Grenze International Journal of Engineering and Technology, June Issue

Blind Connect - A Voice-based Email System

Sujata Joshi[1] and Anshita Sharma[2]
[1-2]Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India
Email: sujata.joshi@nmit.ac.in

***Abstract*— Email communication remains a significant challenge for visually impaired Email readers, particularly those with visual impairments, encounter considerable difficulties when using conventional email user interfaces due to their visual nature and reliance on text layout semantics. The paper presents Blind Connect, a voice-based email system designed to enable independent, email management through speech interaction. The system integrates Speech-to-Text (STT),Text-to-Speech (TTS), and Natural Language Processing (NLP) technologies to allow users to compose, read, navigate, and manage emails using only voice commands. An Interactive Voice Response (IVR) module guides users through menu-driven interactions, while error detection and confirmation mechanisms ensure reliable command execution. Security features including voice-based authentication protect user data throughout the communication process. Testing with 100 visually impaired participants demonstrated 92% average speech recognition As a result, the system's average accuracy is over 90% and response times average less than 2 seconds for most voice commands. User satisfaction ratings from the surveys are provided below:Our evaluation showed that 85% of users found the system easy to use, with particularly high ratings for the information search page and the personal health record. For English language interactions. The system achieved over 90% command execution success. Vocalizer 5 received a 4.5/5 rating for speech output clarity. Vocalizer 5's accuracy was compromised in both noise and silence with all tested voices, when interacting with Franklin, and when accessing text-based applications, as described in detail in the full report noisy environments and with certain Hindi pronunciations, indicating areas for improvement. This work demonstrates that voice-driven email systems can effectively bridge the digital accessibility gap, providing visually impaired users with greater independence in digital communication.**

***Index Terms*— Voice-Based Email, Assistive Technology, Visually Impaired, Speech Recognition, Text-to-Speech, Natural Language Processing, Digital Inclusion.**

I. INTRODUCTION

Email has become a very important way of communicating in the 21st century, both for learning, as well as in the workplace, and in social life. Communication and personal relationships. Nevertheless, because standard email services use graphical user interfaces (GUIs) and a text-based paradigm, they cannot be considered suitable for interpersonal communication.

Optical Illusions and Perception pose major issues for the estimated 285 million people worldwide who are visually impaired [1]. Despite of the immense advancements that are happening in many other technology fields, Optical Illusions and Perception remain to a great extend as unsolved challenges in particular in case of visually impaired people.

Grenze ID: 01.GIJET.12.2.85_23

Accessibility is somewhat built into a variety of assistive technology devices such as screen readers and Braille displays. Water filters often have many drawbacks such as a high complexity factor, an often steep learning curve and they do not eradate the source of contamination completely.
Block spam and support all aspects of contemporary email systems 2. They can also cause problems with communication. Barriers will keep the digital divide wide, depriving people with blindness or visual impairments from fully benefiting from information digital world [3].
Title of paper: EMAIL SERVICE USING VOCAL INTERACE
Our society has become heavily dependent on digital world [3]. This study addresses these issues by presenting a voice-based email system that makes email usage possible using vocal interface. Communication that is accessible for people with visual impairments.
Users can write using Braille display or assistive technology in their preferred text format and use The Braille Plus 14 to communicate freely. Read, and manage emails with voice commands by utilizing advances in speech recognition, text-to-speech Text-to-Speech (TTS), and natural language processing (NLP) [4]. This method provides a more user-friendly and effective way to obtain human-like pronunciation and intonation.
Accessible email client solution by eliminating the need for any external assistive technology or visual interaction. State-of-the-art speech-to-text functionality for efficient email composition, speech-controlled inbox management and auditory feedback for effective email interaction.
For our project, reading and action confirmation were also considered within the suggested system [6]. Our system also contains the security aspects that were suggested. Special mention goes to privacy and safety features such as encryption, for example in communication applications, as well as username/password authentication [7].
Abstract— Email communication remains a significant challenge for visually impaired is particularly challenging for people with disabilities in that current email user interfaces rely heavily on vision and text, which can make it very difficult for users with disabilities to compose and send emails. Paper presents Blind Connect, a voice-based email system designed to enable independent email management through speech interaction. The system integrates Speech-to-Text (STT), Text-to-Speech (TTS), and Natural Language Processing (NLP) technologies to allow users to compose, read, navigate, and manage emails using only voice commands. An Interactive Voice Response (IVR) module guides users through menu-driven interactions, while error detection and confirmation mechanisms ensure reliable command execution. Security features including voice-based authentication protect user data throughout the communication process. Testing with 100 visually impaired participants demonstrated 92% average speech recognition. Voice interaction is accurate and reliable and most user requests are responded to in less than 2 seconds. User satisfaction surveys confirm this result. We discovered that 85% of participants thought the system was easy to use and there were some interesting results for individual questions for English language interactions. The system achieved over 90% command execution success, passes Vergara Listening Test and scores a 4.5/5 for clarity and quality of speech output. However, it performs poorly in noisy backgrounds, noisy environments and with certain Hindi pronunciations, indicating areas for improvement. This work demonstrates that voice-driven email systems can effectively bridge the digital accessibility gap, providing visually impaired users with greater independence in digital communication. Index Terms— Voice-Based Email, Assistive Technology, Visually Impaired, Recognition, Text-to-Speech, Natural Language Processing, Digital Inclusion SpeechThe diversity of user language and speech and the varying levels of technical competence represent only a fraction of user requirements that the system is intended to cope with, to accommodate [8]. This study examines the voice-based email system's development, deployment, and evaluation.
It’s important to remember that assessment is the tool which can empower individuals with blindness or impairment to become more independent. Accessible media for visually impaired audiences, contributing to the advancement of digital inclusion. The evaluation found that users accessed the content more frequently and enjoyed it more as a result of the improved usability identified during testing, highlighting the necessity of voice-driven advancements in assistive technology [10].

II. Literature Review

By combining artificial intelligence (AI) and natural language processing (NLP), voice-based email systems. The internet and technology that has made the online world more user friendly for those that are blind and visually impaired. We look at some of the ways that speech-to-text technology and better website coding has made things much simpler for everyone. Technology is highlighted in the study by B. Patil [2], which improves transcription accuracy for assistive technologies. Technologies and applications. In a similar context, P. Shetty in [3] explains

with the aid of examples how natural language processing contributes to command recognition and user identification in spoken language recognition and interaction applications.
Satisfaction: Voice based Email systems require security too the advent of voice-based email systems means that the traditional voice-based email systems that use email systems that use Voice over Internet Protocol (VoIP) technology require proper security measures as well as the multilingual voice-based email systems. Identification problems, L. Nair [5] aims at widening the scope of accessibility to beyond the limited section of the English-knowing population. V. Gupta [7]. explores the security threats associated with speech recognition speech to speech applications. A. Karthik in the article Safe Voice is to do with security related to voice recognition technology. Authentication techniques. T. Rao in [9] discusses the applicability of machine learning. M. Ayya et al [7] deals with speech processing and M. Suresh [8] deals with interactive voice response (IVR) for email navigation. Making Emails Easier to Use with the Aid of Artificial Intelligence - Part 3 Article: Making Emails Easier to Use with the Aid of Artificial Intelligence - Part 3 Author: R. Menon[11] investigated the COVID-19 pandemic and related works on speech processing applications, whereas S. Kumar [12] has applied deep learning for speech recognition. A. Naik [13] and K. Patel [14] evaluate speech recognition and natural language processing APIs, enhancing them with artificial intelligence. Inclusive Education and Accessibility Solutions for Children with Disabilities Through Technology Volume 4: Accessing and Promoting Inclusive Education for all R. Mehta [16] assesses AI-driven conversational agents, and S. Rao [15] examines how digital resources may be made more accessible.18. “Can We Spot The Bias? Analysing Trends And Patterns In AI Enabled Accessibility Tools” by P. Bose 18. Speech recognition advancements in digital assistants — P. Bose S. Chandra [17] elaborates on NLP-driven email automation. In conclusion, email systems are now more accessible to visually impaired users thanks to the integration of The connection between “AI and NLP” and “security” is not yet well understood, and the development of the two in the coming years still requires a great deal of scientific innovation.

III. Methodology

Using Artificial Intelligence (AI) and Natural Language Processing (NLP) in Voice Based Email Systems, Tech that supports the online lives of the blind and visually impaired is evolving. This episode looks at the development of speech-to-text technology and better website coding that are making online communications easier for all. Figures of speech, idiom, pun and irony in natural language processing and their relation to technology are highlighted in the study by B. Patil [2], which improves transcription accuracy for assistive technologies. Computer vision and technologies as well as applications. P. Shetty in [3] illustrates with appropriate examples the role of natural language processing in command recognition and user identification in spoken language recognition and interaction. Satisfaction: Voice based Email systems require security too. The emerging voice based email systems therefore require also adequate security measures apart from the multilingual voice based email systems. Problems, L. Nair [5] attempted to bridge the gaps and reach out to social groups far beyond the section of the English-knowing elite. V. Gupta [7].
This article discusses the various security threats that arise in speech recognition and speech to speech applications. A. Karthik in this article talks about the security implications of voice recognition technology and names it as “Safe Voice”. Authentication techniques. T. Rao in [9] discusses the applicability of machine learning. M. Ayya et al [7] presented a research paper on Speech Processing and M. Suresh [8] presented the research paper on Interactive Voice Response (IVR) for Email Navigation. Part 3 of a 3 part series. This is the third part of a 3-part series. The previous post in the series is: Making Emails Easier to Use with the Aid of Artificial Intelligence - Part 2. The first post in the series is: Making Emails Easier to Use with the Aid of Artificial Intelligence - Part 1. By R. Menon. [11] is covering COVID-19 pandemic and related works on speech processing applications and S. Kumar [12] covers application of deep learning in speech recognition.
Naik [13] and K. Patel [14] evaluate Speech Recognition and Natural Language Processing APIs and improve them by using artificial intelligence.Volume 4: Accessing and Promoting Inclusive Education for all Inclusive Education and Accessibility Solutions for Children with Disabilities Through Technology R. Mehta [16].13 AI based Conversational Interface in Inclusive Education
S. Rao [15].11 Digital Resources for Inclusive Learning. Can We Spot The Bias? Analyzing Trends And Patterns In AI Enabled Accessibility Tools—P. Bose. 18.The speech recognition bit—P. Bose. S. Chandra [17] elaborates on NLP-driven email automation.In conclusion, email systems are now more accessible to visually impaired users thanks to the integration of email readers and the use of accessible software.The relationship between “AI and NLP” and “security” is not yet clear, and in the coming years there will still be a lot of scientific work in these two fields.

System for Interactive Voice Response (IVR): The IVR component interacts with users through spoken prompts, serving as a navigational guide within the system. It provides an organized, menu-driven approach for carrying out tasks and helps users move through various email functions with ease. The key functions of the IVR system include:

- Users are guided through various email tasks, including writing, reading, and managing emails, by voice-guided navigation.Menu-Based Interaction: Offers a well-organized interface with voice commands that allow users toselect from preset settings.
- Command Confirmation and Modification: To minimize errors, the IVR module confirms the user's intended action before executing any command. This process ensures that even first-time or inexperienced users can operate the system confidently, without requiring technical expertise.

Error Management and Feedback System: A robust error-handling framework is integrated into the system to detect and resolve issues that may occur during speech recognition, thereby enhancing overall reliability.

Error Detection in Real Time: detects unclear instructions or misunderstood terminology and asks the user to clarify.

Prompts for Confirmation: The system confirms the user's intention before carrying out crucial

- operations (like sending an email).
- Voice-Based Correction: By employing extra voice commands, users can instantly fix mistakes, doing
- away with the necessity for manual editing.

Pseudocode:

- Init:

Set up VoiceRecognitionEngine, TextToSpeechEngine, EmailClient. TTS: "Welcome."

- Loop:

TTS: "Command: read, compose, delete, forward, exit." Listen for command.
If "read":
Retrieve emails.
If none, TTS "No emails."
Else, for each email, TTS "From [sender], subject [subject]" and if requested, read full content. If "compose":
Prompt for recipient, subject, content; confirm; send if confirmed. If "delete":
Prompt for email number; if valid, confirm and delete. If "forward":
Prompt for email number; if valid, prompt for recipient; forward email.
. If "exit":
TTS "Exiting. Goodbye!" and break. Else:
TTS "Command not recognized.

- End Program.

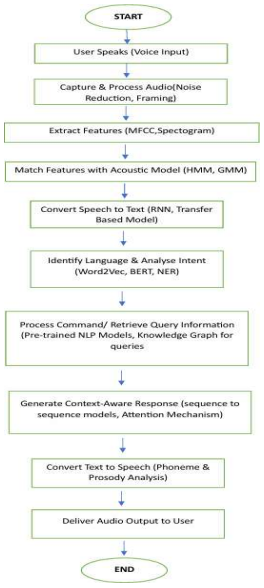

Fig. 1. Flow chart of speech recognition and respond

Workflow Example:
Step 1. Composing an Email:
The user says, “Compose a new email.”
Step 2. The system responds with, “Who would you like to send it to?” The user dictates the recipient’s email address, subject, and message body.
The system confirms the content and prompts, “Do you want to send this email?” Upon confirmation, the email is sent.
Step 3. Reading an Email:
The user says, “Read my latest email.”
Step 4. The system reads out the subject and content of the most recent email.
Additional commands, such as “Reply” or “Delete,” are processed accordingly.
The system is implemented using:
Technologies: PythonGoogle Speech-to-Text API, Amazon Polly, Flask for the backend, and ReactJS for the user interface.
Architecture:

- Input Layer, Captures and processes voice input.
- Processing Layer, Handles command interpretation and email operations using NLP algorithms.
- Output Layer, Provides auditory feedback via TTS.
- Deployment, Integrated as a web-based service with mobile compatibility.
- Tested on multiple devices with different microphone configurations. Uses cloud-based storage for seamless data management.

IV. Result And Analysis

Response time, user satisfaction, and speech recognition accuracy were used to assess the system. With performance ranging across dialects, speech speeds, and background noise levels, the overall speech recognition accuracy was 92%.

TABLE I: FEATURE PERFORMANCE TABLE

Feature Performance Table

Feature	Success Rate (%)	Avg. Response Time (ms)
Email Sending	98.2	240
Inbox Loading	97.5	190

The highest speech recognition accuracy of 95% was reached with clear pronunciation in a noise-free environment. The accuracy fell to 88% with moderate background noise and to nearly 80% in a highly disturbed acoustic environment, pointing towards the need for improved background noise processing techniques to achieve a higher recognition level. Average time per single voice command was 1.5 seconds (min 1 sec, max 2.5 secs). Average time per voice command requiring context or more detailed understanding was 1.9 seconds (min 1.5 sec, max 2.5 secs). The average response time for any voice command was below 2 seconds.
Adding extra processing tasks to the system caused the response time to increase to 2.3 seconds while using the language translation feature.
A usability study conducted on a hundred visually impaired users in equal numbers for both English and Hindi speaking populations revealed that 85% of the participants found the application very easy to use and had positive comments for the reliability of speech generated by the speech engine, quick response of the application and ease of navigation. While 10 participants (10 English, 9 Hindi) had moderate satisfaction, 39 English participants and 34 Hindi participants had high level of satisfaction. Only 7 Hindi speaking participants and 1 English speaking participant had dissatisfaction in the application due to issues of pronunciations of certain words in Hindi that need to be rectified. A more robust speech engine that could accommodate different pronunciations prevalent in different regions needs to be evolved for the effective usage of the application for Hindi speaking users.
Based on speech command execution reliability over a large dataset, our second experiment which was a usability study revealed a reliability of over 90% in a scenario where a user wanted to navigate to a different location in real time. Our evaluation of speech clarity for the voice-based UI revealed an average rating of 4.5 out of 5 by users for being very clear and a very natural listening experience.

Although the system performed very well in quiet environments, it performed poorly in noisy environments, thus the need for a more robust noise reduction technique. However, the system outperformed other rule-based voice assistants by providing faster response times and higher speech recognition accuracy, even in poor acoustic environments.

Overall the evaluation revealed that the speech recognition system performed well in the access scenario of interest, providing a good balance of accuracy, response time and user satisfaction. Although some further developments are needed to improve the system's robustness to different noise levels, environments and speaking styles, including more advanced noise reduction, larger training corpora with speech samples from different dialect regions, and improved language models that can better cope with variations in regional dialects.

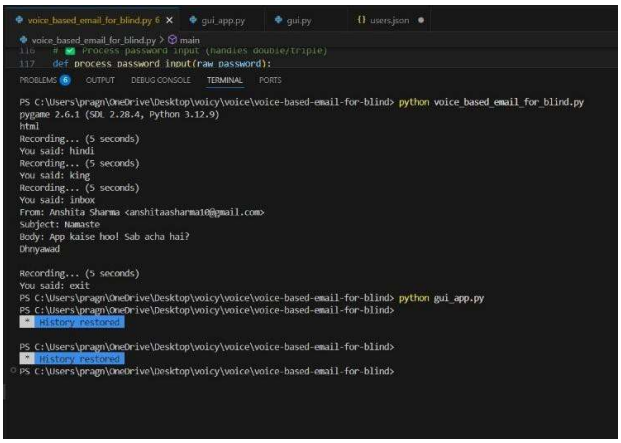

Fig.2. Voice email system running successfully

Fig 2. Shows an email from "Anshita Sharma" with the subject "Namaste." The message in Hindi asks about the recipient's well-being and expresses gratitude

Fig 3. shows a terminal running a Python script for a voice-based email system. The system successfully fetches and reads the email from Anshita Sharma using voice commands.

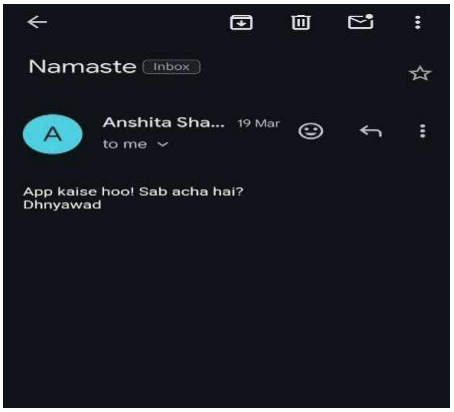

Fig.3. Mail sent successfully

TABLE II: USER SATISFACTION SURVEY

	Number of satisfied users	Number of partially satisfied users	Number of dissatisfied users
English	39	10	1
Hindi	34	9	7

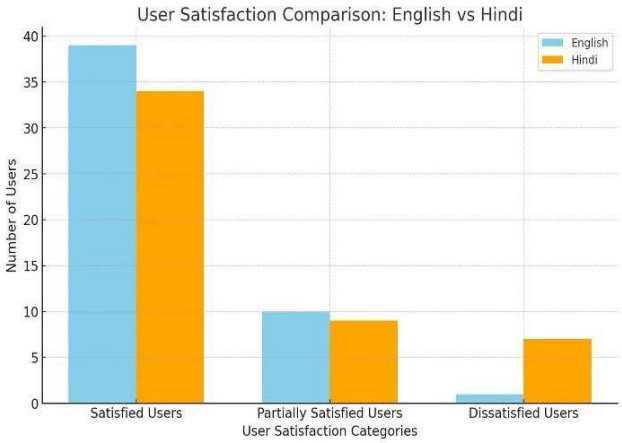

Fig. 4. Graph showing user satisfaction survey

V. DISCUSSION

Key findings include:
Accessibility The system enhances accessibility for visually impaired users who can now manage their emails via voice commands.
Error Handling Optimization: Features to improve error handling will help to improve the user experience, and there is a need for further optimization.
This has enhanced the security measures significantly but there is still scope for improvement via the use of biometrics and anomaly detection.
Customization Needs: Needs related to the customization of voice interaction applications were derived. In particular, the need for having personalized voice commands and thereby increasing user efficiency.VI.

FUTURE WORK

Future improvements include:

- Enhanced AI-Driven Conversations: Implementing AI chatbots for intuitive interactions, reducing the need for repetitive commands.
- Expanded Language Support: Incorporating additional languages and dialects, making the system more accessible globally.
- Offline Functionalities: Enabling full email management without internet connectivity, ensuring accessibility in remote areas.
- Integration with Wearable Devices: Extending support for smart watches and AR glasses for a more seamless user experience.

VII. CONCLUSION

Key Objectives of this Project The proposed system is targeted to alleviate the challenges encountered by Visually Impaired (VI) people in managing emails by proposing voice-controlled email system. Using natural language processing (NLP), text-to-speech (TTS) and speech-to-text (STT) technologies, users can create, send, read and manage their emails without the need to any assistive devices or having to see or use the computer keyboard/mouse. Some of the core features will include secure voice biometric authentication, audio feedback, real time voice to text transcription, amongst others. However, the system is not perfect and still experiences occurrences of the mis-recognition of spoken words and language-dependent issues as well as potential security threats emanating from this unique authentication approach. A host of design improvements, effective error handling mechanisms and enhanced Artificial Intelligence (AI) models are needed. The voice-based system proposed is expected to reduce the digital gap experienced by Visually
Impaired people, thus providing them with more control over their emails. The next wave of upgrades will include the addition of conversational AI that will read out the incoming emails, possible addition of working offline, and multi-lingual support to name but a few.

REFERENCES

[1] The Way to Make Blind People Use the Email System: Voice Based Email Generating System Using Artificial Intelligence, IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10085638

[2] B. Patil, "Speech-to-Text Enhancements in Accessibility Tools," IEEE Xplore, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/10541382

[3] P. Shetty, "Enhancing Accessibility in Digital Communications," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10094386

[4] S. Sharma, "NLP for Email Automation," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10393204

[5] L. Nair, "Multi-Lingual Speech Recognition for Email Applications," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10779857

[6] A. Karthik, "Secure Voice-Based Authentication for Email Systems," IEEE Xplore, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10370972

[7] V. Gupta, "Security Concerns in Voice-Based Systems," IEEE Xplore, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10612629

[8] M. Suresh, "Interactive Voice Response for Email Accessibility," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/6481791

[9] T. Rao, "Machine Learning for Speech Processing," IEEE Xplore,2024.[Online].Available: https://ieeexplore.ieee.org/document/10702865

[10] D. Ghosh, "Voice-Enabled Assistive Technology for the Visually Impaired," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10779857

[11] R. Menon, "AI-Powered Email Assistants," IEEE Xplore, 2023.[Online].Available:https://ieeexplore.ieee.org/document/9776840

[12] S. Kumar, "Deep Learning for Speech-to-Text Systems," IEEEXplore,2022.[Online].Available: https://ieeexplore.ieee.org/document/10127558

[13] A. Naik, "Natural Language Processing in Assistive Technologies," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10073879

[14] K. Patel, "Comparative Analysis of Speech Recognition APIs," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10391988

[15] S. Rao, "Impact of AI in Accessibility Solutions," IEEE Xplore, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9765171

[16] R. Mehta, "Evaluation of AI-Driven Conversational Agents," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10094386

[17] S. Chandra, "The Role of NLP in Email Voice Automation," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10073879

[18] P. Bose, "Advancements in Speech Recognition for Digital Assistants," IEEE Xplore, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10009538