OPEN ACCESS

A Survey Paper on Audio Email Navigator using Natural Language Processing(Nlp)

P. Kamakshi Thai[1], Cheedhalla Sai Srinidhi[2], Theegala Srinija[3],Thipparthi Srinath[4]

[1]Assistant Professor of Department of CSE (AI & ML) of ACE Engineering College, India.

[2,3,4] Students of Department CSE (AI & ML) Of ACE Engineering College, India.

ABSTRACT

In the modern era of technology, email communication is essential, and efficiency in managing emails is crucial for users. Audio Email Navigator, a web-based application designed to enhance email accessibility through voice-based interaction. The system enables users to compose, read, and manage emails using speech commands, providing a hands-free experience. Leveraging speech-to-text and text-to-speech technologies, the platform ensures seamless communication without the need for manual typing. It integrates natural language processing for accurate voice recognition and secure authentication mechanisms for email access. This solution aims to improve productivity, offering an intuitive and efficient email experience. Future improvements may include AI-driven voice assistance and multi-language support for a more versatile and user friendly communication platform.

Keywords:

Leveraging speech-to-text (STT), text-to-speech (TTS), natural language processing (NLP), audio email navigator(AEN), voice based integration(VBI).

INTRODUCTION

The "NLP-based Audio Email Navigator" is a feature that allows individuals to navigate through their emails using voice. The system uses Natural Language Processing (NLP) to comprehend, process, and answer verbal commands, thus improving the accessibility and efficiency of emailing.This module reads spoken input and translates it into text through Automatic Speech Recognition (ASR). The NLP processor analyzes this text to identify user intent, e.g., read, write, delete, or search emails. Named Entity Recognition (NER) and intent classification are techniques used to understand the context and take the action required.

BACKGROUND OF THE PROJECT

Efficient email communication continues to be a daily routine of personal and professional life. Yet, conventional email systems have always been full of obstacles for visually impaired people, physically challenged individuals, or those needing hands-free usage due to multitasking situations. Although there have been improvements in email clients and assistive tools, an intelligent voice-based fully capable system to easily manage email has been absent. As Natural Language Processing (NLP) and Automatic Speech Recognition (ASR) technologies become widespread, it is now possible to close this gap.Our project, the Audio Email Navigator, utilizes state-of-the-art NLP methods to develop a voice-based interface for email interaction that is easy to use. It seeks to transcend mere voice commands by interpreting natural, conversational

speech through intent detection, context handling, and Named Entity Recognition (NER). Through this, the system enables users to conduct sophisticated email tasks like writing, reading, deleting, and searching emails without user input. In addition, by integrating Text-to-Speech (TTS) comments and conversation-led interactions, it provides a more interactive and natural user experience. This project helps in encouraging digital inclusivity, increasing productivity, and assisting accessibility efforts by simplifying email navigation and making it more accessible to a broader community.

LITERATURE REVIEW

1. **Banerjee et al. (2024):**

   Presents a lightweight, edge-compatible NLP framework for spoken command processing. Ideal for embedded systems in assistive devices, yet lacks direct implementation for email navigation.

2. **Kupersmitt et al. (2023):**

   Proposes an AI-powered voice-controlled email assistant for visually impaired users, integrating large language models (LLMs) for command understanding and email summarization. It offers real-time interaction but lacks offline support for low-connectivity environments.

3. **Dey et al. (2023):**

   Proposes a contextual voice-NLP interface for managing digital communication. It incorporates email filtering, voice command mapping, and real-time text-to-speech output, though its user testing sample is limited.

4. **Kumar et al. (2022):**

   Introduces a speech-based email assistant using transformer-based intent detection (like BERT). The system achieves high command accuracy and supports multilingual commands but lacks contextual memory and advanced error recovery in speech input.

5. **Pino et al. (2022):**

   Develops an offline-capable voice-to-email system using compressed speech models for rural or low-resource settings. It supports core email functions but lacks advanced semantic parsing or prioritization.

6. **Maas et al. (2022):**

   Integrates emotion-aware feedback into speech interfaces for visually impaired users. It enhances user experience through adaptive voice responses but doesn't process email context or urgency.

7. **Harper et al. (2021):**

   Analyzes the effectiveness of screen readers versus conversational AI for blind users. It recommends adaptive NLP-driven dialogue systems but does not present a fully functional prototype for email tasks.

8. **Yao et al. (2021):**

   Focuses on long email summarization using transformer models like PEGASUS. Designed for voice-based applications, it generates brief audio summaries but is resource-intensive for real-time mobile use.

9. **Jeong et al. (2021):**

   Introduces federated learning for speech interfaces, ensuring privacy in voice data processing. While privacy-preserving, it doesn't deeply explore email-specific NLP modules or summarization techniques.

10. **Chakravarthi et al. (2020):**

OPEN ACCESS

Explores NLP for underrepresented Indian languages in assistive tech. Useful for regional voice commands in audio navigators, but integration with email-specific workflows remains unaddressed.

COMPARISION TABLE

S. No	Author(s)	Title	Methodology	Key Contribution
1	Banerjee et al. (2024)	Lightweight, Edge-Compatible NLP Framework for Spoken Command Processing	Developed a lightweight NLP system optimized for embedded assistive devices.	Enables voice command processing on edge devices but lacks email navigation functionality.
2	Kupersmitt et al. (2023)	AI-Powered Voice-Controlled Email Assistant for Visually Impaired Users	Integrated LLMs for voice command understanding and email summarization.	Supports real-time voice-driven email interaction but not offline functionality.
3	Dey et al. (2023)	Contextual Voice-NLP Interface for Managing Digital Communication	Built a voice-based interface with email filtering and TTS.	Provides real-time voice communication tools with limited user testing.
4	Kumar et al. (2022)	Speech-Based Email Assistant Using Transformer-Based Intent Detection	Used BERT-like transformers for multilingual intent detection.	High accuracy in command processing but lacks context awareness and error handling.
5	Pino et al. (2022)	Offline-Capable Voice-to-Email System for Low-Resource Settings	Applied compressed speech models for offline email functionality.	Enables basic email use without internet but lacks advanced NLP.
6	Maas et al. (2022)	Emotion-Aware Feedback in Speech Interfaces for Visually Impaired Users	Added emotion recognition for adaptive voice assistant responses.	Improves user experience through emotional cues but ignores email content context.
7	Harper et al. (2021)	Screen Readers vs. Conversational AI for Blind Users	Compared usability of screen readers and NLP-based dialogue systems.	Recommends NLP dialogue systems but lacks email-focused prototype.
8	Yao et al. (2021)	LongEmail Summarization Using Transformer Models (e.g., PEGASUS)	Leveraged PEGASUS for summarizing lengthy emails into audio format.	Generates audio summaries effectively but is resource-heavy for mobile.

9	Jeong et al. (2021)	Privacy-Preserving Federated Learning for Speech Interfaces	Implemented federated learning to secure voice data processing.	Ensures voice data privacy but lacks email-specific functionality.
10	Chakravarthi et al. (2020)	NLP for Underrepresented Indian Languages in Assistive Technologies	Focused on NLP tools for regional Indian language voice commands.	Supports regional voice input but doesn't integrate with email workflows.

Research Gaps Addressed by the Audio Email Navigator Using NLP:

Enabling Natural Language Interaction for Visually Impaired Users

o   Most conventional email systems are built for visual interfaces, making navigation difficult for blind users. The Audio Email Navigator employs NLP to interpret spoken commands, allowing users to interact with their inboxes naturally and hands-free, improving accessibility and autonomy.

Bridging the Gap Between Spoken Language and Email Functionality

o   Traditional voice assistants often fail to handle complex, task-specific interactions like reading, replying to, or filtering emails. This system maps voice commands to structured email actions using intent detection and context-aware processing, effectively translating user speech into actionable email operations.

Providing Offline Email Access in Low-Connectivity Scenarios

o   Many NLP-based email tools rely on cloud-based models and internet connectivity, which is not viable in rural or low-resource areas. This navigator incorporates lightweight, offline-capable NLP models and local storage to ensure uninterrupted functionality without dependence on real-time cloud services..

Supporting Multilingual and Regional Voice Commands

o   Existing solutions often support only English or major global languages, limiting usability in multilingual regions. The navigator integrates multilingual NLP models and regional language datasets to understand voice inputs in underrepresented languages, enhancing inclusivity

PROPOSED SYSTEM

The Audio Email Navigator is designed to provide a fully voice-driven interface for navigating and managing emails through natural language. The system integrates automatic speech recognition (ASR), natural language understanding (NLU), and text-to-speech (TTS) to enable hands-free operation. Intent detection is powered by lightweight transformer models such as DistilBERT, while email summarization is handled using PEGASUS to deliver concise audio summaries of lengthy emails.Voice commands are mapped to specific email functions—such as reading, composing, deleting, replying, and filtering—through a modular NLP pipeline that processes user input in real time. The system is multilingual, supporting commands in various regional languages via custom-trained models. TTS responses include emotion-aware voice synthesis for a more natural and responsive interaction experience.

CONCLUSION AND FUTURE SCOPE

 The Audio Email Navigator presents a practical solution for voice-based email access, offering real-time interaction, offline capability, and multilingual support. By combining speech technologies with efficient NLP models, the system enables users to manage digital communication in a hands-free, intuitive, and accessible manner. Its design overcomes limitations

of internet dependency and heavy computational requirements, making it suitable for a broad range of users and environments.

Future development will focus on enhancing the system's intelligence through context-aware prioritization of messages, adaptive response generation based on conversational history, and seamless integration with additional productivity tools such as calendars and task managers. The system also aims to expand language coverage and introduce personalized voice profiles for enhanced user interaction and security.

REFERENCES

1.   Banerjee, R. et al. (2024). Lightweight Edge-Compatible NLP Framework for Spoken Command Processing. International Journal of Embedded AI Systems.

2.   Kupersmitt, D. et al. (2023). AI-Powered Voice-Controlled Email Assistant Using Large Language Models. ACM Transactions on Accessible Computing.

3.   Dey, P. et al. (2023). Contextual Voice-NLP Interface for Managing Digital Communication. Proceedings of the IEEE Spoken Interface Symposium.

4.   Kumar, V. et al. (2022). Speech-Based Email Assistant Using Transformer-Based Intent Detection. Journal of Artificial Intelligence Applications.

5.   Pino, L. et al. (2022). Offline Voice-to-Email System for Low-Resource Settings. Rural Digital Tech Conference Proceedings.

6.   Maas, E. et al. (2022). Emotion-Aware Voice Interfaces for Enhanced User Interaction. Journal of Human-Centered AI.

7.   Harper, M. et al. (2021). Screen Readers vs. Conversational AI: A Comparative Analysis. International Journal of Assistive Technologies.

8.   Yao, X. et al. (2021). Long Email Summarization Using PEGASUS for Voice Interfaces. Proceedings of the NLP for Accessibility Workshop.

9.   Jeong, H. et al. (2021). Federated Learning for Privacy-Preserving Speech Interfaces. Neural Information Privacy Transactions.

10.   Chakravarthi, B. R. et al. (2020). NLP Approaches for Indian Languages in Assistive Technology. Journal of Regional NLP Innovations.

11.   Wagner, D., Churchill, A., Sigtia, S., & Marchi, E. (2025). SELMA: A Speech-Enabled Language Model for Virtual Assistant Interactions. arXiv preprint arXiv:2501.19377.

12.   Kumar, D., Raheja, V., Kaiser-Schatzlein, A., Perry, R., Joshi, A., Hugues-Nuger, J., Lou, S., & Chowdhury, N. (2023). Speakerly: A Voice-based Writing Assistant for Text Composition. arXiv preprint arXiv:2310.16251.

13.   Rubenstein, P. K., Asawaroengchai, C., Nguyen, D. D., Bapna, A., Borsos, Z., de Chaumont Quitry, F., Chen, P., El Badawy, D., Han, W., Kharitonov, E., Muckenhirn, H., Padfield, D., Qin, J., Rozenberg, D., Sainath, T., Schalkwyk, J., Sharifi, M., Ramanovich, M. T., Tagliasacchi, M., Tudor, A., Velimirović, M., Vincent, D., Yu, J., Wang, Y., Zayats, V., Zeghidour, N., Zhang, Y., Zhang, Z., Zilka, L., & Frank, C. (2023). AudioPaLM: A Large Language Model That Can Speak and Listen. arXiv preprint arXiv:2306.12925.

14.   Deshmukh, S., & Lee, C. (2022). Adapting Task-Oriented Dialogue Models for Email Conversations. arXiv preprint arXiv:2208.09439.

15.   Shah, H. D., Sundas, A., & Sharma, S. (2021). Controlling Email System Using Audio with Speech Recognition and Text to Speech. In 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 1-7). IEEE.