

Volume 2 Issue 4

Article Number: 23069

Enhancing Email Accessibility for Visually Impaired Individuals: A Voice-Based Email System using Artificial Intelligence

Jaspreet Kaur* and Rohit Agnihotri

Department of Computer science and Engineering, JSS Academy of Technical Education, Affiliated to AKTU, Noida, India 201301

Abstract

This paper presents an innovative, voice-based email system designed to improve email accessibility for visually impaired individuals. The proposed system leverages Artificial Intelligence and speech recognition technologies to convert speech to text and text to speech, enabling visually impaired individuals to send and receive emails using voice commands. The system offers an intuitive user interface, secure authentication measures, and robust database architecture to ensure seamless and secure user experience. Comparative analysis with traditional email systems reveals the superior functionality and inclusivity of the proposed system. Despite certain limitations, future enhancements promise to further refine the system, paving the way for a more inclusive digital communication environment.

Keywords: Artificial Intelligence; Speech Recognition; Email Accessibility; Visually Impaired; Voice-Based Email System

1 Introduction

With rapid technological advancements and increased internet accessibility, numerous aspects of our lives have become digitalized, including communication [1]. Indeed, communication is one field that has significantly evolved due to technological advancements, making distance a minor factor [2]. One of the most reliable methods for transmitting essential information in this digital age is email [3], a tool used globally. However, not everyone can equally access this beneficial tool. To access the internet and use email, one must be able to see, a prerequisite that poses challenges to a significant number of visually impaired or outwardly impeded individuals globally [4, 5]. Visual impairment restricts individuals from interacting with standard web interfaces that typically require visual input and output [6, 7]. Unfortunately, this means that a significant number of people are effectively cut off from the conveniences of email and the broader web [8–10]. Visually impaired individuals face difficulties in sending and receiving emails, understanding the content provided by email, and using the existing email systems due to their inherent visual interface [10–12]. In the current scenario, a visually impaired person has only one choice for sending an email: they must verbally provide a third person with the entire content of the mail, who then types and sends the mail on their behalf [13]. This practice, however, neither guarantees privacy nor empowers visually impaired individuals, instead reinforcing their dependency on others. In light of these challenges, the authors propose a concept designed to make the digital world, particularly email communication, more accessible to visually impaired individuals. This innovative solution allows a visually impaired person to send and receive emails using voice commands rather than relying on visual devices or keyboards. The proposed system aims to increase societal inclusivity and independence for visually impaired individuals, transforming the way they interact with the digital world.

*Corresponding author: jaspreetkaur@jssaten.ac.in

Received: 12 July 2023; **Revised:** 19 July 2023; **Accepted:** 20 July 2023; **Published:** 30 September 2023

© 2022 Journal of Computers, Mechanical and Management.

This is an open access article and is licensed under a [Creative Commons Attribution-Non Commercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

DOI: [10.57159/gadl.jcmm.2.4.23069](https://doi.org/10.57159/gadl.jcmm.2.4.23069).

Artificial Intelligence (AI) plays a crucial role in making this concept a reality. AI is a technology used to develop intelligent systems and robots that mimic human intelligence [14]. Expert systems, natural language processing (NLP), machine vision, and speech recognition are some applications of AI [15–17] that are particularly relevant to the proposed solution in this study. NLP is the process of understanding and analyzing human language, such as English, by extracting information from keywords, emotions, relationships, and concepts. This technology enables the transformation of voice commands into executable actions, opening a world of opportunities for user interfaces [18, 19]. The proposed system in this article also uses AI to convert speech to text (STT) and text to speech (TTS). Google’s Cloud STT service provides developers with a straightforward API for converting audio to text. It includes robust neural network models that recognize over a hundred and twenty languages and variations [20–22]. The TTS engine, conversely, reproduces spoken language from written text, making our system interact with users in a natural, human-like way [23]. This technology allows computers to talk to users, making the system more interactive and user-friendly.

2 Related Work

The proliferation of the internet and digitization has led to an exponential growth in the use of email as a mode of communication [24]. According to the Email Statistics Report of 2014-2018 by a tech market research firm based in Palo Alto, CA, the number of email accounts worldwide has increased from 4.1 billion in 2014 to over 5.2 billion by the end of 2018 [25, 26]. This substantial growth underscores the prevalence of email as a primary mode of communication. However, it’s essential to consider the challenges that prevent certain demographics from using this medium effectively. Based on studies conducted by the Vision Loss Expert Group (VLEG), approximately 253 million people worldwide are visually impaired or blind, indicating that a large number of individuals are currently unable to access email [25, 27, 28]. Several existing systems provide email access and management features to users via web services, thereby enhancing email’s popularity as a communication medium [29]. However, most of these systems lack voice command or audio capabilities, rendering them unsuitable for visually impaired users. These traditional systems typically present information in textual format which isn’t accessible for visually impaired individuals. Although some internet browsers have the capability to play music and video, users must first input textual commands to request such media [7, 30]. This requirement of text-based interaction with web services is a significant impediment for blind users. A noteworthy mention here is the role of screen readers in assisting visually impaired individuals in accessing digital content. Screen readers interpret and read aloud the text displayed on a screen [31]. However, these tools have significant limitations. They read the text sequentially [32], which can be inefficient for complex pages with lots of content. Screen readers can only interpret content provided in basic HTML [33, 34]. Since many modern web pages use advanced languages like CSS, Bootstrap, JavaScript, and others to enhance appearance and usability, screen readers often fail to read and understand these pages [35–37]. Some applications have been developed to aid visually impaired individuals, like mobile apps capable of interpreting and reading data encoded as a barcode on a product [38–40]. Despite being innovative, these solutions are device-specific and not universally applicable, limiting their use. Moreover, they are ineffective in scenarios like accessing emails, which are primarily text-based. The crux of the issue lies in the fact that traditional systems do not offer effective, intuitive, and inclusive solutions for visually impaired users. This gap is what the authors’ proposed system seeks to address. Leveraging AI and speech recognition technologies, the proposed system aims to provide a voice-based interface for email communication. This would significantly enhance email accessibility for visually impaired individuals, empowering them to send and receive emails independently. The proposed system innovatively transforms the traditional, visual, and text-based email experience into an auditory one, marking a substantial step forward in inclusivity in the digital world. The rest of the article sections shall delve into the methods and technologies behind the proposed system, detailing how it aims to revolutionize email usage for visually impaired individuals.

3 Methods

The methodology behind the proposed voice-based email system involves several crucial steps ranging from the design of the system’s user interface, the architecture of the database, the overall system design, and the development of the mail programming module. The system is subsequently implemented through various features including login, dashboard, send mail, and inbox operations. Finally, the system ensures user authentication and data security.

3.1 Proposed work design

UI Design

The development process begins with designing the system’s user interface (UI). This includes the creation of all the web content with which users will interact. An intuitive, user-friendly UI is critical for the success of the application, especially given that it’s intended for visually impaired individuals. To make the system universally accessible, the design process employs HTML5 and CSS3 to create a seamless, interactive, and responsive interface.

Database architecture

As the system stores user credentials and email data, a reliable database is a necessity. The architecture of the database includes the construction of various tables designed to store user credentials for authentication purposes and to hold user emails securely. This database architecture serves as the backbone of the system, enabling the efficient storage and retrieval of data.

System design

The system design incorporates all the modules necessary for the framework, including the Text-to-Speech (TTS) and Speech-to-Text (STT) modules, and a Mail programming module. The design ensures that each module complements the others and collectively contributes to the system's seamless operation. In our system, the Text-to-Speech (TTS) conversion is handled by the Google Text-to-Speech (GTTS) service, which provides high-quality and natural-sounding speech output. The working of the GTTS algorithm is depicted in Figure 1 (a) and Figure 1 (b) provides an overview of the system design, showing the connection between different modules in the proposed system.

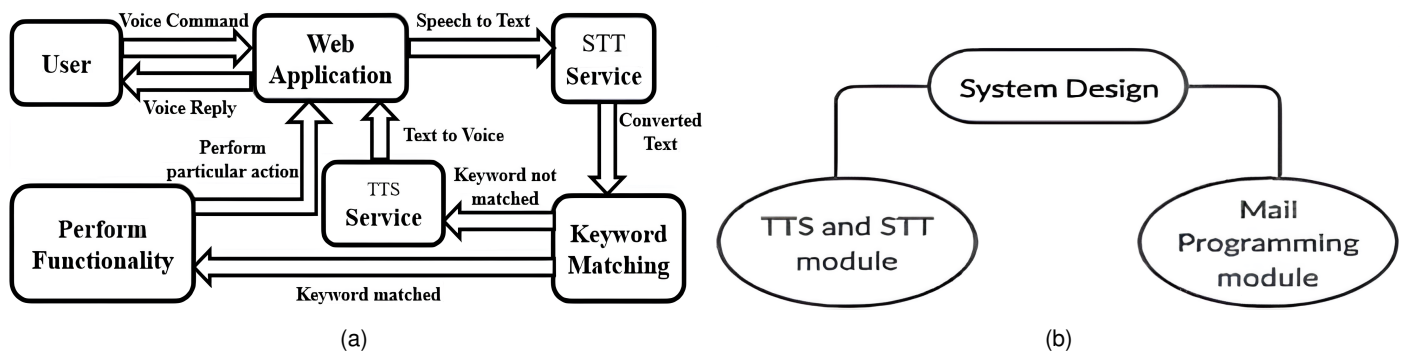


Figure 1: (a) Working principle of the Google Text-to-Speech (GTTS) algorithm; (b) Comprehensive system design highlighting the interconnections between various system modules.

Mail Programming Module

As email becomes an increasingly important web service, many internet systems utilize Simple Mail Transfer Protocol (SMTP) to send emails from one user to another [41, 42]. SMTP is responsible for sending emails, while the receiving end uses the Post Office Protocol (POP) or Internet Message Access Protocol (IMAP) to fetch the message [41, 43]. Figure 2 (a) presents the architecture of the proposed voice-based email system, illustrating how the various components interact with each other and Figure 2 (b) illustrates how the Simple Mail Transfer Protocol (SMTP) works, which is a key component in the sending and receiving of emails.

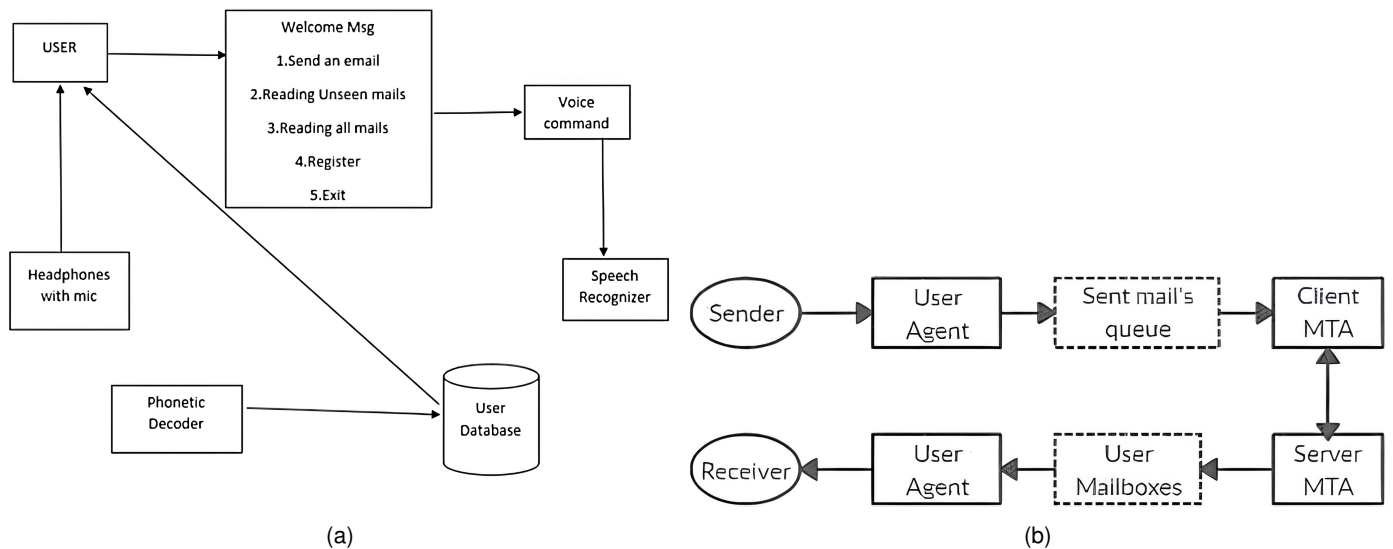


Figure 2: (a) Architecture Diagram of Voice-based e-mail system; (b) Functioning mechanism of simple mail transfer protocol (SMTP) in email transmission.

3.2 Implementation

The workflow of the proposed system is depicted in Figure 3, which shows how users navigate the application and utilize its features.

Login

Users start by logging into the system using voice commands. The user's Gmail account is the primary method of authentication. If the login is successful, the user is granted access to the system's features. The user interface of the login page, which is the entry point to the system, is shown in Figure 4 (a).

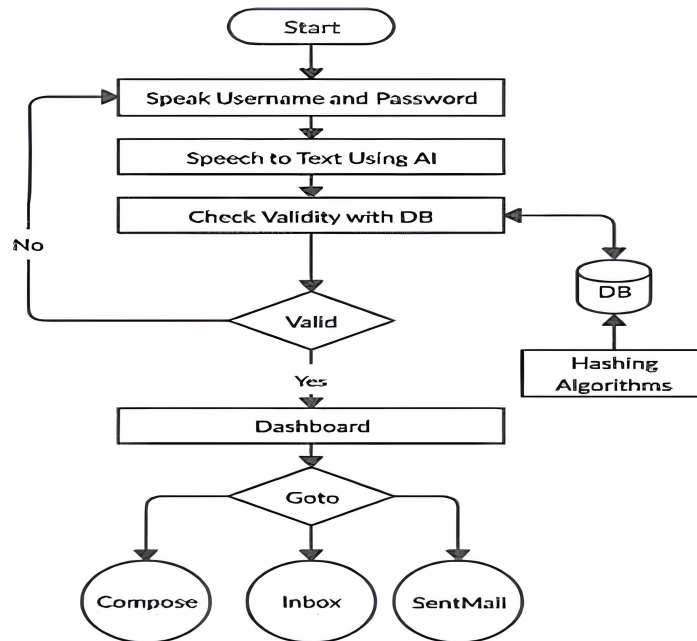


Figure 3: Workflow diagram for Voice-based e-mail system.

Dashboard

Once logged in, the user is directed to the dashboard, which offers various options including 'Inbox', 'Compose New Mail', 'Sent Mail', and 'Trash'. The system performs the corresponding actions when the user issues voice commands. Figure 4 (b) displays the system's dashboard, which offers a variety of options for the user to select using voice commands.

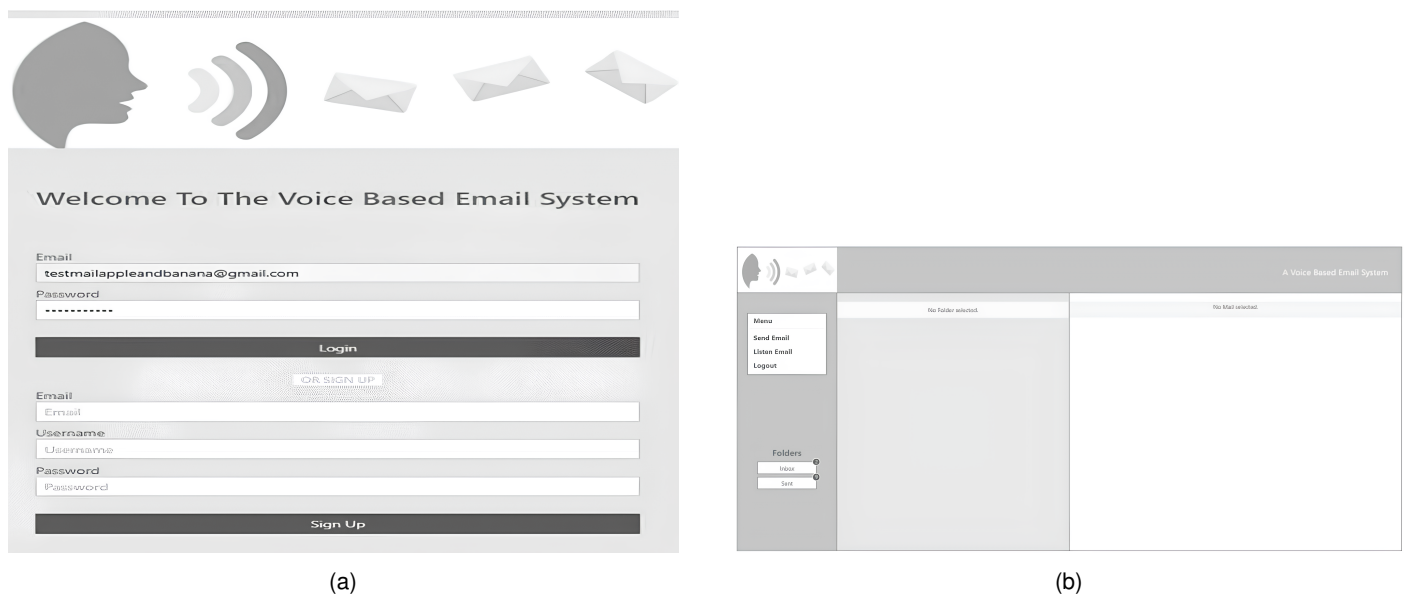


Figure 4: (a) User-interface design of the Login page for the designed system; (b) Snapshot of the dashboard interface offering multiple options to the user.

Send Mail

When a user wants to send an email, they issue a voice command saying "SEND EMAIL". The system then opens a form where the user fills out necessary details using voice commands. The system re-reads all the details to confirm their accuracy before the user sends the email. The process of composing and sending an email using the designed system is depicted in Figure 5a.

Inbox

The inbox feature reads out new emails to the user. The system alerts the user of any new email received and reads out the senders' names one by one. The user can then specify whose email they want to listen to first. Figure 5b shows the inbox interface of the designed system, where incoming emails are listed.

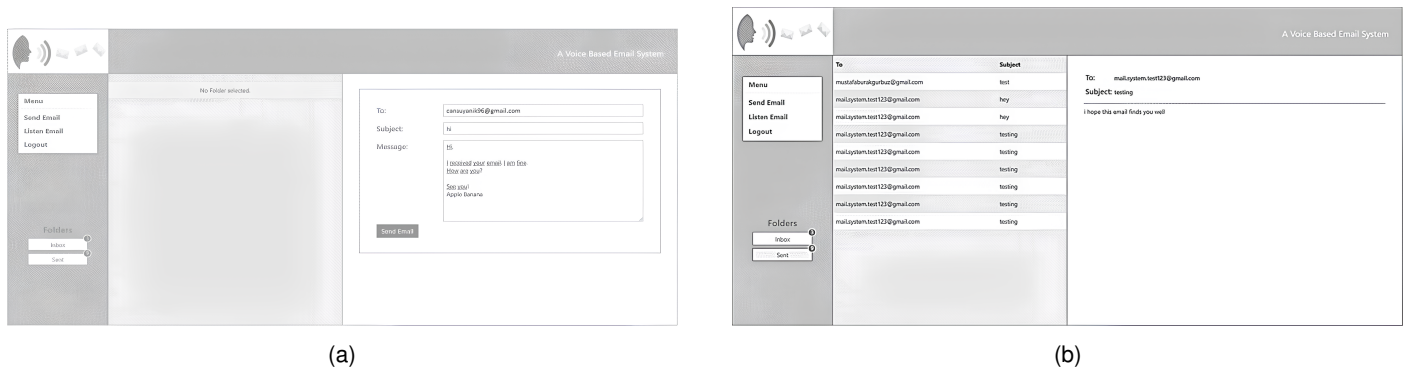


Figure 5: (a) Demonstration of voice-commanded drafting and sending of emails in the designed system; (b) The inbox interface displaying incoming emails in the voice-based email system.

3.3 Authentication and Security

Authentication and security are critical elements of the system [44]. The system implements authentication by requiring users to provide credentials, such as a username and password. These credentials are securely stored in a database and used to verify the user's identity each time they access the application. To ensure security, the system uses a hashing technique. Hashing transforms passwords into a form that cannot be converted back to the original password, significantly enhancing the security of the stored user credentials [45, 46]. The system uses common hashing algorithms like Message Digest Algorithms (MD5) and Secure Hash Algorithms (SHA) [47] to maintain data integrity and security. Figure 6 represents the hashing algorithm used in the system, a crucial component of the system's security measures.

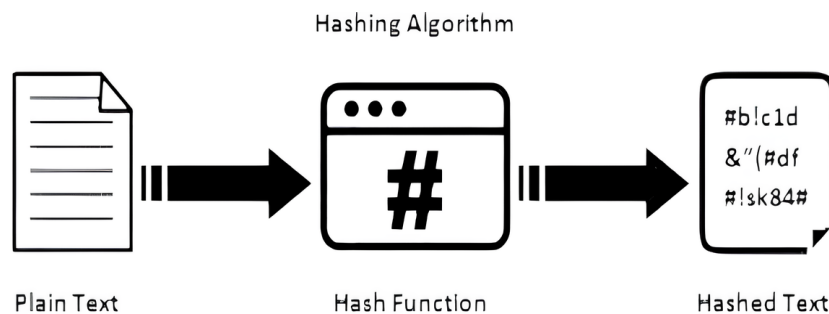


Figure 6: Depiction of the hashing algorithm used for enhancing user data Security in the System.

4 Results and Discussion

The proposed voice-based email system showcases clear advantages over traditional systems. A comparative analysis reveals that the proposed model offers unique features not present in many of the industry's established tools. The model introduces voice command and control, facilitating use by visually impaired individuals, a significant improvement over traditional systems. The model works seamlessly across all email platforms, which is a considerable advantage over other systems that may be platform-specific. In terms of speech-to-text transcription, the Google Text-to-Speech (gTTS) service used in this system provides robust performance. When compared to other services available in the market, gTTS stands out due to its versatility, language support, and accuracy.

The reliability of gTTS in the proposed system contributes to its superior functionality, helping visually impaired users navigate their emails effortlessly. The proposed voice-based email system has a broad scope for future enhancement. Potential improvements could include the integration of various languages and access to additional email categories such as deleted and spam emails. Incorporating a sign-language interpretation module could further increase the system's adaptability, making it even more robust and inclusive. The system finds its application primarily among visually impaired individuals, who can utilize this Android application for a quick and efficient email experience. The system also serves as a beneficial tool for individuals who have difficulty typing or navigating traditional email interfaces. Despite the promising features and applications of the proposed system, some limitations exist. For instance, the system's effectiveness can be hampered if the user struggles with pronunciation, as the system's operation relies heavily on voice commands. At present, the application is limited to working with Google accounts, restricting its use with other email platforms. Furthermore, the system currently lacks fingerprint authentication, which could potentially compromise user security and privacy if they inadvertently disclose their passwords and textual information. Addressing these limitations in future iterations of the system would significantly improve its effectiveness and user-friendliness. The comparison of the proposed voice-based email system with traditional email systems is demonstrated in Figure 7, highlighting the distinctive advantages of the proposed system.

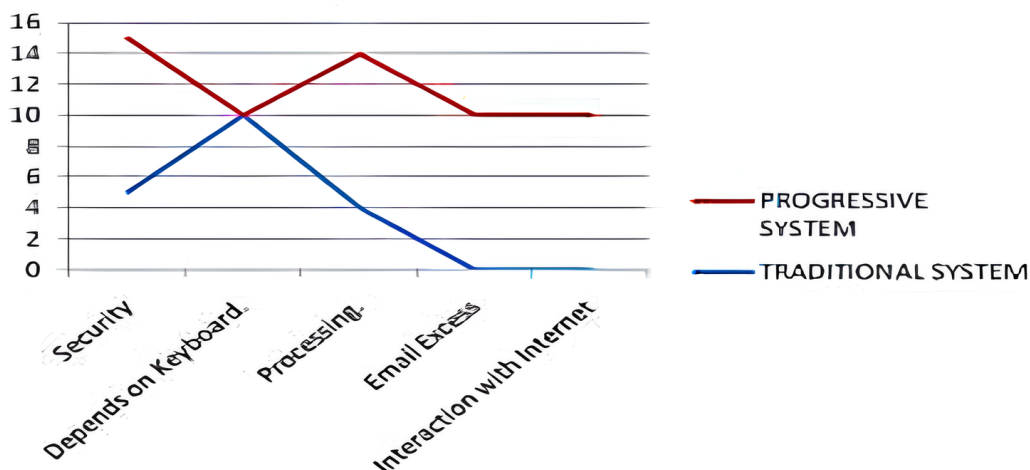


Figure 7: Comparative analysis of the proposed voice-based email system and traditional email systems.

5 Conclusion

The proposed voice-based email system is an innovative and inclusive solution that enhances email accessibility for visually impaired individuals. The system, leveraging Artificial Intelligence and speech recognition technologies, offers an auditory email experience, enabling visually impaired individuals to independently send and receive emails. While traditional systems rely heavily on visual input and output, this proposed system emphasizes the transformation of speech to text and text to speech, making the system user-friendly and practical for visually impaired users. Furthermore, the system effectively eliminates the need for keyboard shortcuts and screen readers, reducing the cognitive load of remembering keyboard shortcuts. With a user-friendly interface and the added value of security features, this system marks a significant step in enhancing the digital experience for visually impaired individuals. Although certain limitations currently exist, future enhancements and modifications hold the promise to make the system more robust and adaptive, leading to a more inclusive digital world.

Declaration of Competing Interests

The authors declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Declaration

This research did not receive any grants from governmental, private, or nonprofit funding bodies.

Author Contribution

Jaspreet Kaur: Conceptualization, Supervision, Writing- Reviewing and Editing, Project Administration; **Rohit Agnihotri:** Methodology, Data curation, Investigation, Software, Validation; Writing–Original draft preparation.

References

- [1] M. Akour and M. Alenezi, "Higher education future in the era of digital transformation," *Education Sciences*, vol. 12, no. 11, p. 784, 2022.
- [2] A. Pregowska, K. Masztalerz, M. Garlińska, and M. Osial, "A worldwide journey through distance education—from the post office to virtual, augmented and mixed realities, and education during the covid-19 pandemic," *Education Sciences*, vol. 11, no. 3, p. 118, 2021.
- [3] T. J. Blank, *Folklore and the Internet: Vernacular expression in a digital world*. University Press of Colorado, 2009.
- [4] K. Manjari, M. Verma, and G. Singal, "A survey on assistive technology for visually impaired," *Internet of Things*, vol. 11, p. 100188, 2020.
- [5] A. Webster and J. Roe, *Children with visual impairments: Social interaction, language and learning*. Psychology Press, 1998.
- [6] G. R. Hayes, S. Hirano, G. Marcu, M. Monibi, D. H. Nguyen, and M. Yeganyan, "Interactive visual supports for children with autism," *Personal and ubiquitous computing*, vol. 14, pp. 663–680, 2010.
- [7] B. Shneiderman and C. Plaisant, *Designing the user interface: Strategies for effective human-computer interaction*. Pearson Education India, 2010.
- [8] Y. Yu, S. Ashok, S. Kaushik, Y. Wang, and G. Wang, "Design and evaluation of inclusive email security indicators for people with visual impairments," in *2023 IEEE Symposium on Security and Privacy (SP)*, pp. 2885–2902, IEEE, 2023.
- [9] F. A. Inan, A. S. Namin, R. L. Pogrund, and K. S. Jones, "Internet use and cybersecurity concerns of individuals with visual impairments," *Journal of Educational Technology & Society*, vol. 19, no. 1, pp. 28–40, 2016.
- [10] A. M. Piper, R. Brewer, and R. Cornejo, "Technology learning and use among older adults with late-life vision impairments," *Universal Access in the Information Society*, vol. 16, no. 3, pp. 699–711, 2017.
- [11] R. Brewer, R. C. Garcia, T. Schwaba, D. Gergle, and A. M. Piper, "Exploring traditional phones as an e-mail interface for older adults," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 8, no. 2, pp. 1–20, 2016.
- [12] J. Hailpern, L. Guarino-Reid, R. Boardman, and S. Annam, "Web 2.0: blind to an accessible new world," in *Proceedings of the 18th international conference on World wide web*, pp. 821–830, 2009.
- [13] C. A. Beverley, P. Bath, and A. Booth, "Health information needs of visually impaired people: a systematic review of the literature," *Health & Social Care in the Community*, vol. 12, no. 1, pp. 1–24, 2004.
- [14] S. Kumar, U. Gupta, A. K. Singh, and A. K. Singh, "Artificial intelligence: Revolutionizing cyber security in the digital era," *Journal of Computers, Mechanical and Management*, vol. 2, no. 3, pp. 31–42, 2023.
- [15] M. Kocaleva, D. Stojanov, I. Stojanovic, and Z. Zdravev, "Pattern recognition and natural language processing: State of the art," *Tem Journal*, vol. 5, no. 2, pp. 236–240, 2016.
- [16] W. Nam and B. Jang, "A survey on multimodal bidirectional machine learning translation of image and natural language processing," *Expert Systems with Applications*, p. 121168, 2023.
- [17] K. M. N. Win, Z. Z. Hnin, Y. M. K. K. Thaw, *et al.*, "Review and perspectives of natural language processing for speech recognition," *International Journal Of All Research Writings*, vol. 1, no. 10, pp. 112–115, 2020.
- [18] L. Rajput and S. Gupta, "Sentiment analysis using latent dirichlet allocation for aspect term extraction," *Journal of Computers, Mechanical and Management*, vol. 1, no. 2, pp. 30–35, 2022.
- [19] J. Kaur, P. Verma, and S. Bajoria, "Sashakt: A job portal for women using text extraction and text summarization," *Journal of Computers, Mechanical and Management*, vol. 1, no. 2, pp. 22–29, 2022.
- [20] S. S. Priya, P. Rachana, and D. Chellani, "Augmented reality and speech control from automobile showcasing," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1703–1708, IEEE, 2022.
- [21] M. Yan, P. Castro, P. Cheng, and V. Ishakian, "Building a chatbot with serverless computing," in *Proceedings of the 1st International Workshop on Mashups of Things and APIs*, pp. 1–4, 2016.

- [22] M. Javed and L. Xudong, "Sos intelligent emergency rescue system: Tap once to trigger voice input," in *Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence*, pp. 187–193, 2020.
- [23] L. Qiu and I. Benbasat, "Online consumer trust and live help interfaces: The effects of text-to-speech voice and three-dimensional avatars," *International journal of human-computer interaction*, vol. 19, no. 1, pp. 75–94, 2005.
- [24] C. K. Mishra, "Digital marketing: Scope opportunities and challenges," *Promotion and Marketing Communications*, p. 115, 2020.
- [25] R. Khan, P. K. Sharma, S. Raj, S. K. Verma, and S. Katiyar, "Voice based e-mail system using artificial intelligence," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 9, no. 3, 2020.
- [26] A. J. Cerminara, "Stranger danger: The perils of cold-contact emailing to market nonfiction books," 2015.
- [27] P. Sivakumar, R. Vedachalam, V. Kannusamy, A. Odayappan, R. Venkatesh, P. Dhoble, F. Moutappa, and S. Narayana, "Barriers in utilisation of low vision assistive products," *Eye*, vol. 34, no. 2, pp. 344–351, 2020.
- [28] Y. D. Sapkota, S. Marmamula, and T. Das, "Population-based eye disease studies," in *South-East Asia Eye Health: Systems, Practices, and Challenges*, pp. 109–121, Springer, 2021.
- [29] H. Bhuiyan, A. Ashiquzzaman, T. I. Juthi, S. Biswas, and J. Ara, "A survey of existing e-mail spam filtering methods considering machine learning techniques," *Global Journal of Computer Science and Technology*, vol. 18, no. 2, pp. 20–29, 2018.
- [30] L. Manovich, *Software takes command*. Bloomsbury Academic, 2013.
- [31] H. Petrie, C. Harrison, and S. Dev, "Describing images on the web: a survey of current practice and prospects for the future," *Proceedings of Human Computer Interaction International (HCII)*, vol. 71, no. 2, 2005.
- [32] U. Sarwar and E. Eika, "Towards more efficient screen reader web access with automatic summary generation and text tagging," in *Computers Helping People with Special Needs: 17th International Conference, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part I 17*, pp. 303–313, Springer, 2020.
- [33] V. Sorge, C. Chen, T. Raman, and D. Tseng, "Towards making mathematics a first class citizen in general screen readers," in *Proceedings of the 11th web for all conference*, pp. 1–10, 2014.
- [34] S. Sandhya and K. S. Devi, "Accessibility evaluation of websites using screen reader," in *2011 7th International Conference on Next Generation Web Services Practices*, pp. 338–341, IEEE, 2011.
- [35] K. Williams, T. Clarke, S. Gardiner, J. Zimmerman, and A. Tomasic, "Find and seek: Assessing the impact of table navigation on information look-up with a screen reader," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 12, no. 3, pp. 1–23, 2019.
- [36] S. C. Baker, "Making it work for everyone: Html5 and css level 3 for responsive, accessible design on your library's web site," *Journal of Library & Information Services in Distance Learning*, vol. 8, no. 3-4, pp. 118–136, 2014.
- [37] R. Larsen, *Mastering SVG: Ace web animations, visualizations, and vector graphics with HTML, CSS, and JavaScript*. Packt Publishing Ltd, 2018.
- [38] D. Freitas and G. Kouroupetroglou, "Speech technologies for blind and low vision persons," *Technology and Disability*, vol. 20, no. 2, pp. 135–156, 2008.
- [39] M. C. Domingo, "An overview of the internet of things for people with disabilities," *journal of Network and Computer Applications*, vol. 35, no. 2, pp. 584–596, 2012.
- [40] M. Klasson, C. Zhang, and H. Kjellström, "A hierarchical grocery store image dataset with visual and semantic labels," in *2019 IEEE winter conference on applications of computer vision (WACV)*, pp. 491–500, IEEE, 2019.
- [41] V. V. Riabov, "Smtpl (simple mail transfer protocol)," *River College*, 2005.
- [42] J. Postel, "Simple mail transfer protocol," tech. rep., 1982.
- [43] J. Rhoton, *Programmer's guide to internet mail: SMTP, POP, IMAP, and LDAP*. Digital Press, 1999.
- [44] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel, "A survey on essential components of a self-sovereign identity," *Computer Science Review*, vol. 30, pp. 80–86, 2018.
- [45] A. Sadeghi-Nasab and V. Rafe, "A comprehensive review of the security flaws of hashing algorithms," *Journal of Computer Virology and Hacking Techniques*, vol. 19, no. 2, pp. 287–302, 2023.
- [46] R. Biddle, S. Chiasson, and P. C. Van Oorschot, "Graphical passwords: Learning from the first twelve years," *ACM Computing Surveys (CSUR)*, vol. 44, no. 4, pp. 1–41, 2012.
- [47] P. Gupta and S. Kumar, "A comparative analysis of sha and md5 algorithm," *architecture*, vol. 1, no. 5, 2014.