2023 International Conference on Advanced Computing Technologies and Applications (ICACTA)

Speech Recognition Email System for People with Impaired Sights

Pratik Kanani
Computer Engineering
Dwarkadas J. Sanghvi College of Engineering, Mumbai
pratikkanani123@gmail.com

Ayush Jain
Computer Engineering
Dwarkadas J. Sanghvi College of Engineering, Mumbai
ayushjain4702@gmail.com

Preet Gada
Computer Engineering
Dwarkadas J. Sanghvi College of Engineering, Mumbai
preet.djsce@gmail.com

Aditya Rakshe
Computer Engineering
Dwarkadas J. Sanghvi College of Engineering, Mumbai
adityarakshe1011@gmail.com

Yogesh Jha
Computer Engineering
Dwarkadas J. Sanghvi College Of Engineering,Mumbai
jhayogesh30@gmail.com

2023 International Conference on Advanced Computing Technologies and Applications (ICACTA) | 979-8-3503-4834-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICACTA58201.2023.10393204

***Abstract*— - In today's culture, where having access to the internet and digital communication is essential for daily life, it is crucial to recognize the importance of accessibility for those with visual disabilities. The proposed email system that utilizes speech recognition technology is a significant advancement that can greatly benefit individuals with impaired sight. It allows these individuals to easily compose, read, and navigate emails using audio and other non-visual means. The project's contribution in enabling blind people to send and receive voice-based e-mail messages is commendable, and the system's GUI has been evaluated and found to perform better than that of existing GUIs. The use of voice-to-text and text-to- voice techniques for accessing emails is also an important feature that makes it user-friendly for individuals with impaired sight. The fact that the system is based on interactive voice response makes it efficient and easy to use. The speech recognition email system represents a significant advancement for individuals with impaired sight, enabling them to stay connected with friends, family, and colleagues without the need for visual assistance. It is important to continue developing such technologies to ensure that digital communication is accessible to all individuals, regardless of their abilities.**

Keywords- GUI, Speech recognition, voice-to-text, text-to-voice, user-friendly, commendable

I. INTRODUCTION

Email is a ubiquitous tool for communication, used by individuals and businesses worldwide. However, visually impaired individuals face significant barriers when using email, as traditional email systems are not designed with their needs in mind. This presents a unique challenge for email developers, who must balance the need for accessibility with usability for sighted users. Visually impaired individuals are a diverse group, with varying levels of visual impairment, and therefore require different levels of assistance when using email. Some individuals may be able to use standard email systems with the assistance of screen readers or other assistive technology, while others may require more specialized email systems designed specifically for their needs. One approach to improving email accessibility for visually impaired individuals is to incorporate assistive technologies into email systems. Screen readers, for example, are software programs that can read the contents of a computer screen aloud, allowing visually impaired individuals to navigate and interact with email in a way that is similar to sighted users. However, screen readers may not be sufficient for all visually impaired users, and more specialized technologies.

Another approach to improving email accessibility is to develop specialized email systems designed specifically for visually impaired users. These systems may incorporate features such as text-to-speech capabilities, and simplified interfaces that make it easier for visually impaired individuals to navigate and interact with email. Additionally, these systems may incorporate voice recognition or other alternative input methods, allowing visually impaired individuals to compose and send emails without the need for a traditional keyboard or mouse. One challenge in developing email systems for visually impaired individuals is ensuring that they are both accessible and usable. While it is important to incorporate assistive technologies and specialized features to improve accessibility, these features should not compromise the usability of the email system for sighted users. Additionally, email systems for visually impaired individuals must be able to accommodate a wide range of user needs and preferences, and therefore must be highly customizable. In addition to compatibility with assistive technologies, email systems for visually impaired individuals should also consider the accessibility of attachments. Often, attachments are not designed with accessibility in mind and may present a significant barrier for visually impaired users. Email systems should, therefore, provide alternative formats for attachments that are accessible to visually impaired individuals, such as audio or text-based descriptions of images and videos.

Furthermore, another important aspect to consider is privacy and security. Email systems must ensure that visually impaired users are not at a disadvantage when it comes to privacy and security concerns. Email systems should provide features that allow visually impaired individuals to easily identify and block spam emails, phishing emails, and emails with malicious attachments. As such email developers shouldmake a long-term commitment to accessibility and continuously work to improve the accessibility of their products. It is essential to

involve visually impaired individuals in the design and development of email systems. Their feedback and input are crucial in ensuring that the email systems are accessible and meet their needs. Involving visually impaired individuals in the design and development process can also help raise awareness of the challenges faced by this community and promote the development of more accessible digital products.

Finally, it is worth noting that email systems for visually impaired individuals have the potential to benefit not only this community but also the wider population. Many of the features and functionalities that are designed to make email systems accessible to visually impaired individuals can also enhance the usability of email systems for sighted users. For example, simplified interfaces, alternative input methods, and customizable settings can make email systems easier and more convenient to use for all users. By prioritizing accessibility and inclusivity in the design of email systems, we can create products that not only meet the needs of visually impaired individuals but also improve the overall user experience for all users.

II. LITERATURE SURVEY

The Email application described in Paper [1] is an example of how email systems can be designed specifically for visually impaired individuals utilizing technology to create a more accessible and user-friendly experience. By incorporating features such as text-to-voice and speech recognition engines, the Email application allows visually impaired individuals to interact with their email in a way that is intuitive and efficient. Additionally, the use of microphones and other input devices can enable users who are blind to easily navigate and select emails, eliminating the need for traditional visual interfaces.

While the Email application is a promising development in the field of email accessibility, there is still much work to be done in this area. As previously mentioned, ensuring compatibility with assistive technologies and addressing the accessibility of attachments are two critical considerations that must be taken into account when designing email systems for visually impaired individuals. Privacy and security concerns also need to be addressed to ensure that visually impaired individuals can use email systems with confidence and without fear of cyber threats. An approach to utilizes spoken commands to navigate and interact with emails was presented in Paper [2]. This approach places a significant emphasis on the usability of a wide range of users, including both sighted and blind individuals, which is critical for promoting greater inclusivity and accessibility in the digital world.

The use of the IMAP protocol to obtain email content from the mail server is another interesting aspect of this system. IMAP is a standard protocol used by email clients to retrieve and access email messages from a mail server. By utilizing this protocol, the system can efficiently and securely retrieve email content, enabling users to interact with their email in real-time. Moreover, the system's ability to interpret spoken commands and execute the appropriate action is a critical feature that enhances its usability for visually impaired individuals. For example, users can simply say "SEND" to create a new email or "READ" to listen to the content of an email. These simple and intuitive commands can significantly reduce the mental effort required to navigate and interact with emails, promoting a more efficient and effective user experience.

While the system described in Paper [2] is a promising development in the field of email accessibility, it is important to note that the success of this system and others like it depends on ongoing support and maintenance. Developers must ensure that these systems are regularly updated and tested to ensure continued accessibility and usability for visually impaired individuals. Additionally, addressing privacy and security concerns is critical to promoting confidence and trust among visually impaired individuals who use email systems and other digital media.

The importance of voice message design in enabling visually impaired individuals to access email and other interactive media components of their working environment is highlighted in Paper [3]. This design has the potential to significantly enhance the accessibility and usability of email systems for visually impaired individuals, especially those who are illiterate and face challenges in remembering and inputting characters on traditional keyboards. However, it is important to note that implementing voice message design in email systems and other digital media requires ongoing support and maintenance. Developers must ensure that these systems are regularly updated and tested to ensure continued accessibility and usability for visually impaired individuals.

Use of technologies such as automatic speech recognition (ASR) and text-to-speech (TTS) that allow users to send and receive emails using their voice was described in Paper [4]. The system also includes a command-based interface that allows users to manage their emails using voice commands. The authors conducted user studies to evaluate the usability and effectiveness of the system, and the results showed that the system was easy to use and improved the accessibility of email communication for visually impaired individuals. In order to develop desktop applications that enable visually impaired individuals to send and receive emails using their voice is designed in Paper [5]. The application includes a login capability for user authentication, ensuring that only authorized users are able to access their email accounts.

Overall, the email system presented in this paper serves as a valuable tool for visually impaired individuals, enabling them to access and manage their email accounts more easily and effectively. The voice-based interface and login capability ensure that the system is both accessible and secure, promoting greater inclusivity and privacy for visually impaired individuals in the digital world.

III. RESEARCH GAP

While the papers mentioned above make significant contributions to the field of email accessibility for visually impaired individuals, there are still some research gaps that need to be addressed:

While some email systems may be accessible for visually impaired individuals to read and compose emails, accessing andunderstanding email attachments may still be challenging.

Some email systems designed for visually impaired individuals may not be fully compatible with all types of assistive technologies, limiting their effectiveness and accessibility.

There is a need to ensure that visually impaired individuals can use email systems with confidence and without fear of cyber threats.

Developing an accessible email system is not enough; it must be regularly updated and tested to ensure continued accessibility and usability for visually impaired individuals.

The accuracy of speech recognition systems may still be a challenge for some users, resulting in frustration and errors.

Many visually impaired individuals may not be familiar with using email or other digital media, so there is a need for training and education to promote effective use of accessible email systems.

IV. METHODOLOGY

The suggested system is not at all like other mail systems already in use and is based on an entirely original concept. The accessibility of the proposed system has been given top priority in its development. A web system is only considered fully accessible if it can be effectively utilized by all types of users, able-bodied or disable. Our method places a greater emphasis on user friendliness for all types of people, including visually impaired and illiterate persons, in contrast to the current system, which places more emphasis on user friendliness for regular users. IVR, or interactive voice response, is the foundation of the entire system.

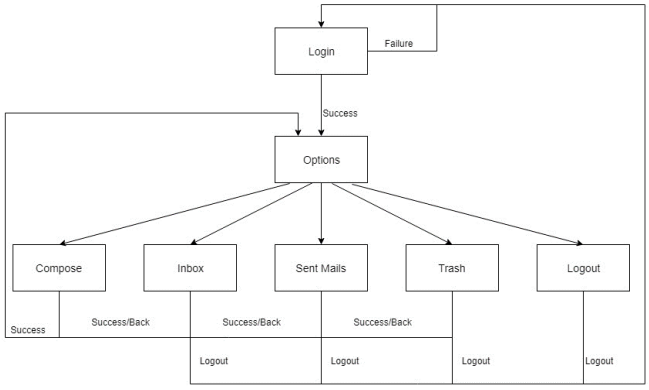

Fig 1 – Architectural Design of the Proposed System

A. SPEECH TO TEXT CONVERTER:

Speech to text converter refers to the process of turning audio or spoken words into text. Often, the method is referred to as voice recognition. Speech comprehension, a more general process of obtaining content from speech, is referred to as speech recognition. It is incorrect to refer to this process as voice recognition or speaker recognition since we frequently relate it with the act of recognizing a person by their voice. Two models dominate speech to text converters. linguistic model, and 1. acoustic model. Typically, systems follow the pronunciation model. Speech recognition is not perfect, just like other pattern recognition technologies. The quality of the speaker's voice, the nature of the speech, and the surrounding circumstances all have a significant impact on the accuracy of speech transcripts. Contrary to popular belief, speech recognition is a more difficult approach to determine a person's personality. Humans are not born with the ability to accurately transcribe speech, and only well-developed speech will do so.

The Speech Recognition library was used to implement the speech-to-text conversion feature in our voice-based email system. The Google Speech Recognition API [6], which was chosen for its precision and usability, was one of the speech recognition engines this library provided an interface to. The recorded audio data was sent to the API in order to use it, and it returned a text transcription of the speech. The relevant data from the transcription, including the email message's content and the recipient, was then extracted.

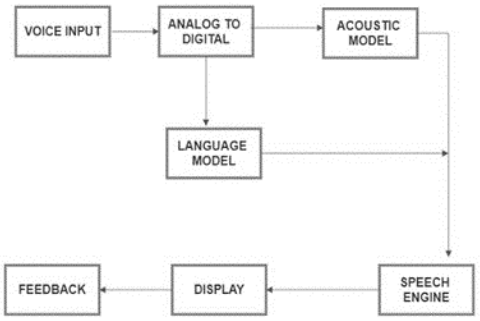

Fig 2 – System Block Diagram For Speech Recognition

B. TEXT TO SPEECH CONVERTER:

Alternative systems, rather than text-to-speech (TTS) systems, produce symbolic linguistic representations that translate speech text to speech. The text-to-speech system consists of two components: a front end and a back end.. First of all, it converts raw text containing symbols like numbers and abstractions into a form that resembles written words. The front end breaks down and marks the text into language units such as phrases, clauses and sentences and assigns spoken transcriptions to each word. Text-to- phoneme or grapheme-to-phoneme conversion is the process of phonetically transcribing words. The symbolic linguistic representation generated by the front-end consists of phonetic transcriptions and prosody data. The symbolic verbal representation is then translated into sound via the backend, also known as a synthesizer. In some systems, this step also includes calculating the desired prosody, which is then applied to the output speech.

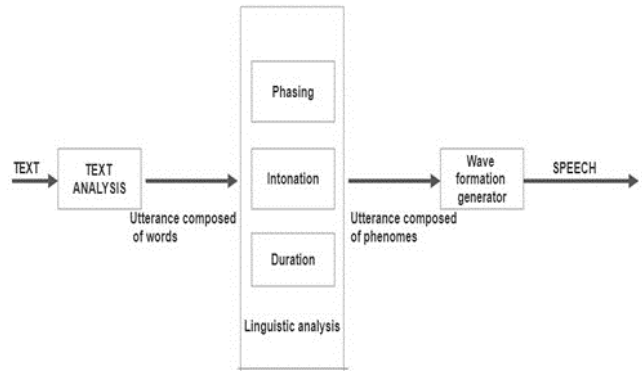

Fig 3 – Text to Speech Conversion

For the text-to-speech conversion functionality, the gTTS [7] library was used, which provided an interface to the Google Text-to-Speech API. The email message content and the email sender information were processed to generate a text string, which was passed to the gTTS library. The library generated an audio file containing the spoken version of the text. The playsound library [8] was used to play the audio file to the user. The gTTS library allowed customization of the voice and language used for the text- to-speech conversion, which was important for making the system more user-friendly for blind users.

V. IMPLEMENTATION AND RESULTS

A. *Login Page :*

Fig 4 – GUI of User Login Page

It is the first page that the user encounters. This asks for the user's email address and authenticates the same. After the email address of the user is confirmed, the user has to enter his password with the use of voice commands. If the user is successfully authenticated the system opens the options page else it asks the user to login again.

B. *Options* Page:

Fig 5 – GUI of Options page

According to the flowchart above, after the user has successfully logged into the system using voice commands to enter his Email-Id and password. The system offers the user a selection of actions to take. Compose, inbox, sent, trash, and logout are the options. The user must make a decision and express what action they wish to take. If a user says compose, the system will take him to the compose page. Similar steps must be taken to read the inbox, send an email, access the trash, and logout. If the user takes an invalid action, the system prompts the user for input again

C. *Compose Page:*

When the user is redirected to the compose page, the system asks for the recipient's email address, and after confirming it, the system asks if he wants to send the email to more recipients, and so on. Following that, the system requests the email's subject and conforms to it. Then it proceeds to the body of the email. The body of the email is composed using voice commands, and the user is then asked if there are any attachments. If the user wants to attach a file to the email, he must speak the filename, and the file will be attached. After that the user has to say yes and the mail will be successfully sent and the system will be redirected to the options page.

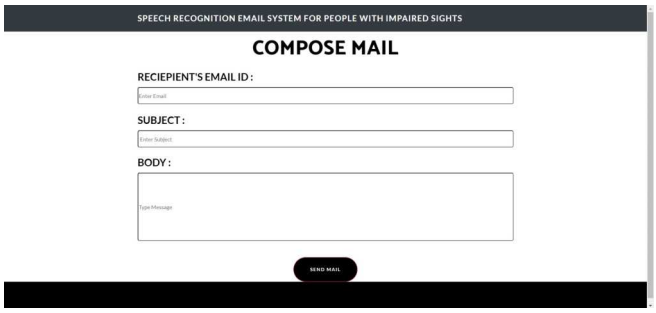

Fig 6 – GUI of Compose Email Page

D. *Inbox Page:*

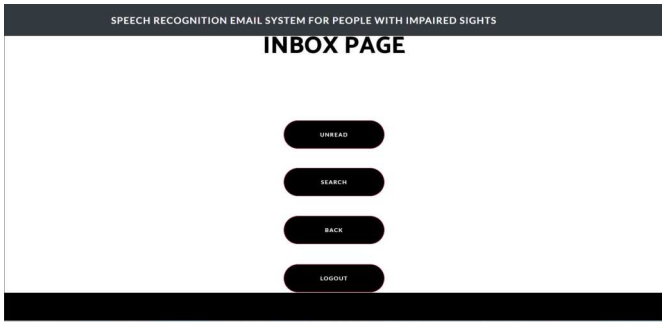

Fig 7 – GUI of Inbox Page

The inbox page in the voiced-based email system offers the user several options, including the unread messages option for quick access to urgent emails, the search option to find specific emails based on keywords, and the back option for easy navigation to the previous page. These options provide a user-friendly interface for the user to efficiently manage their emails.

E. *Sent Page:*

The voice-based email system's sent page has a search function that enables users to quickly locate particular sent emails using keywords, as well as a back button for simple page navigation. Additionally, the logout option enables the user to leave the system securely after they are done. The user can manage their sent emails with the help of these options, which offer a practical and user-friendly interface.

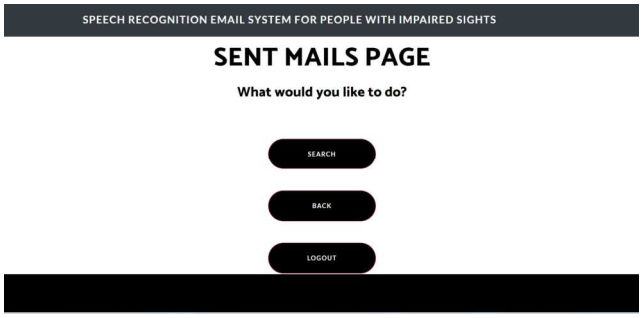

Fig 8 – GUI of Sent Email Page

F. Trash Page:

The trash page in the voiced-based email system is where all deleted emails are stored. The user can easily access this page to recover any accidentally deleted emails or to permanently delete unwanted emails. The trash page also features a search option to help the user find specific deleted emails based on keywords. This option is particularly helpful when the user is looking for a specific email that they deleted in the past. The search option saves the user time and effort by allowing them to find the email they are looking for quickly. Overall, the trash page provides a convenient and user-friendly interface for the user to manage their deleted emails.

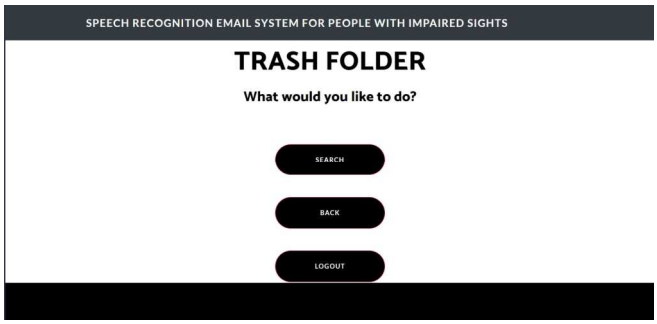

Fig 9 – GUI of Tash Folder Page

VI. USER SATISFACTION SURVEY

To gather feedback from potential users of the voiced-based email system, a survey was conducted using Google Forms. The survey included questions about the user's experience with email management, their familiarity with voice-based technology, and their interest in using a voice-based email system for blind people. The survey received responses from a diverse group of participants, including blind and visually impaired individuals, as well as sighted individuals interested in supporting accessibility efforts.

Below are screenshots of the survey questions and responses:

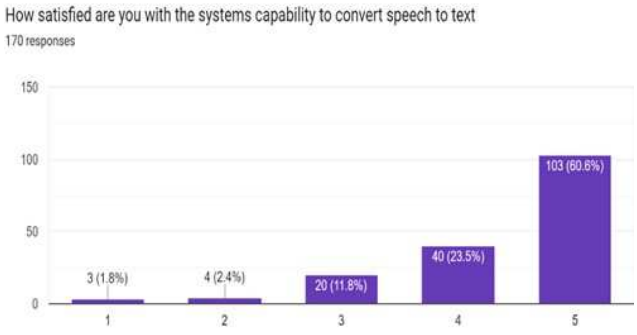

Fig 10 – Survey Result 1

According to the responses received, it can be concluded that a significant percentage of users, approximately 65%, were highly satisfied with the system's capability to convert text to speech, rating it 5 out of 5. Additionally, 21% of the users rated the system 4 out of 5, suggesting that they were also satisfied with the system's performance.

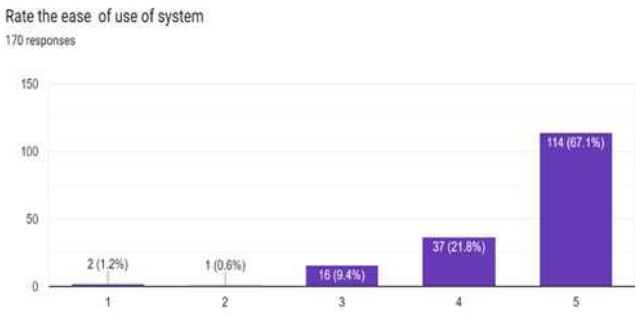

Fig 12 - Survey Result 3

The responses received suggest that a majority of users, around 67%, rated the ease of use of the system 5 out of 5, indicating that they found the system to be very easy to use.

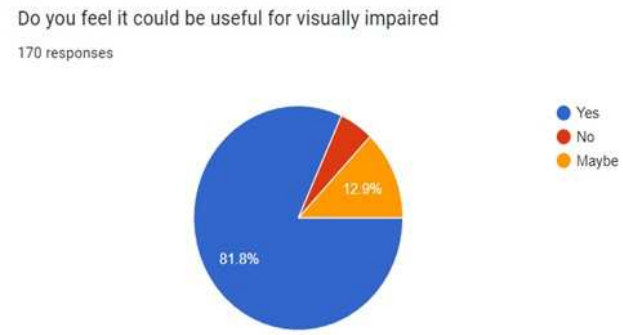

Fig 13 – Survey Result 4

Based on the responses received, it can be inferred that a significant percentage of users, approximately 81.8%, believed that the system could be useful for visually impaired individuals. This suggests that the system has the potential to be beneficial for a broad range of user

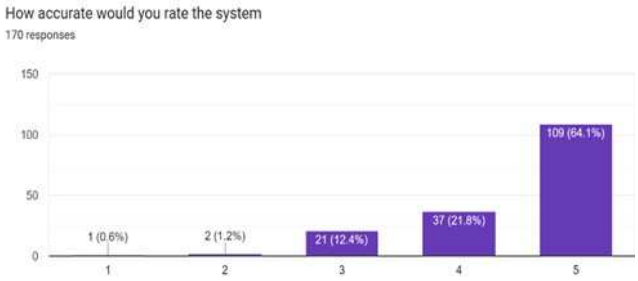

Fig 14 – Survey Result 5

According to the responses received, it can be concluded that a majority of users, approximately 64%, rated the system's accuracy as 5 out of 5, indicating that they found the system to be highly accurate. Additionally, 21% of users rated the system 4 out of 5, suggesting that they found the system to be reasonably accurate.

VII. CONCLUSION AND FUTURE SCOPE

A voice-based email system designed for visually impaired individuals has the potential to revolutionize the way people with disabilities communicate. With advancements in natural language processing, machine learning, and artificial intelligence, the technology can accurately transcribe spoken words into written text and vice versa, making email communication more accessible and inclusive.

Furthermore, the future scope of voice-based email systems is promising. As the technology advances, we can expect more features to be added to the system, such as voice recognition and intelligent assistance, which will make the system more user-friendly and efficient. The integration of other assistive technologies, such as screen readers and Braille displays, can further enhance the accessibility of the system. The benefits of a voice-based email system for visually impaired individuals are enormous. It allows them to communicate more efficiently, thereby improving their productivity, reducing their dependence on others, and enhancing their overall quality of life. The system provides equal opportunities and access to visually impaired individuals, empowering them to participate more fully in social and professional interactions.

REFERENCES

[1] Sherly Noel."Human computer interaction (HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU)". In 2020 Third International Conference on Smart Systems and Inventive Technology (ICCSIT) .

[2] Paulus A. Tiwari; Pratiksha Zodawan; Harsha P. Nimkar; Trishna Rotke; Priya G. Wanjari; Umesh Samarth, "A Review on Voice based EMail System for Blind". In the 2020 International Conference on Inventive Computation Technologies (ICICT)

[3] Jagtap Nilesh, Pawan Alai, Chavhan Swapnil and Bendre M.R.. "Voice Based System in Desktop and Mobile Devices for Blind People". In International Journal of Emerging Technology and Advanced Engineering (IJETAE), 2014.

[4] Ravichandran, R., & Kumar, S. S. (2016). "Voice-Based Email System for Blind and Visually Impaired People". International Journal of Emerging Trends & Technology in Computer Science, 5(5), 255-259.

[5] Nivedita Bhore, Shraddha Mahala "Email System for Visually Impaired People", In International Journal of Engineering Research and Technology (IJERT), 2021.

[6] SpeechRecognition library: Huang, A., et al. (2021). SpeechRecognition (Version 3.8.1) [Computer software]. Retrieved from https://pypi.org/project/SpeechRecognition/

[7] Google Text-to-Speech API: Google LLC. (n.d.). Text-to- Speech API. Retrieved April 29, 2023, from https://cloud.google.com/text-to-speech

[8] Playsound library: Bastian Bechtold (2020). playsound (Version 1.2.2) [Computer software]. Retrieved from https://pypi.org/project/playsound/