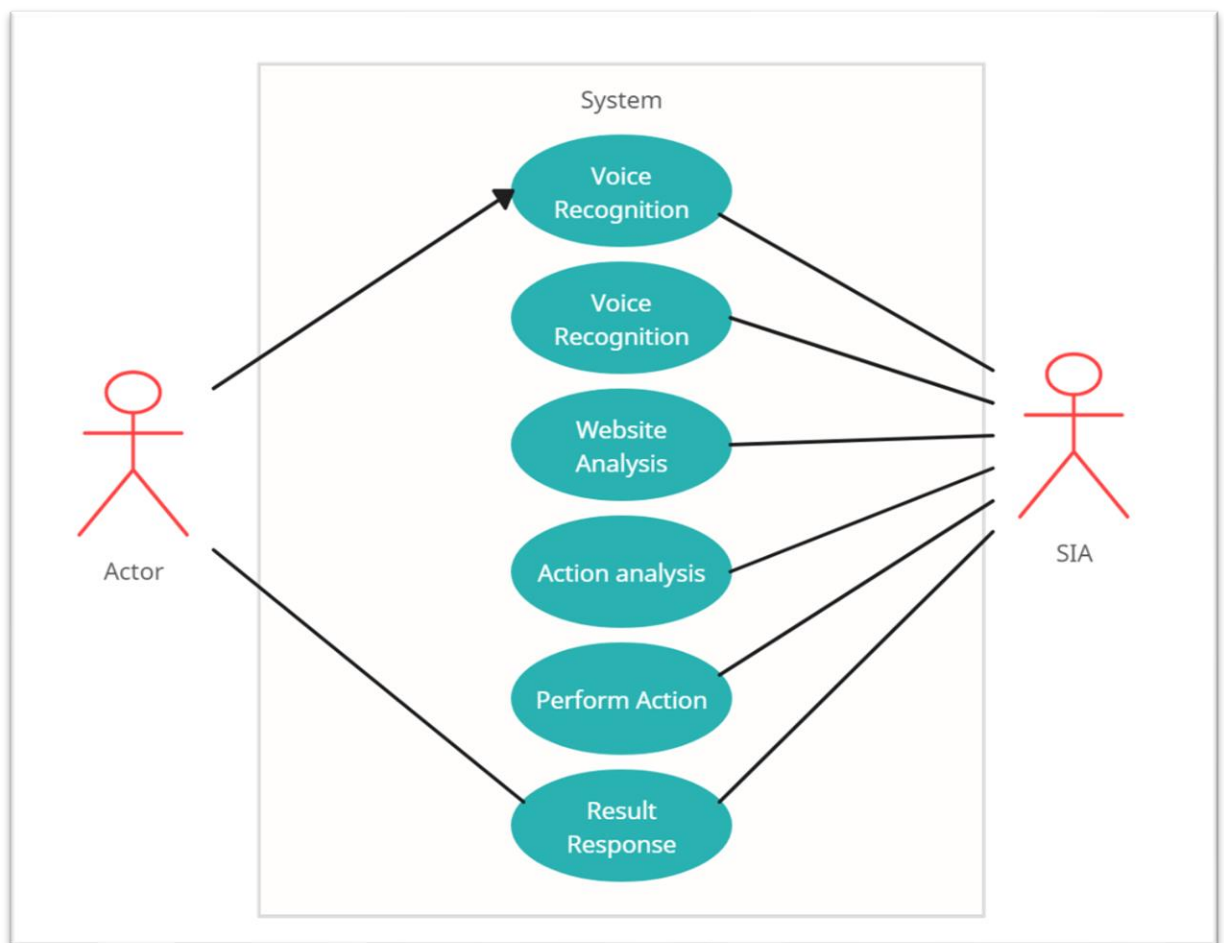


Figure 1: Use Case Diagram

## 4.2. SIA State Diagram:

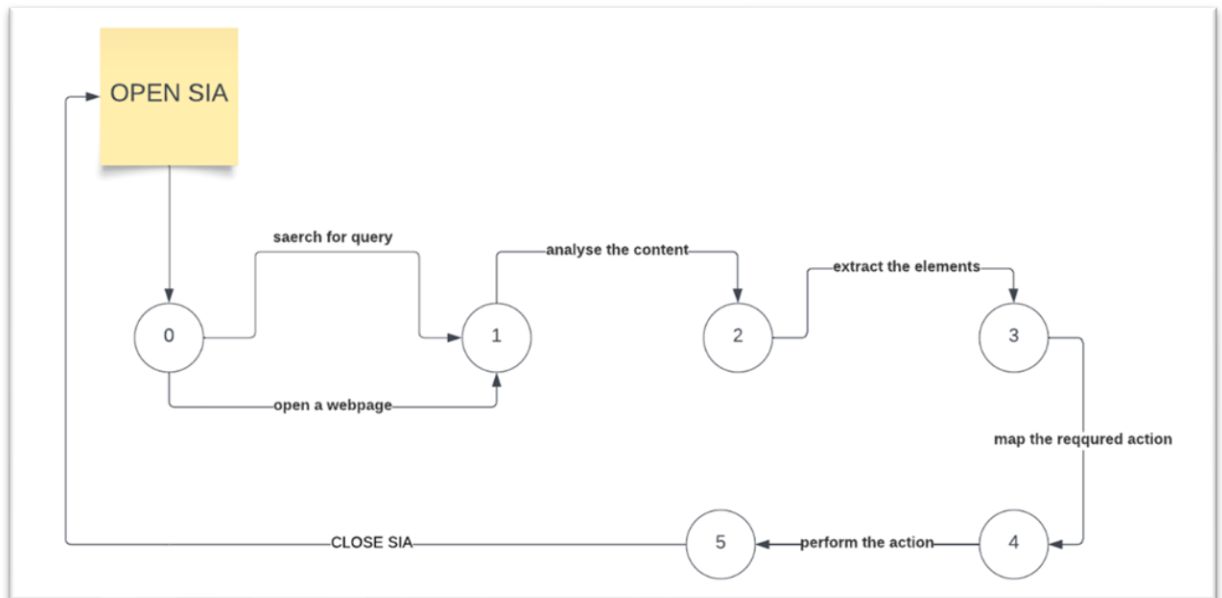


Figure 2: SIA State Diagram

### 4.3. Sequence Diagram

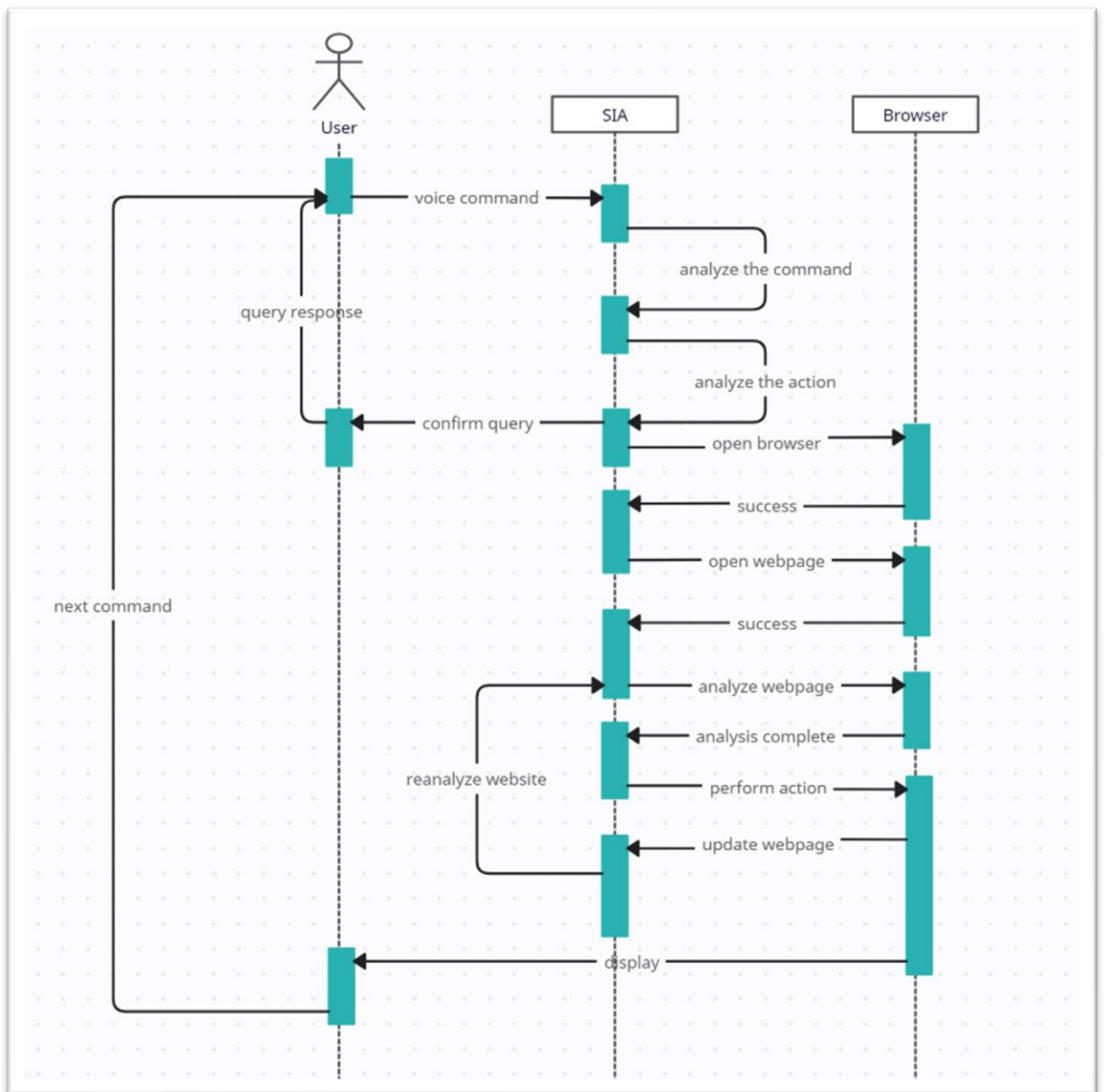


Figure 3: Sequence Diagram

## CHAPTER 5: IMPLEMENTATION

### 5.1. Algorithm:

Steps:

1. Start
2. Identify the user command.
3. Analyse the command using the Natural processing tools.
4. Understand the action to be performed.
5. In case of any confirmation or additional response is required:
  - a. Ask using a voice command.
  - b. Go to step 2.
6. If the command is to Quit processing:
  - a. Go to step 12.
7. Open web browser or connect to cloud or internet.
8. Open or modify the web page
9. Analyse the webpage.
10. Provide output response using the voice command.
11. Go to step 1.
12. Stop

## 5.2. Flow Chart

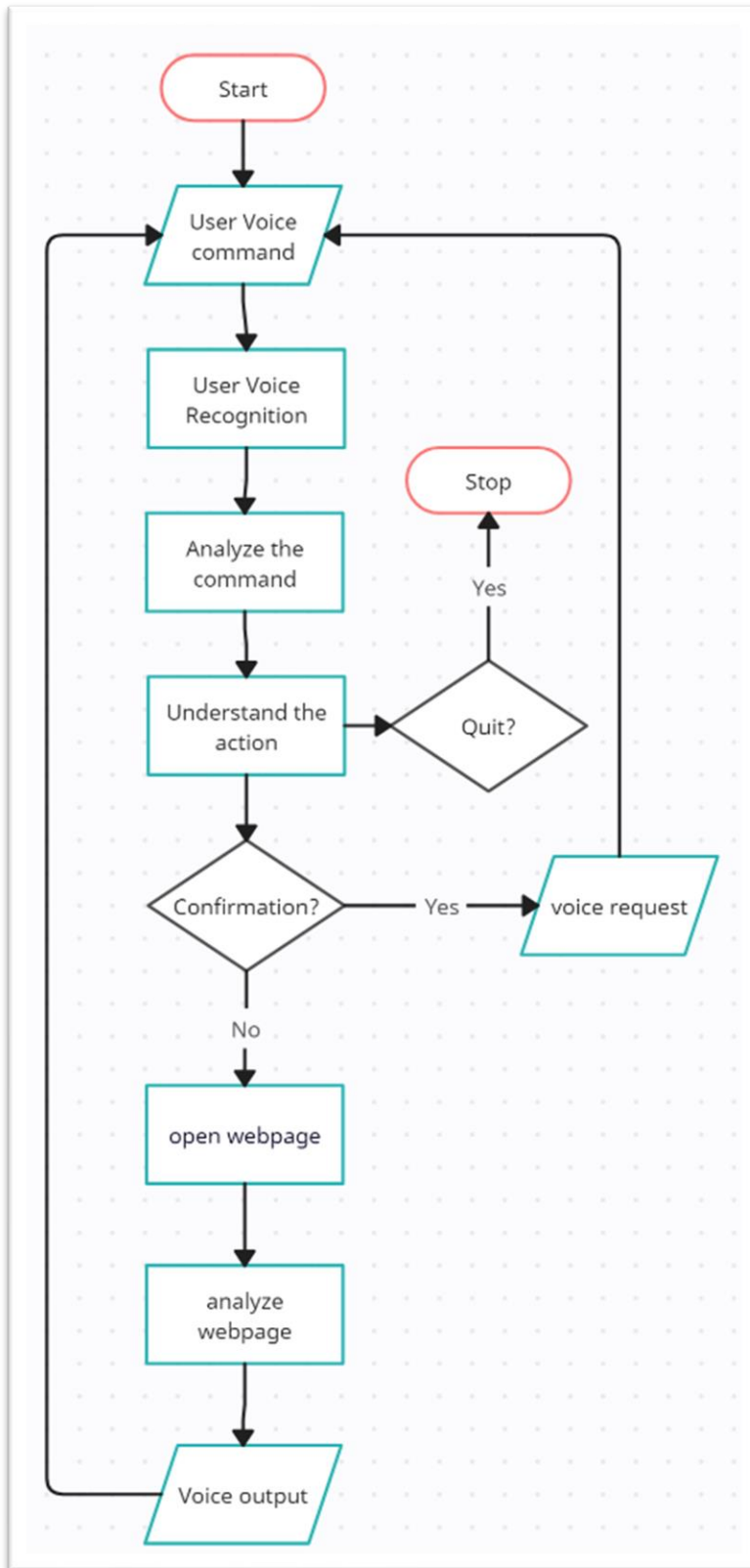


Figure 4: Flow Chart

## CHAPTER 6: OUTPUT

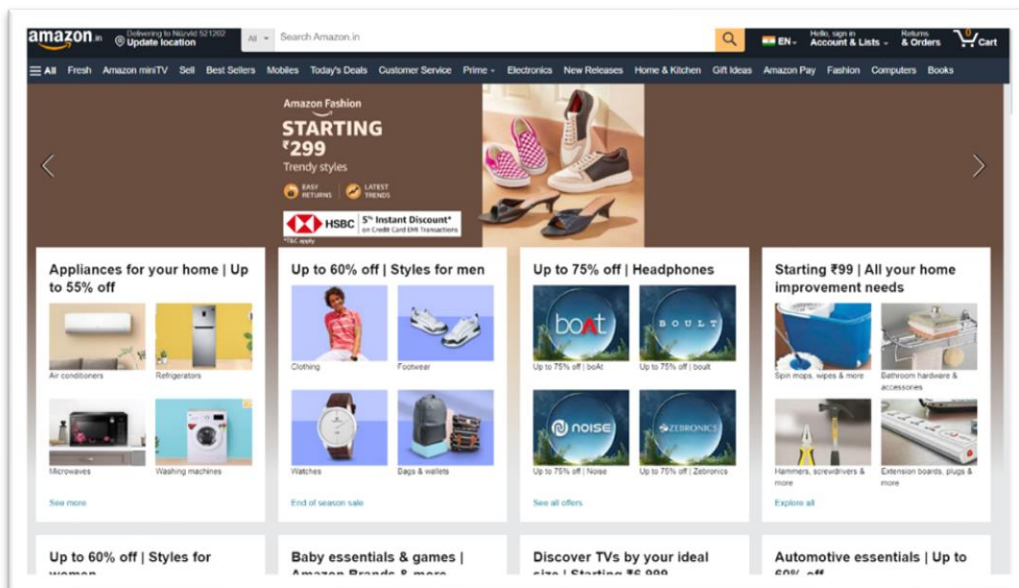


Figure 5: output 1

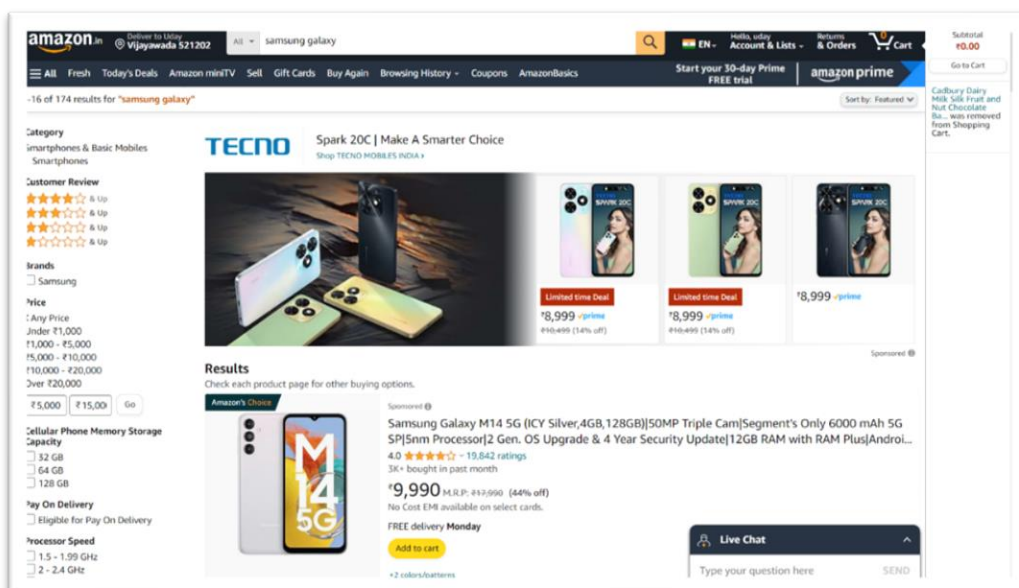


Figure 6: output 2

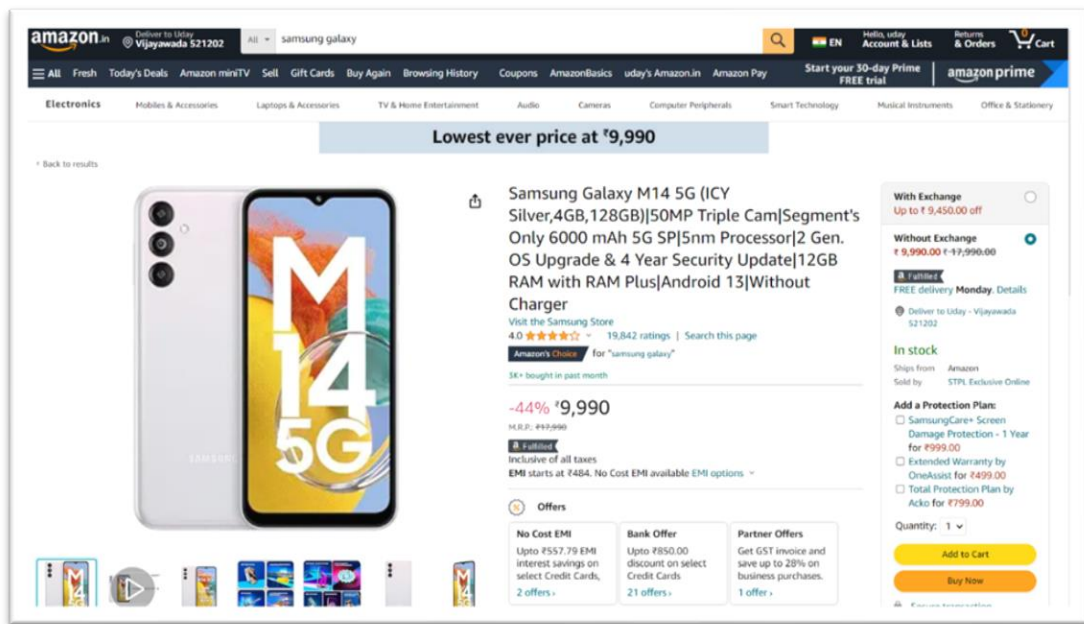


Figure 7: output 3

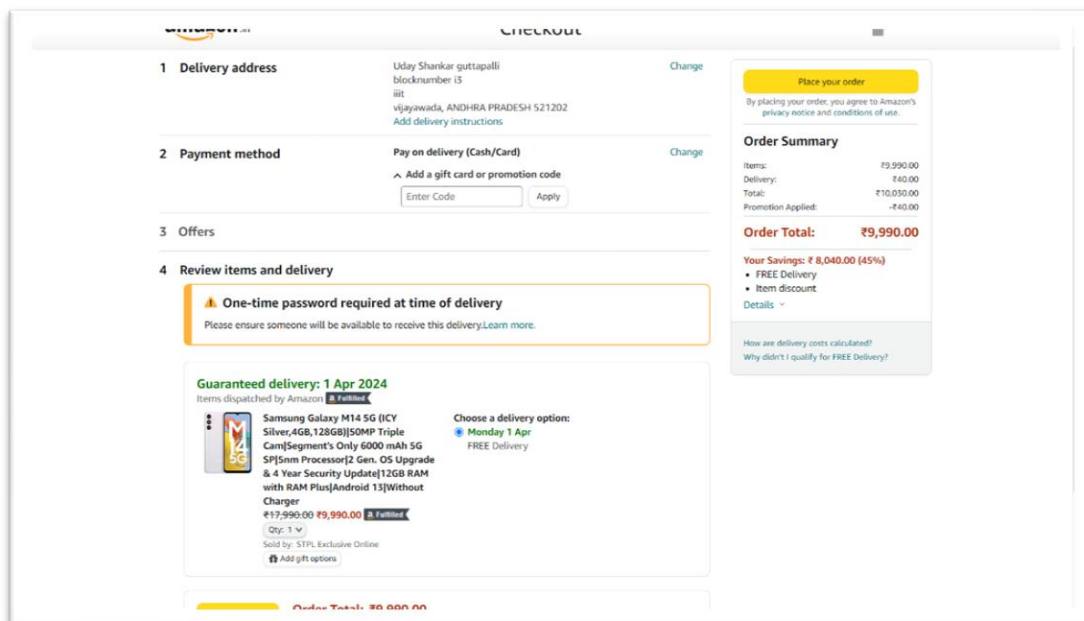


Figure 8: output 4

## CHAPTER 7: HARDWARE

### 1. Raspberry Pi



*Figure 9: Raspberry pi*

### 2. Speaker



*Figure 10: Speaker*

### 3. Microphone



*Figure 11: Microphone*



## CHAPTER 8: NON-FUNCTIONAL REQUIREMENTS

### 8.1. Performance:

- The voice assistant should respond to user commands within a reasonable time frame, preferably less than 1 second.
- It should be capable of handling concurrent user interactions without experiencing performance degradation.
- The system should be scalable to accommodate an increasing number of users and web traffic.

### 8.2. Usability:

- The voice assistant interface should be intuitive and user-friendly, with clear prompts and feedback mechanisms.
- It should support natural language understanding to interpret user commands accurately and respond appropriately.
- The system should provide customizable preferences for users, such as voice speed, volume, and language.

### 8.3. Reliability:

- The voice assistant should have high availability, with minimal downtime or service interruptions.
- It should be resilient to errors and capable of gracefully handling unexpected inputs or system failures.
- Regular backups and redundancy measures should be implemented to safeguard against data loss or service disruptions.

## CHAPTER 9: RESULT

### 9.1. Progress

- The project works completely for the E-commerce website: Amazon for most of the features provided in it.
- A hardware device has been designed for the project that would not require a personal computer to run.

### 9.2. Drawbacks:

#### 1. Handling CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart):

CAPTCHA mechanisms are commonly used on websites to prevent automated bots from accessing or interacting with web content. However, these mechanisms often rely on visual or audio challenges that are difficult for voice assistants to solve. As a result, the voice assistant may struggle to bypass CAPTCHA barriers, limiting its ability to access certain webpages or perform specific actions.

#### 2. Handling OTP (One-Time Password) Authentication:

Many websites utilize OTP authentication as an additional layer of security for user login or verification processes. OTPs are typically sent via email or SMS and require user input to complete the authentication process. Voice assistants may face challenges in retrieving OTPs from emails or messages and entering them accurately on behalf of the user, potentially hindering access to protected content or functionalities.

#### 3. Handling Dynamic Objects:

Webpages often contain dynamic elements and interactive components that change in response to user actions or real-time data updates. These

dynamic objects, such as dropdown menus, sliders, or pop-up windows, may pose challenges for voice assistants in accurately identifying and interacting with them. As a result, the voice assistant may struggle to navigate or manipulate dynamic webpage elements, impacting the user's browsing experience.

#### **4. Handling Personal Passwords:**

User authentication and password management are critical aspects of web security and privacy. While voice assistants can assist users in accessing and interacting with web content, they may face limitations in securely managing personal passwords and sensitive account information. Storing or transmitting passwords through voice commands raises concerns about privacy and security risks, potentially exposing users to unauthorized access or data breaches.

## CHAPTER 10: CONCLUSION

In conclusion, while the development of an adaptive voice assistant for web accessibility holds promise in enhancing online inclusivity for visually impaired individuals, it's important to address its limitations. Challenges such as handling CAPTCHA, OTP authentication, dynamic webpage elements, and personal passwords must be overcome to ensure effectiveness, usability, and security. With careful planning, resources, and adherence to legal and ethical considerations, the project remains both practical and viable. By addressing these challenges and leveraging innovative solutions, the voice assistant has the potential to significantly improve the browsing experience for visually impaired users, enabling more effective and independent access to web content. Ongoing collaboration, user feedback, and advancements in technology will be crucial for maximizing the impact and success of the adaptive voice assistant for web accessibility, ultimately creating a more inclusive online environment for all users.

## CHAPTER 11: FUTURE SCOPE

As the further extension of the project, we are planning to:

- Make this assistant available for various websites.
- Real time object detection and finding.
- Visually impaired user personal assistant for identifying items near them.

## CHAPTER 12: REFERENCES

1. Speech Recognition Library:  
<https://pypi.org/project/SpeechRecognition/>
2. Selenium Library:  
<https://pypi.org/project/selenium/>
3. Speech to text library:  
<https://pypi.org/project/pyttsx4/>
4. Chrome Driver:  
<https://chromedriver.chromium.org/downloads>
5. Raspberry PI:  
<https://www.raspberrypi.com/documentation/computers/raspberry-pi.html#content>
6. Chrome user manual:  
<https://support.google.com/a/users/answer/9310144?hl=en>
7. Output Reference Video link:  
[https://drive.google.com/file/d/10U0--IDorKX5XFFJA7IJmIxuA2mIGOE8/view?usp=drive\\_link](https://drive.google.com/file/d/10U0--IDorKX5XFFJA7IJmIxuA2mIGOE8/view?usp=drive_link)