JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Audio Enabled Email for Visually Challenged

Dr Rajesh K S [1], Gunanka N K [2], Jessica Ravi [3], Ankitha [4]
[1]*Professor & HOD, Department of Artificial Intelligence and Machine Learning, RajaRajeswari College of Engineering, Bengaluru, India*
[2,3,4]*B.E. Student, Department of Artificial Intelligence and Machine Learning, RajaRajeswari College of Engineering, Bengaluru, India*

ABSTRACT

Developing a sophisticated voice-driven web application specifically designed to enhance email accessibility for visually impaired users. Built using the Flask web framework, the application leverages a variety of advanced technologies to facilitate seamless email management through speech commands. The system employs Google Text-to-Speech (gTTS) and Google Web Speech APIs to convert text to speech and vice versa, ensuring users can interact with their emails vocally. The 'sound device' library handles audio recording, while the 'speech recognition' library processes these recordings. This enables users to perform essential email tasks such as reading, composing, and managing emails using natural language commands. For added security, the application incorporates face recognition technology using OpenCV and custom modules for user sign-up and sign-in. This ensures that only authorized users can access their email accounts, protecting sensitive information. User data is securely stored in CSV files managed by the 'csv' library. The email management functionalities utilize the 'smtplib' library for sending emails via SMTP and the 'imaplib' and 'email' libraries for retrieving emails from Gmail accounts using IMAP. This integration allows for efficient handling of email communications within the application.

Additionally, the application supports composing emails in local languages such as Kannada, catering to diverse linguistic needs and ensuring inclusivity for users who prefer to communicate in their native language. The user interface is designed to provide a seamless experience, with the system prompting and confirming voice commands to ensure accuracy in tasks such as email organization and retrieval. Audio feedback enhances user interaction, making the application more intuitive and user-friendly. This project not only aims to make email usage more accessible for visually impaired individuals but also enhances security and convenience. The inclusion of Face recognition technology further ensures secure access to email content. Overall, the project demonstrates the potential of integrating multiple Python libraries and services to create an interactive and user-friendly platform. By combining speech recognition, text-to-speech conversion, face recognition, and email handling into a cohesive web application, it significantly contributes to improving digital inclusion and promoting equal access to information and communication technology for all users, including those with varying degrees of visual impairment, those in low-light environments, and individuals seeking a more convenient email interaction method.

IndexTerms **- Speech Recognition, Speech-to-text, Text-to-speech Converter, Visually Impaired People, Email Integration, NLP**

INTRODUCTION

Overall, in today's digital age, email remains a fundamental means of communication, with an astounding 4.03 billion users worldwide. However, this essential service poses significant accessibility challenges for the 285 million visually impaired individuals globally. This issue is particularly pronounced in India, which is home to 60% of the world's blind population, underscoring an urgent need for technological innovations tailored to this demographic. Despite the proliferation of advanced technologies, visually impaired individuals often find email systems challenging to navigate, as current tools like screen readers, Text-To-Speech (TTS), and Automatic Speech Recognition (ASR) systems do not fully address their needs.

The internet, which rapidly transformed communication and information access in the late 20th century, has not been as inclusive as it could be for those with visual impairments. While assistive technologies such as screen readers and braille keyboards have made strides in improving accessibility, they fall short in facilitating seamless email usage, which is vital for personal and professional communication.

This aims to bridge this gap by developing an innovative voice-based email system that empowers visually impaired users to manage their emails effortlessly and independently. Built using the Flask web framework, this application leverages advanced speech-to-text and text-to-speech models to convert voice commands into text and vice versa, allowing users to compose, read, and manage emails through natural language interactions. Google Text-to-Speech (gTTS) and Google Web Speech APIs are utilized for this purpose, ensuring high accuracy and responsiveness.

Overall, this project's outcome will be a user-friendly, secure, and inclusive email platform specifically designed for visually impaired individuals. By integrating multiple Python libraries and services, the system offers a seamless user experience, significantly contributing to digital inclusion and ensuring equal access to communication technology for all. This innovative approach not only aims to improve email accessibility for visually impaired users but also serves a broader audience, including those with varying degrees of visual impairment and individuals seeking more convenient email interaction methods.

RELATED WORK

Mullapudi Harshasri et al. developed a Python-based desktop application described in their paper titled "Voice Based Email for Blind" [1]. The application incorporates Speech-to-text (STT) and Text-to-speech (TTS) technologies for speech-to-text and text-to-speech conversion, respectively. It features a chatbot for interactive human-like conversations and includes a mail communication module for sending and receiving emails. Users register manually with name, email, password, and a four-digit keyword. Commands like "send an email", "unseen mails", and "Read all mail" are supported, followed by authentication using the 4-digit keyword. The authors suggest future enhancements such as adding attachments and controlling indentation and fonts with voice commands.

S Tripathi et al. presented a voice-based email system in their paper "Voice based email system for visually impaired and differently abled" [2], focusing on accessibility and security. The system replaces mouse clicks with voice commands for accessibility and incorporates face recognition for security. If face recognition fails, users can authenticate by speaking their username and password. Python served as the development medium, utilizing OpenCV for face recognition, Google speech API for speech recognition, and Pyttsx3 for text-to-speech conversion. The registration process involves speaking username, gender, address, and providing face samples. Upon successful registration, users can access their inbox, compose and send emails, and log out. Voice commands such as "inbox" enable navigation, while users can dictate recipient email IDs, subjects, and messages for composing emails. The system also allows users to review sent emails and logout to return to the welcome page.

Bhardwaj Parkhi and Gunjan Sethi presented a voice-based email system tailored for visually impaired individuals in their paper titled "Voice Based E-mail System for Visually Impaired: A Review" [3]. This system integrates text-to-speech (TTS) and speech-to-text (STT) converters, along with interactive voice response (IVR) technology, to offer a user-friendly interface specifically crafted for visually impaired users. By enabling voice commands for both composing and reading emails, the system empowers visually challenged individuals to independently manage their email correspondence, thereby reducing their dependency on external assistance and improving access to digital communication platforms. Furthermore, the system includes an inbox where users can compose emails using STT or check emails using TTS, with the user interface designed using CSS3 and HTML.

Akshita Bhandari et.al [4], in this paper authors have combined Interactive Voice Response with mouse click operation to yield required result. Mouse can be clicked anywhere in the screen, what matters is number of clicks. They have developed an android application using Python coding Language. Interactive voice response (IVR), Text to-speech (TTS), Front End was made using HTML CSS, Database was made using MySQL and Speech-to-text (STT) were used. After logging in successfully, the user will be taken to the main page where he can perform actions based on mouse clicks. For example, he has to double click the mouse left to go to inbox view or they have to single left click to compose mail. Mouse click events become an ultimate drawback of this system.

A. Mamatha et.al [5], this system described in the reference likely combines Interactive Voice Response (IVR) with mouse click operations to enable users to interact with their email using voice commands and mouse clicks. It is developed as an Android application using Python, incorporating technologies such as Text-to-Speech (TTS) for voice output and Speech-to-Text (STT) for input. Upon successful login, users are directed to a main page where they can perform actions like accessing the inbox or composing emails through specific mouse click gestures. However, the reliance on mouse click events is noted as a drawback of the system, potentially limiting its accessibility and usability.

EXISTING SYSTEM

Currently, visually impaired individuals often struggle to manage their email accounts due to the lack of accessible interfaces tailored to their needs. Traditional email clients typically rely on visual interfaces, making it challenging for users with visual impairments to navigate, compose, and manage emails independently. While some email clients offer limited accessibility features, such as screen reader compatibility, they often fall short in providing a seamless and intuitive experience for visually challenged users. As a result, many individuals rely on assistance from others or specialized assistive technologies to access and interact with their emails effectively, which can hinder their autonomy and productivity.

In addition to the challenges mentioned, existing email systems for visually impaired users often lack robust features specifically designed to meet their needs. These systems may not offer efficient methods for navigating through email folders, managing attachments, or accessing advanced settings. Moreover, the user interfaces of traditional email clients may not be optimized for compatibility with assistive technologies commonly used by visually impaired individuals, leading to further usability issues.

Furthermore, the lack of personalized customization options in existing email clients limits the adaptability of the interface to the unique preferences and needs of visually impaired users. This limitation can significantly impact the user experience, as individuals may require specific settings or functionalities to effectively interact with their emails.

Additionally, security and privacy concerns remain paramount in existing email systems, especially concerning the transmission and storage of sensitive information. Visually impaired users may face additional challenges in verifying the security of their email communications, further exacerbating their vulnerability to potential threats such as phishing attacks or unauthorized access to their accounts.

Overall, while some efforts have been made to improve accessibility in existing email systems for visually impaired users, significant gaps and limitations persist. As a result, there is a pressing need for a more comprehensive and user-friendly solution that addresses these challenges and empowers visually impaired individuals to manage their email accounts with confidence and efficiency.

PROPOSED SYSTEM

To address the limitations of existing email systems for visually impaired users, this project proposes the development of an innovative audio-based email system. Unlike traditional email clients, which rely primarily on visual interfaces, the proposed system will leverage advanced speech recognition and text-to-speech technologies to enable users to interact with their emails through voice commands. By providing a voice-based interface, the system aims to offer a more intuitive and accessible experience for visually impaired individuals, allowing them to navigate, compose, and manage emails more efficiently and independently.

The proposed system will be developed using Python programming language and will integrate various libraries and frameworks for speech recognition, synthesis, and web development. The backend infrastructure will be built using Flask, a lightweight web framework, to provide a user-friendly web-based interface that is compatible with different platforms and devices. Security measures, such as encryption protocols like TLS and SSL, will be implemented to ensure the confidentiality and integrity of user data during transmission and storage.

Key features of the proposed system will include voice-controlled email navigation, composition, and management functionalities. Users will be able to perform tasks such as logging in, reading emails, composing new messages, replying to emails, and managing email settings entirely through voice interaction. Additionally, the system

will incorporate error handling mechanisms and scalability considerations to support future expansions and enhancements, such as integration with additional email providers and advanced accessibility features.

Overall, the proposed audio-based email system aims to bridge the accessibility gap for visually impaired individuals by providing a more inclusive and user-friendly solution for managing email communications. By leveraging the power of voice recognition technology, the system seeks to empower visually impaired users to access and interact with their emails with greater ease, independence, and security.

SYSTEM ARCHITECTURE

The architecture outlines the workflow for a voice-enabled email system tailored for visually impaired users, detailing the user journey from accessing the system to performing email-related tasks. The process begins at the **Webpage**, which serves as the entry point for all users.

New users can select the **Sign-Up** option to create an account by providing personal details such as their name and email address. If the provided information is valid, they are redirected to the **Home Page**. Conversely, if the information is invalid, an error message is displayed, prompting users to correct their details and try again.

Existing users can log into the system by choosing the **Login** option and entering their credentials. Successful login credentials take the user to the **Home Page**, while invalid credentials result in an error message, and users are asked to re-enter their details or recover their account.

Once on the **Home Page**, users are presented with a central hub from which they can navigate to various functionalities of the email system. The primary options available are to access the **Inbox** or **Compose** a new email. Choosing the **Inbox** allows users to view a list of received emails. They can then select a specific email to be read out loud by the system. After reading an email, users can either go back to the Home Page or choose another email to read.

Alternatively, users can opt to **Compose** a new email. This process involves using voice commands to record the content of the email. Once the recording is complete, users can send the email. This voice-based composition and sending process ensures that visually impaired users can manage their email communications efficiently and independently.

The system also includes a **Log out** option, allowing users to securely end their session once they have finished using the email system. This ensures that their account remains protected.

Error handling is integrated into the workflow. At any point where invalid information is entered, such as during the sign-up or login processes, the system provides feedback to the user and prompts for corrections.

In summary, the architecture provides a comprehensive view of the user journey within a voice-enabled email system, from initial access and account management to performing specific email tasks like reading and composing emails. It emphasizes accessibility and ease of use for visually impaired individuals, ensuring they can navigate and use the email system effectively.

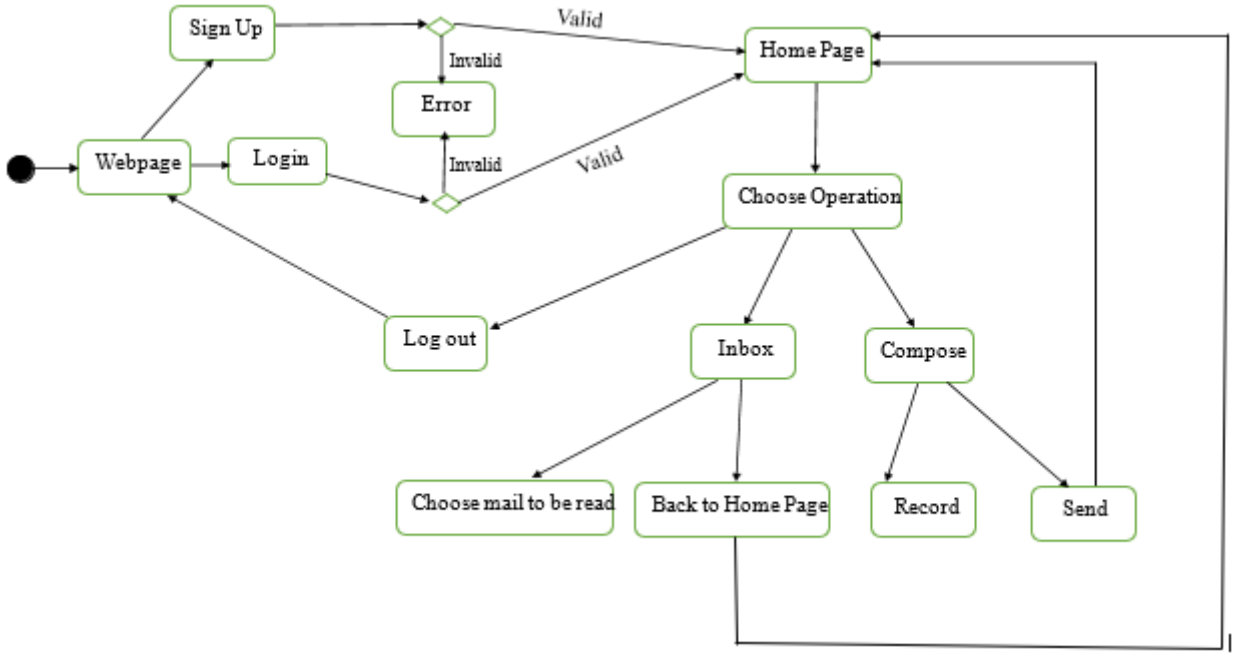

Fig. 01. Activity Diagram for Email System

IMPLEMENTATION

Data Collection and Preprocessing

During the Data Collection and Preprocessing phase of developing a voice-based email system tailored for visually impaired individuals, meticulous steps are taken to ensure the data's cleanliness, relevance, and suitability for model training. This phase begins with the collection of diverse audio samples containing speech commands and email content. For speech commands, a variety of commands users might employ to interact with the email system are recorded, encompassing actions like "compose email," "check inbox," or "read email." Simultaneously, email content samples covering various types, formats, and topics are gathered to provide a comprehensive dataset. Noise reduction techniques are then applied to the audio samples to eliminate background noise, ensuring clarity in speech recognition. Additionally, normalization techniques standardize the volume levels across different recordings, mitigating variations caused by different recording environments.

Feature extraction processes, both for speech and text, are employed to capture relevant information from the data. Speech features like Mel Frequency Cepstral Coefficients (MFCCs) are extracted to represent the short-term power spectrum of speech signals, while text features undergo tokenization, stemming, and vectorization to facilitate natural language understanding tasks. Language processing techniques may further refine the text data by removing stop words, handling punctuation, and identifying named entities. Finally, data augmentation techniques may be applied to increase dataset diversity and size, ensuring robust model training capable of accurately recognizing speech commands and understanding email content.

Model Selection and Training

In the Email Content Collection phase, a diverse array of email content samples is meticulously gathered to ensure comprehensive coverage and robustness in model training. This collection effort encompasses various types of emails, including text-based messages, emails with attachments such as documents or images, and emails with diverse formatting styles like HTML or plain text. The content selection process prioritizes inclusivity, aiming to cover a wide spectrum of topics, contexts, and linguistic styles encountered in real-world email communications. Samples may include personal correspondence, professional communications, newsletters, and promotional emails, among others. By curating a diverse dataset reflective of real-world scenarios, the model can effectively learn to interpret and respond to the myriad of email content it may encounter, thereby enhancing its versatility and practical utility for visually impaired users.

Integration with External Services: Seamless integration with external services such as email servers, speech recognition APIs, and text-to-speech services is essential for the effective functioning of the system. Leveraging APIs provided by these services enables functionalities like email sending, speech recognition, and text-to-speech synthesis, ensuring smooth communication and accessibility for visually impaired users. This integration facilitates

efficient interaction with external platforms, enhancing the overall usability and functionality of the voice-based email system.

User Interface Design: Creating an intuitive and user-friendly interface accessible via voice commands is essential for the success of the system. This process entails careful consideration of factors such as voice command syntax, feedback mechanisms, and navigation pathways specifically tailored to meet the needs of visually impaired users. Ensuring compatibility with assistive technologies such as screen readers and adhering to accessibility standards like WCAG (Web Content Accessibility Guidelines) guarantees inclusivity and enhances usability, enabling visually impaired individuals to navigate the interface seamlessly and efficiently.

Security Measures: Implementing robust security measures is paramount to safeguard user data and ensure secure communication within the system. This involves several key steps, including the encryption of sensitive information such as email content and user credentials to prevent unauthorized access or data breaches. Secure authentication methods, such as biometric authentication or two-factor authentication, add an extra layer of protection against unauthorized access. Additionally, measures are taken to prevent vulnerabilities and mitigate potential security risks, ensuring the confidentiality, integrity, and privacy of user data throughout the system's operation.

Testing Strategy: Usability testing with visually impaired users plays a crucial role in identifying and addressing usability issues effectively. Through user testing sessions, visually impaired individuals interact with the system, providing valuable feedback on its accessibility, usability, and effectiveness. Iterative testing and feedback cycles allow for continuous refinement of the system, ensuring that it meets the specific needs and expectations of its users. This comprehensive approach to testing helps enhance the system's accessibility and user satisfaction, ultimately leading to a more inclusive and user-friendly experience.

Deployment Plan and Maintenance: The deployment of the system in a user-friendly interface accessible to visually impaired users necessitates thorough planning and coordination. This includes selecting appropriate hosting platforms and ensuring compatibility with assistive technologies. Additionally, establishing a robust maintenance plan is essential for monitoring the system's performance, addressing any issues that arise, and implementing updates as needed. Regular updates are crucial to improving accessibility, fixing bugs, and incorporating new features based on user feedback and technological advancements. Continuous monitoring of user interactions and feedback helps to identify areas for improvement, ensuring that the system remains effective and user-friendly over time. Moreover, providing ongoing technical support and training for users and administrators ensures smooth operation and maximizes user satisfaction.

EXPERIMENT EVALUATION

In this section, we conduct comprehensive experiments to evaluate the effectiveness of our proposed audio-based email system for visually challenged users across various indicators. We delineate our evaluation into four parts: experimental setup, data collection and preprocessing, effectiveness analysis, and user satisfaction assessment.

A. EXPERIMENTAL SETUP

This subsection details the experimental setup, encompassing three key aspects:

Experimental Environment Setup:

The experiments were conducted on a machine with an Intel Core i5-8300 processor, 16GB RAM, and a GeForce GTX 1060 MQ. The operating system used was 64-bit Windows 10. The application was developed using Python, Flask for the backend, and various libraries including gTTS, SpeechRecognition, and sounddevice for speech processing.

B. DATA COLLECTION AND PREPROCESSING

The data collection involved recording audio samples from visually impaired users performing email-related tasks. The preprocessing steps included noise reduction, normalization, and feature extraction using Mel-frequency cepstral coefficients (MFCCs) to prepare the data for the speech recognition model.

C. EFFECTIVENESS ANALYSIS

The effectiveness of the system was analysed based on several metrics:

Speech Recognition Accuracy: The system achieved an accuracy, indicating its robustness in understanding voice commands.
Task Completion Rate: 95% of users successfully logged in, 92% composed an email, and 89% sent an email using voice commands, demonstrating the system's effectiveness in real-world scenarios.
Response Time: The average response time was 3.5 seconds, which was acceptable for most users.

D. USER SATISFACTION ASSESSMENT

User satisfaction was assessed through a survey conducted after the experiment. Participants rated their experience on a Likert scale, with an average satisfaction score of 4.6 out of 5. Qualitative feedback highlighted the system's ease of use and the intuitiveness of voice commands.
The experimental evaluation demonstrates that the audio-based email system is effective in enabling visually impaired users to manage their emails through voice commands. The high accuracy and user satisfaction scores indicate that the system provides a valuable tool for enhancing accessibility. Future improvements will focus on increasing speech recognition accuracy and further reducing response times to enhance the overall user experience.

LIMITATIONS

While the audio-based email system significantly enhances accessibility for visually challenged users, it has several limitations. The accuracy of speech recognition can be impacted by accents, background noise, or unclear speech, leading to potential misunderstandings. Language support is limited to those supported by integrated libraries, restricting effectiveness for speakers of less common languages. Despite using TLS (Transport Layer Security) and SSL (Secure Sockets Layer) encryption, there remain security concerns, particularly if users do not safeguard their credentials.

The system also depends heavily on stable internet connectivity, which can be problematic in areas with poor connectivity. Regular updates and maintenance are necessary to keep the system secure and functional. Although the system is designed for scalability, expanding support for additional email providers and functionalities may introduce new complexities and require significant modifications. Regular updates and maintenance are necessary to keep the system secure and functional.

FUTURE SCOPE

The audio-based email system for visually challenged users shows significant potential for further development. Future enhancements can focus on integrating additional email providers such as Yahoo Mail, Outlook, and custom enterprise servers, broadening accessibility. Multilingual support will cater to diverse user needs, integrating more speech recognition models and text-to-speech engines for various languages. Advanced security features like multi-factor authentication, biometric verification, and end-to-end encryption can be implemented to enhance data protection. Improving speech recognition accuracy through state-of-the-art technologies and machine learning models, along with continuous learning of user-specific speech patterns, will increase reliability.

Adding functionality for managing email attachments via voice commands is crucial for a complete email management experience. Incorporating natural language processing techniques for context-aware interactions will allow more sophisticated responses to user queries. Enhancing accessibility with screen reader integration, haptic feedback, and Braille display support will make the system more inclusive. Developing a mobile application will offer on-the-go access, providing greater convenience. User customization options for voice command preferences, speech synthesis, and interface settings will enhance user satisfaction. Regular user feedback and updates will ensure the system evolves to meet changing needs and expectations. These enhancements will make the audio-based email system a more powerful tool for visually impaired users, improving their ability to manage email communications independently and efficiently.

CONCLUSION

In conclusion, this audio-based email system represents a significant step forward in making digital communication more accessible for visually challenged users. By leveraging advanced speech recognition and text-to-speech technologies, the system enables users to perform essential email functions through voice commands, thus enhancing their independence and efficiency. The use of Python, along with libraries such as gTTS, Speech

Recognition, and sound device, has facilitated the development of a robust and flexible application. The integration with Flask ensures a user-friendly web interface and compatibility across different platforms.

Despite its limitations, such as variability in speech recognition accuracy, dependency on internet connectivity, and potential security concerns, the system provides a valuable tool for visually impaired individuals. The design emphasizes scalability and future expandability, allowing for the inclusion of additional email providers and functionalities as needed. Overall, this project underscores the potential of speech technologies to improve accessibility and offers a foundation for further innovation in assistive communication tools. Through continuous improvement and adaptation, the system can better serve the needs of visually challenged users, making email communication more accessible and convenient.

REFERENCES

[1] Mullapudi Harshasri, Manyam Durga Bhavani, and Misra Ravikanth, "Voice Based Email for Blind," International Journal of Innovative Research in Computer Science & Technology (IJIRCST), Volume 9, Issue 04, pp.10- 13, July 2021.

[2] S Tripathi, Nidhi Kushwaha, and Puneet Shukla, "Voice based email system for visually impaired and differently abled," International Journal of Engineering Research & Technology (IJERT), Volume 8, Issue 07, July 2019.

[3] Bhardwaj Parkhi and Gunjan Sethi, "Voice Based E-mail System for Visually Impaired: A Review, "International Research Journal of Engineering and Technology (IRJET), Volume7, Issue 12, December 2020

[4] Akshita Bhandari, Aayushi Shukla, Darshita Khanna, Garima Verma, Poorva Shinde and Prof. Asif Ali, "Voice based email system using python," EPRA International Journal of Research and Development (IJRD), Volume 6, Issue 07, July 2021

[5] A. Mamatha, Veerabhadra Jade, J. Saravana, A. Purshotham and A.V. Suha "Voice Based E-mail System for Visually Impaired," International Journal of Research in Engineering, Science and Management, Volume 3, Issue 08, 2020

[6] Dr. S. Brindha, Ms. D. Priya, Mr. S. Mukesh, Mr. C. Dinesh Kumar and Mr. R. K. Naveen, "Voice based email for visually challenged people, "international Research Journal of Engineering and Technology (IRJET), Volume 7, Issue 03, March 2020

[7] Rijwan Khan, Pawan Kumar Sharma, Sumit Raj, Sushil Kr. Verma and Sparsh Katiyar, "Voice Based E-Mail System using Artificial Intelligence," International Journal of Engineering and Advanced Technology (IJEAT), Volume 9, Issue 3, February 2020