International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Voice-Based Email for Visually Challenged Using Python

Krishna Mehar[1], Pooja S Rao[2], S Deshitha[3], P Pavan Raj[4], Achyuth S[5].

[1] Assistant Professor, Department of AI & ML, Impact College of Engineering and Applies Sciences, Bengaluru, India

[2-5] Students, Department of AI & ML, Impact College of Engineering and Applies Sciences, Bengaluru, India

Abstract: Email has become a vital tool due to the growing reliance on digital communication, but because traditional email systems rely on graphical user interfaces and keyboard inputs, visually impaired people experience considerable challenges while utilizing them. The Voice-Based Email System shown in this research allows visually impaired users to send, receive, and navigate emails using only voice instructions. The system creates an accessible and user-friendly communication environment through the employment of Interactive Voice Response (IVR), Text-to-Speech (TTS), and Speech- to-Text (STT) technologies. The system, which is implemented in Python and makes use of modules like SpeechRecognition, Pyttsx3, SMTPLIB, and IMAPLIB, guarantees independence, lessens cognitive strain, and improves digital inclusivity for users who are visually impaired.

Keywords: Voice-based email, Speech-to- Text, Text-to-Speech, IVR, Accessibility, Visually impaired users, Assistive technology.

I. INTRODUCTION

Email communication has become essential to modern life since it makes information sharing easy. However, because typical email systems rely on visual cues, screen navigation, and keyboard-based inputs, visually challenged people have difficulty utilizing them. Despite the availability of assistive devices, such as screen readers, many visually impaired people still rely on others for digital communication since they have steep learning curves.

To remove these obstacles, this proposal suggests a voice-based email system. The system provides total accessibility, independence, and usability by utilizing speech commands for all email functions, including writing, sending, reading, and managing messages.

Implemented in Python with packages that allow for email automation and speech processing, the solution combines STT, TTS, and IVR technologies.

1.1 Objectives:

- To develop a fully voice-controlled email system that enables visually impaired users to send, receive, and manage emails independently.
- To integrate Speech-to-Text (STT) and Text-to-Speech (TTS) technologies for accurate voice input and natural audio feedback.
- To reduce reliance on keyboards, screens, and assistive devices by enabling complete hands-free email interaction.
- To improve accessibility and digital inclusivity by simplifying email operations through intuitive voice commands.
- To provide real-time audio guidance and confirmations using IVR, ensuring a smooth and error- free user experience.

II. SYSTEM REQUIREMENTS

2.1 Hardware Requirements

Input Devices

- **Microphone (External or Built-in):**
  Required for capturing voice commands clearly. A noise- cance ling microphone is recommended for optimal speech recognition accuracy.

Output Devices

- **Speakers/Headphones:**
  Used to deliver audio output from the Text-to-Speech engine, system prompts, and read-out email content.

Processing Unit

- **Minimum CPU:** Dual- core processor
- **Recommended CPU:** Intel i5 / AMD Ryzen 5 or higher
  Provides smooth handling of speech processing and real-time responses.

Memory & Storage

- **RAM:** Minimum 4 GB (8 GB recommended)
- **Storage:** At least 200 MB of free disk space for Python, libraries, and dependent modules.

Network Requirements

- **Stable Internet Connection:**
  Required for accessing SMTP/ IMAP servers, using cloud-based speech recognition (if selected), and performing email operations.

2.2 Software Requirements

Operating System

- Windows 10 or newer
- Linux distributions (Ubuntu, Mint, Fedora)
- macOS

Programming Environment

- **Python 3.x** (3.7 or above preferred)

Python Libraries & Dependencies

- **SpeechRecognition:**

For converting spoken input into text.

- **PyAudio:**

Captures microphone input; essential for real-time speech acquisition.

- **Pyttsx3:**

Provides offline text-to-speech synthesis.

- **SMTPLIB:**

Handles email sending via SMTP protocol.

- **IMAPLIB:**

Fetches and processes incoming emails from IMAP server.

- **Email & MIME Libraries:**

For constructing messages, formatting subjects, and handling attachments (if extended).

Email Server Requirements

- **SMTP Server:** smtp.gmail.com
- **IMAP Server:** imap.gmail.com
- Account must have:
  - IMAP access enabled
  - App-specific password (if 2-factor authentication is ON)

Additional Utilities

- Updated browser drivers if Selenium is used (optional)
- Latest security permissions for microphone access

III. SYSTEM ARCHITECTURE

The Voice-Based Email System follows a modular, layered architecture designed for reliability, accessibility, and ease of expansion. The architecture combines voice processing, email handling, and interactive guidance to deliver a seamless user experience for visually impaired users.Input layer: The input layer handles user interaction via a Streamlit interface, where users provide a text prompt describing the desired video concept. This layer serves as the starting point of the pipeline, capturing the requirements and context for the subsequent layers.

Components

- **Speech Input Module**

Captures user's voice commands using microphone input.

- **Speech-to-Text (STT) Engine**

Converts spoken commands into text using SpeechRecognition.

- **Processing Engine**
  - Detects commands such as "compose mail," "read inbox," etc.
  - Uses SMTP for sending mails and IMAP for reading mails.
  - Handles authentication, message formatting, inbox retrieval.

1. **Text-to-Speech (TTS) Engine**
Converts email content and system messages into speech using Pyttsx3.
2. **IVR Guidance Layer**
Provides audio prompts so the user is guided throughout the process.

Architecture Flow

1. User speaks a command
2. Voice is converted to text
3. System identifies requested operation
4. Performs SMTP/IMAP action
5. Outputs results via TTS

IV. IMPLEMENTATION DETAILS

Speech Recognition

- Implemented using **SpeechRecognition** library.
- Captures microphone input and converts to text.
- Google Web Speech API ensures high accuracy.
- Noise handling via adjust_for_ambient_noise().

Text-to-Speech Conversion

- Implemented using **pyttsx3** for offline speech synthesis.
- Reads instructions, email content, errors and confirmations aloud.

Sending Email (SMTP)

- Uses SMTPLIB to connect to Gmail's SMTP server.
- Sends emails with subject and body formed from user voice input.
- Follows RFC 822 email formatting.

Reading Emails (IMAP)

- IMAPLIB retrieves unread or important emails.
- Extracts sender, subject, and body.
- Converts them into speech via TTS.

IVR Navigation

System continuously guides user with prompts like:

- "Say the recipient's email ID."
- "Your email has been sent."
- "You have three unread messages."

V. CONCLUSION

The Voice-Based Email System successfully overcomes the accessibility challenges faced by visually impaired individuals by enabling complete email management through speech. With integrated STT, TTS, and IVR technologies, the system eliminates the need for screen navigation and manual typing, thereby offering independence and convenience. Future improvements include support for attachments, multilingual capabilities, enhanced noise filtering, and integration with mobile platforms.

REFERENCES

[1]R. D. Chavare, S. B. Ketkale, "Speech- to-Text and Text-to-Speech Technologies in Assistive Communication," 2022.

[2] Madhuri Rajole, Rajesh Phursule, " Voice- Based Email Systems for Accessibility," 2023.

[3] Aditi Sharma et al., "Advancements in Voice-Driven Email Solutions," 2024.

[4] Sherly Noel, "IVR-Integrated Assistive Email Platforms," 2024.

[5] Python Documentation – SpeechRecognition, Pyttsx3, SMTPLIB, IMAPLIB.

[6]RFC 822 Standard for Internet Text Messages.