

Voice Based E-Mail System for Visually Impaired People

**Manish Rana¹, Sunny Sall¹, Maya Yogesh Patil¹, Manali Patekar²,
Rajesh L. Gaikwad³, Janhavi Ankit Sangoi¹, Kunal Meher⁴**

¹St. John College of Engineering & Management (SJCEM), Palghar, India

²Vidyavardhini College of Engineering and Technology (VCET), Vasai Road, India

³Shree. L. R. Tiwari College of Engineering, Mumbai, Maharashtra -India

⁴Atlas SkillTech University, Kurla, Mumbai-India

Email: dr.manish_rana@yahoo.co.in.

This article talk about the internet and other technological advancements, communication has been easier in recent years. However, visually impaired persons find it difficult to use this equipment. Despite the fact that numerous new developments have been incorporated to assist them in effectively using the computer, this requires some expertise with the current technology. This study aims to develop an email system that will allow even the most unskilled visually impaired person to interact using the services without any prior training. The system provides the user with a voice based assistant which helps users to perform basic tasks related to handling email by giving voice commands thus providing a user friendly and interactive experience to the users.

Keywords: IVR, TTS, STT, LSTM etc.

1. Introduction

With the help of growing technologies, the lifestyle of individuals has become easy as the work is wiped out to less with more accuracy and efficiency. Communication is one of the fields that has advanced in technology and also the accessibility of the Internet. The foremost standard approach to online communication is email. Email is one of the foremost reliable ways to exchange necessary information and it is also used worldwide. Electronic Mail or simply Email is defined as a way of exchanging information among people using electronic devices such as computers, mobile phones, tablets etc. Despite the fact that it is the most reliable and the most efficient way to communicate over the internet, not everyone is capable of using it. There are more than 4 billion email accounts to be estimated and by the end of 2023, the number of email users will be approximately 4.3 billion, which is an improvement of 100 million users each year. But there are also gifted people in our society who are not

able to use email. Every time finding a third person to compose and send mail is not possible, as there might be personal content and to tell the entire content of the mail is not feasible. Therefore, the idea of developing society by helping the physically challenged people by providing the ability to send and receive emails through voice commands without using any keyboards.

2. LITERATURE REVIEW

In paper [1] a vivid methodology and implementation of contextual intent classification is shown. For the implementation they have used s Bi-LSTM, Feed Forward, GRU neural network. The main aspect of the paper is on improvising the intent classification in Natural Language Understanding. Evaluation has been performed on 2550 real user queries. The research also includes a speed comparison assessment of the various models used. It also specifies that performance of the model declines if training data is unable to represent real user queries. The paper also provides light on contextual disambiguation of intents and how important it is while understanding user queries. Scaling the provided approach for a large number of contextual utterances is difficult.

In paper [2] a system enabling the blind and illiterate to interact with the email system is proposed. This machine removes the use of IVR technology which are the usage of screen readers, Braille keyboards. In this paper there has been use of STT and TTS conversion. Voice instructions are also used for activities and for registration, you may use your identity, your email address, and your password. This is functionality to use the function that tells PHP to email. It is the functionality to use the functions that tells PHP to email. The system's environment is clean, the voice-controlled by a feedback system in each step. This scheme is that it uses Gmail as a host server so that we can make use of other email services like Yahoo, Google, etc.

In paper [3], it introduces the voice based email system for visually impaired people which is more effective than existing systems. IVR stands for interactive voice response, and it refers to how a user interacts with a voice message. The proposed System is an opportunity of the current system which has extra capabilities along with voice to text and text to voice conversation. It completely eliminates the use of keyboards and we would be allowed to enter the object only by using our voice. This device is beneficial for visually impaired humans.

The paper [4] has emphasized on the problems faced by the blind people while accessing the general email system and tries to put a solution for it. Speaking about the email system, all these email systems have been designed addressing the normal user. Therefore, they are not very accessible to the blind people. An Automatic Speech Recognition System (ASR) is available, but they have their own set of limitations. All these software are built for a limited number of languages so there is a language barrier. Also they have specific requirements and configuration so there is platform dependency. Speaking about ASR, the performance of it decreases drastically when there is a noisy environment, and it also provides slower results due to the processing involved in it. This paper tries to address all these problems by proposing a simple solution, here instead of ASR, the user's voice is sent as it is. Here since

voice is sent as it is there is no language barrier and local languages can also be used. Since this system doesn't need ASR there is drastic reduction in the computation and thereby light weight application is produced. Here the feedback is given in audio format. For composing mail and doing the rest of the navigation simple mouse clicks shortcuts are enabled. Here recipient mail address has to be written through keyboard so this requires the user to have the knowledge of keyboard. Now due to involvement of keyboard and mouse a complete hands free experience is not provided to the user but still due to its simple and easy to use architecture it provides greater convenience to the blind user than the traditional email system.

The paper [5] majorly focuses on proposing an A.I based voice assistant architecture which enables hands free voice control of their mobile devices. To make the system more robust there has been an idea of using text to speech and facial recognition. In this paper for understanding users speech Google speech-to-text API has been used. This API takes input as audio and processes the audio and returns the corresponding text. Facial recognition has been used to eliminate the systems need of authentication through password. There has been mention of natural language understanding but the paper fails to explain the working of it and how it helps to implement the system. The term facial recognition has been used vaguely and no clear description of algorithm or strategy to implement this has been explained. The term 'Brain' used in this paper has not been clearly described. It also fails to explain the working of response generation.

This paper [6] focuses on the progress made in developing a new technique for visually impaired people to receive email using voice and text. It makes use of Tinker, Python libraries, and an interactive voice response system (IVR). The goal of this project is to make it easier for people who are blind to send and receive voice mail. Low eyesight and visual impairment affect many people across the world. This system gives an area of position where users are interacting by way of communication. The framework utilizes TTS (Text-to-Speech) for visual impairment to give navigation administration through voice. System is based on Google gTTS API speech to text and speech to speech conversion. Can be efficiently used by any person such as a common or blind person. Based on conversion of written text into speech which helps the blind person to listen and read the audio based emails. Python libraries used in voice based email service extended for visually impaired persons. This e-mail system helps the visually impaired by providing audio based emails for easy to understand and access. The required content is converted to voice audio by gTTS and saved in an mp3 file. . This will lessen the cognitive load of the visually impaired person in terms of remembering information and procedures, as well as character typing labor on the keyboard. This enables visually impaired people to take a step forward in a productive manner in order to reach their goals. The disadvantage of the system is that it is not flexible for different user utterances and for that they did not use any NLP model.

The article suggests [7], the application is a web-based tool for visually impaired people that uses IVR (interactive voice response), allowing them to operate their email accounts using only their voice and read, send, and perform all other useful operations. The system will provide the user voice commands to perform a certain action, and the user needs to respond. The main benefit of this technology is that it does dispense the necessity for a keyboard; instead, the user must rely only on speech and mouse clicks to respond. The application uses

Google Speech-to-text and text-to-speech APIs. One of the application's drawbacks is that the user must recall mouse clicks, whether single or several. One of the application's drawbacks is that the user must remember mouse clicks, whether single, double, left or right and what will happen after each click and it is not flexible for different users' utterances like if a user gives different commands then it will not recognize.

3. PROPOSED WORK

Unlike the existing system, which emphasizes user friendliness for normal users, our approach prioritizes user friendliness for all types of people, including normal persons who are visually impaired.

The architecture of our proposed system is shown in figure. The diagram shows the major components of the present system, which are:

- User
- Web Server
- TTS(Text to Speech)
- Predict Response
- Intent Classification
- STT(Speech to Text)
- Name Entity Recognition
- Gmail Database

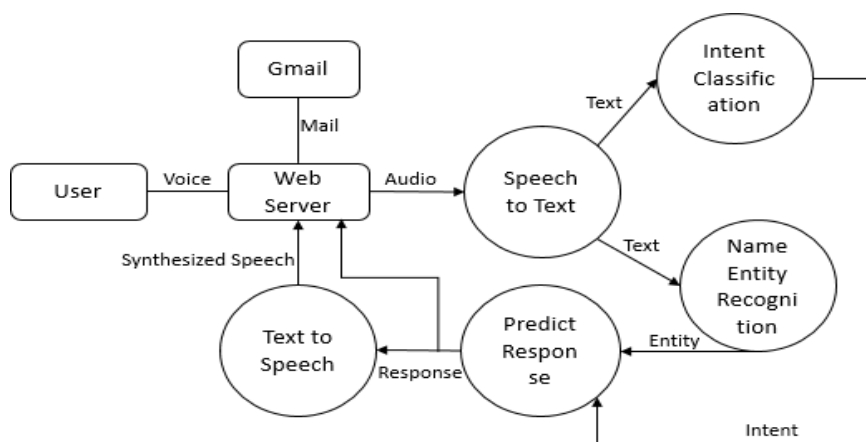


Fig.1.Architecture for Voice Based Email System

1.1 User:

User is a person who will use this architecture to access the services provided by it. The process of the entire system is initiated by the user. Users can use this platform to access

mail through the use of his/her voice. This platform provides users all the functionalities of normal email.

1.2 Web Server:

The web server is an important part of the system since it serves as a link between the user and the email system. The web server's primary function is to take user requests and forward them for processing; it also handles the gmail database and, after receiving data from other components, it is responsible for providing an appropriate response to the user.

1.3 Text to Speech:

Text to speech is the conversion of text into audio and outputting it to the user[8]. This is a very critical task and we need to consider various things about generated speech. Before converting text to speech care should be taken that the given text is free from ambiguity and the same information in the text should not be repeated. The generated speech should be clearly audible and understandable. Along with this we also need to take care of the pronunciations of words with proper pauses in between. The following module will be implemented by the use of Javascript API which can be easily embedded into the front end code.

1.4 Predict Response:

This part deals with what type of response should be given to the user depending upon input query. The predicted response should be relevant to the user command. The response should be such that it gives maximum satisfaction to the user. To achieve this there will be use of both intent classification and entity recognition.

1.5 Intent Classification:

Intent classification is the process of classifying the user's intent by analysing the user's command. This can be thought of as an automatic association of text to a specific purpose or goal[8]. Here we take user's utterances and classify them into one of the several predefined classes. For intent classification we are using deep learning.

1.6 Speech to text:

Speech to text conversion is the process of converting spoken words into written texts[9]. This process is also often called speech recognition. This is used to convert the voice command by the user into text which can be used for extracting information. The speech to text model should be independent of the user's accent and pronunciation. This model should be able to convert speech to text with high accuracy as poor accuracy will degrade overall systems quality and smooth flow between users would be compromised.

1.7 Entity Recognition:

This is the process of identifying and categorizing key information in the text present. This is a two-step process

1. Detect a named entity
2. Categorize the entity

For example, consider the sentence ‘Send this mail to Vinay at 12:15 pm’. Here Vinay is the name of a person and 12:15pm is the time. So entity recognition tries to extract such types of features from the text.

1.8 Gmail Database:

This is used to store all the necessary information of the user which can also be used for user authentication and verification [9]. It further contains all the necessary information needed by the voice assistant to process user commands and generate responses.

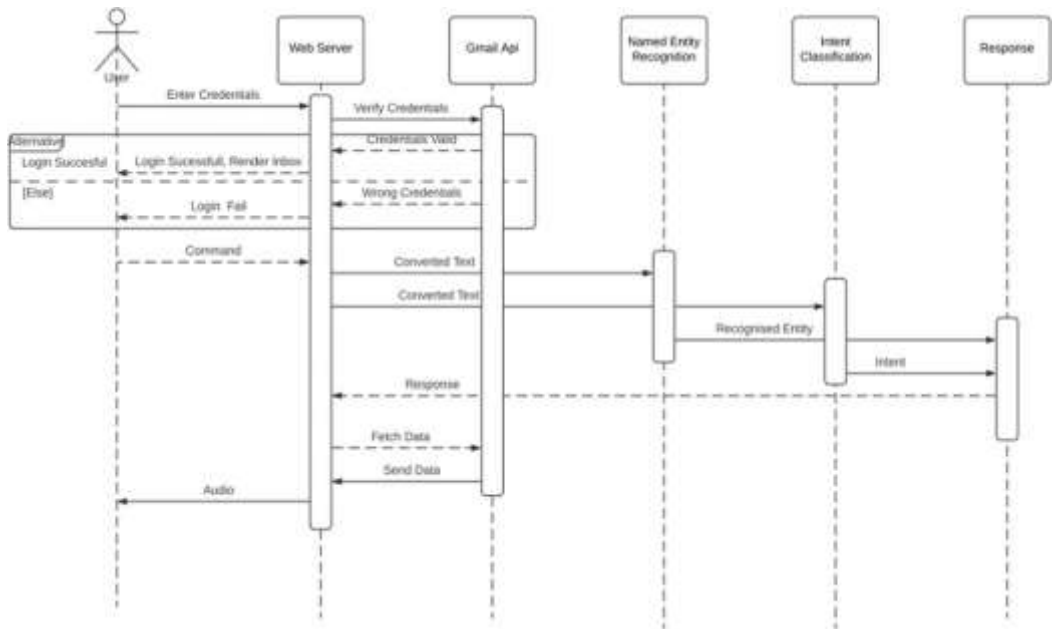


Fig: 3.2. Sequence Diagram for Voice Based Email System

The figure 3.2 represents how each component interacts with each other and the order of interaction.

4. RESULTS & DISCUSSION

If a user is using this app for the first time, he or she must grant the system permission to access their Gmail account. After giving the permission the user is directed to the inbox page of their email.

Here the user will get an immediate voice response of the basic functionalities. If the user is using the website for the first time then he/she will be asked to navigate to the help page, he/she can navigate to that page by giving a voice command and pressing the key on the keyboard. A video demonstrating how to use this application will be available on the help page.

Nanotechnology Perceptions Vol. 20 No.6 (2024)

Similar functionality is present on the sent mail page.

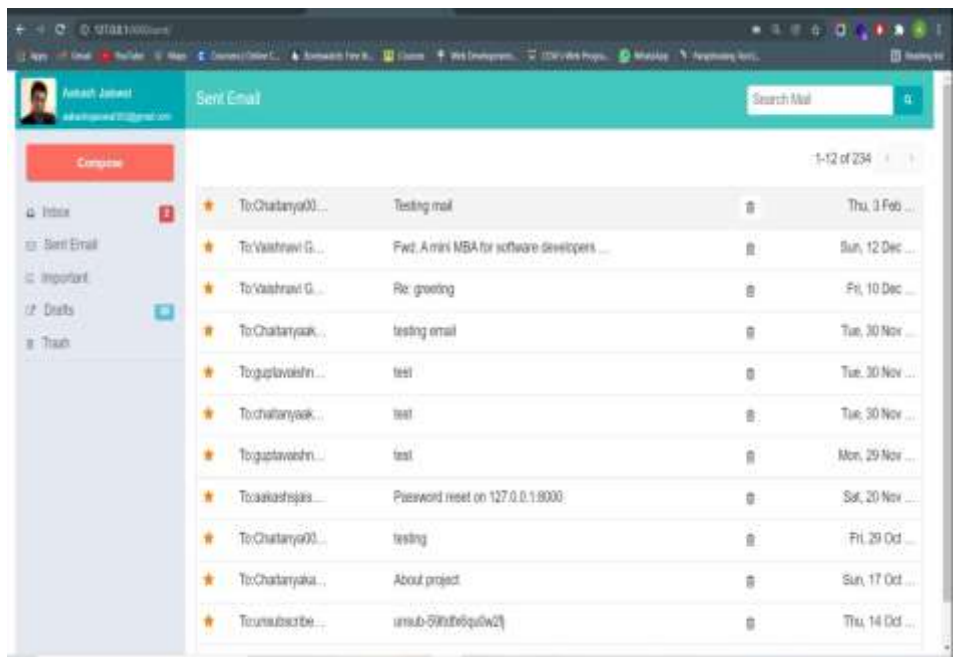


Fig: 4.3 Send Mail for Voice Based Email System.

Search feature for searching a particular mail in either inbox or sent mail is also provided. The user can further traverse mails from the list and also open a particular mail.

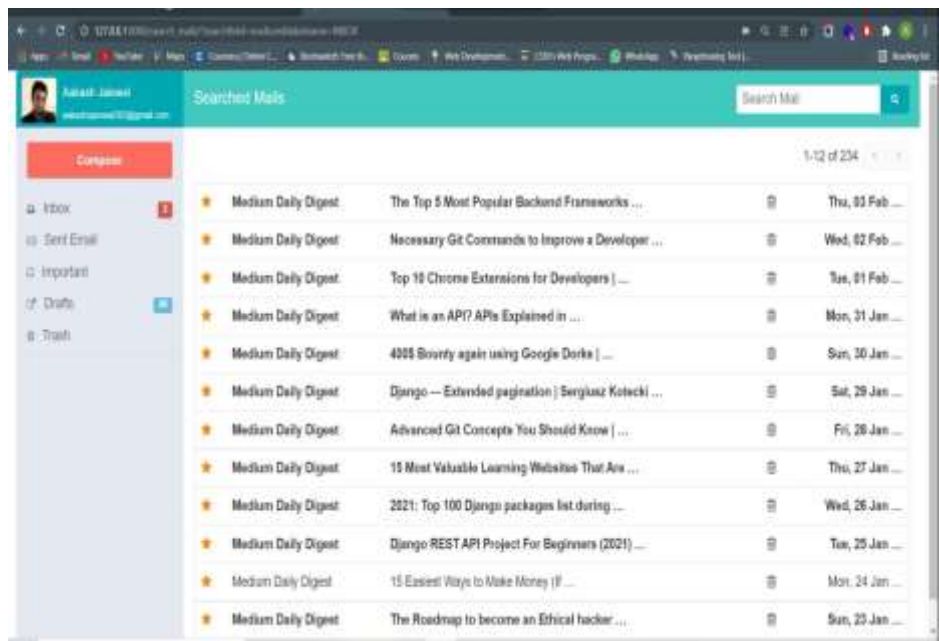


Fig: 4.4 Search mail for Voice Based Email System.

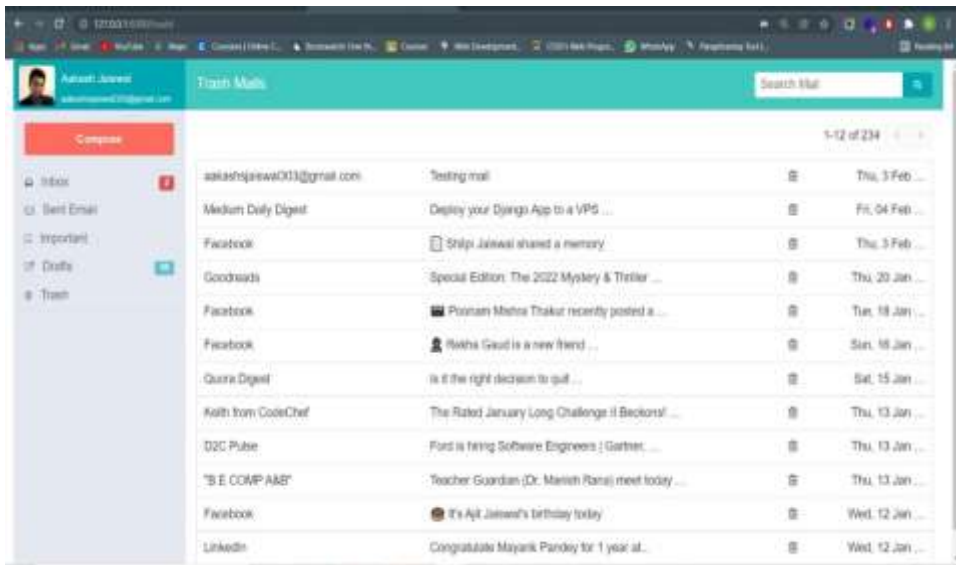


Fig: 4.5 Trash Mail for Voice Based Email System.

A user can also access the mails present in the trash Composing emails can also be done with the help of voice commands. Here for every field in compose mail a response from the user is taken providing a user an interactive experience. After the user has finished writing the message, before sending the mail, the entire message is read again for confirmation.

Conclusion

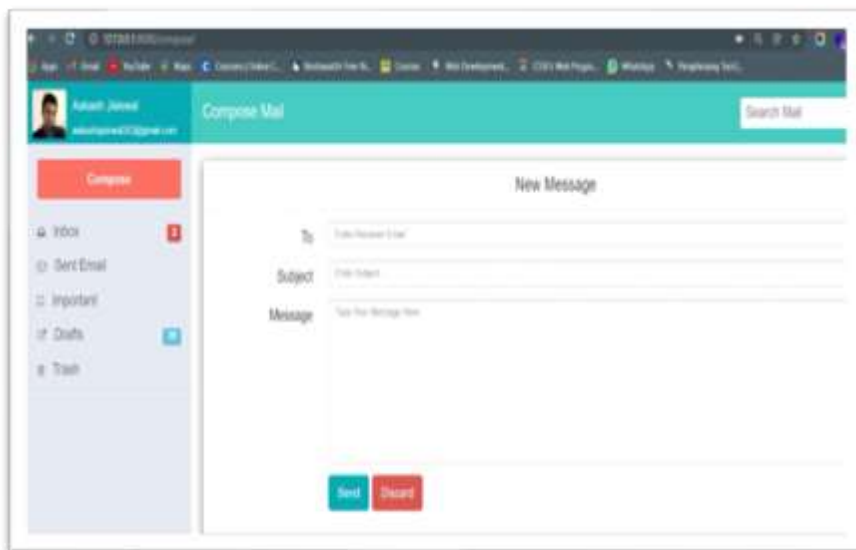


Fig: 4.6 Compose Mail for Voice Based Email System.

The manuscript presents a Voice Based Email System designed to facilitate user interaction with email services through voice commands, particularly focusing on first-time users. The following sections summarize the key findings and discuss their implications.

Key Findings

- User Onboarding:

Upon first use, the requirement for users to grant permission to access their Gmail accounts is standard and crucial for functionality. This initial step directs users to their inbox, where they receive immediate feedback, enhancing their initial experience.

- Voice Interaction Features:

The system's ability to provide an immediate voice response about basic functionalities upon entering the inbox is a notable advantage. This feature supports users in familiarizing themselves with the application without needing extensive reading.

- Help Page Navigation:

New users can easily access a help page via voice commands, complemented by a video demonstration. This dual approach caters to different learning styles, ensuring users can quickly understand how to utilize the app effectively.

- Email Navigation and Management:

Users can navigate their inbox using voice commands and arrow keys, with the system reading out basic information for each email. This feature streamlines email management and enhances the experience for users who may struggle with traditional navigation.

- Automatic Email Reading:

Once an email is opened, it is read aloud automatically, providing a hands-free experience that allows users to engage without having to read visually.

- Composing and Sending Emails:

The voice-activated composing feature engages users interactively by requesting input for each field, followed by a confirmation read-back of the entire message before sending. This reduces errors and increases user confidence in their communications.

- Search and Trash Access:

The inclusion of a search feature for locating specific emails, along with access to the trash folder, enhances the system's usability, making it easier for users to manage their emails effectively.

Discussion

The Voice Based Email System represents a significant advancement in email interaction, particularly for users who may benefit from hands-free operation or who have accessibility needs. The immediate voice responses and structured navigation simplify the user experience, making email management more intuitive.

Accessibility and Inclusivity: This system addresses the needs of diverse users, including those with disabilities, by allowing for voice-based commands. By reducing reliance on visual interfaces, the application can be a powerful tool for enhancing accessibility in digital communication.

Learning Curve: The combination of voice prompts and a help page helps to reduce the learning curve for new users. This aspect is particularly important in fostering user confidence and encouraging ongoing engagement with the application.

Future Enhancements: While the current features are robust, future iterations could include enhanced voice recognition capabilities, support for additional languages, and integration with other communication platforms. User feedback mechanisms could also be established to continually improve functionality based on actual usage patterns.

5. Future Scope

The future scope of the Voice Based Email System for visually impaired users presents exciting opportunities for enhancement and wider application. Here are several key areas to consider:

- **Multilingual Support:** Expanding the system to support multiple languages will make it accessible to a broader audience. This could include regional dialects and languages, ensuring that users from diverse backgrounds can engage with the platform.
- **Enhanced Natural Language Processing (NLP):** Investing in advanced NLP capabilities will allow the system to better understand user commands, enabling more fluid interactions. This could include recognizing context, handling complex queries, and improving overall response accuracy.
- **Personalization and Adaptability:** Developing algorithms that adapt to individual user preferences and behaviors could enhance user experience. Features like customizable voice options, tailored notifications, and learning user habits will make the system more intuitive.
- **Integration with Other Services:** Connecting the email system with other applications (e.g., calendars, task managers) can provide a holistic digital assistant experience. This integration would streamline workflows for visually impaired users.
- **Feedback Mechanisms:** Implementing robust feedback mechanisms will help users report issues or suggest improvements. This input can guide iterative development, ensuring the system evolves based on real user experiences.
- **Cross-Platform Accessibility:** Ensuring the system is accessible across various devices (smartphones, tablets, and computers) and operating systems will increase usability. This includes optimizing for different screen readers and accessibility tools.
- **Community Engagement:** Collaborating with organizations and communities representing visually impaired individuals can provide valuable insights into user needs and preferences, fostering a user-centered design approach.

- **Security and Privacy Features:** Given the sensitivity of email communications, enhancing security features tailored for visually impaired users will be critical. This includes voice-activated security protocols and easy-to-understand privacy settings.
- **Research and Development:** Ongoing research into assistive technologies will keep the system aligned with the latest advancements, ensuring it remains relevant in an evolving digital landscape.
- **Pilot Programs and User Testing:** Implementing pilot programs with real users will allow for testing and refining the system in practical environments, ensuring it meets the actual needs of visually impaired users.

6. Conclusion

The paper tries to analyse the problems faced by the visually impaired people for accessing the primitive mail system and tries to propose a feasible and user friendly solution for it. For addressing this problem a website has been implemented to manage user email, and the user has provided the feature to navigate through the website using voice commands. The paper analyzed various literature and products built for the visually impaired. After the study it can be seen that a major problem faced by the visually impaired people with various applications is the lack of interactivity, ease of control and the uniformity of UI. Our system tries to overcome all these shortcomings by providing a voice assistant which assists the user to handle their mails. The system gives an immediate feedback for every command for users and thus guides them for a better utilization of the application. The system currently supports a single language that can be extended by providing multiple language interaction, and thus will help to overcome language barrier. With more robust architecture for natural language understanding better interactivity and response can be provided to the user.

The system should be designed in such a way that it should provide better interactivity to the user along with a faster response. Users should also be allowed to interact with the system in the local language. If all the problems are successfully addressed then an efficient product can be achieved.

The Voice Based Email System outlined in the manuscript offers a comprehensive solution for modern email management, emphasizing user-friendly design and accessibility. Its innovative features not only enhance the experience for first-time users but also promote broader inclusivity in digital communication. Future developments should focus on refining these capabilities to maintain relevance in an evolving digital landscape.

References

1. P. A. Tiwari, P. Zodawan, H. P. Nimkar, T. Rotke, P. G. Wanjari and U. Samarth, "A Review on Voice based E-Mail System for Blind," 2020 International Conference on Inventive Computation Technologies (ICICT), 2020, pp. 435-438, doi:10.1109/ICICT48043.2020.9112539.
2. Rijwan Khan, Pawan Kumar Sharma, Sumit Raj, Sushil Kr. Verma, SparshKatiyar. "Voice

- Based EMail System using Artificial Intelligence". International Journal of Engineering and Advanced Technology (IJEAT) ISSN: i2249– i8958, Volume-9 Issue-3, February, i2020.
3. Pankaj Kumar Maurya, Prince Kumar, Mukesh Kumar, PramodNath. “ Voice Based E-Mail System”. In International Research Journal of Engineering and Technology (IRJET), 2018 of Vol.5, Issue: 4.
 4. T. Dasgupta, A. Anuj, M. Sinha, R. Ghose and A. Basu, "VoiceMail architecture in desktop and mobile devices for the Blind people," 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI), 2012, pp. 1-6, doi: 10.1109/IHCI.2012.6481791.
 5. Mr. Yash Agarwal, Mr. RanjeetRai, Mr. Punit Kumar Chaubey. “BRAIN– THE A.I. (PERSONAL VOICE ASSISTANT)”. In InternationalResearch Journal of Engineering and Technology (IRJET), 2020 of Volume: 07 Issue: 04.
 6. Vinutha H, AmutharajJoyson, Mythri G, Seema K, Varsha V, “Voice Based Email Service for Visually Impaired”, International Journal of Advanced Science and Technology”, Vol. 29, No. 6, (2020).
 7. PronabMukherje, Harsh Agarwal, Gaurav Miglani, “Voice based Email service for visually challenged people”,International Journal of Advanced Science and Technology”, Vol. 26, No. 4, (2017).
 8. Sharma, Arpit. "Improving Intent Classification in an E-commerce Voice Assistant by Using Inter-Utterance Context." In Proceedings of The 3rd Workshop on e-Commerce and NLP, pp. 40-45. 2020.
 9. Saurav Mishra, AshwaniPawar. “VOICE BASED EMAIL SYSTEM FOR VISUALLY IMPAIRED”. In International Journal of Emerging Technology and Advanced Engineering (IJETA), 2017.