

# Voice Based E-Mail System using Artificial Intelligence



Rijwan Khan, Pawan Kumar Sharma, Sumit Raj, Sushil Kr. Verma, Sparsh Katiyar

**Abstract:** One of the mostly used forms of communication among the people is Email. Lot of confidential and urgent information is exchanged over emails in today's time. There are about 253 million visually impaired people worldwide. These visually impaired people are facing a problem of communication. Since, technology is growing day by day these types to visually challenged people feel that they are more challenged. So authors proposed a Voice based Email System using AI that will make email system very easily accessible to visually challenged people and also help society. Accessibility is the most important feature that is considered while developing this system. Any system is called accessible only if both able and disable people can use it easily.

**Keywords:** Speech recognition, Text to speech, Voice mail, visually challenged people.

## I. INTRODUCTION

The technologies are growing very fast day by day this has made the lifestyle of people so easy as mostly all work can be done in less amount of time with accuracy and efficiency. Communication is one of those fields that have grown to next level with the advancement in technology and the availability of Internet. Technologies have made communication so easy that distance has become a negligible parameter in communication [9]. When we think of communication using internet, the first thing that we come in our mind is communication via email. Email is one of the most reliable ways for exchange of some important information and also email is used worldwide, but for accessing internet a person must be able to see. There are millions of people who are blind or visually challenged who are not able to see the screen; keyboards therefore they are not able to access the internet [12]. In this way, they are very

far away from email communication and internet world. These blind people cannot use the existing email system, they cannot send, receive emails and cannot read the information shared through email; therefore the existing systems are not easily accessible to them.

To access the internet the person must be able to read what is written on the screen so, this makes internet useless technology for visually challenged people. There is only one way by which a visually challenged person can send an E-mail is, they have to tell the entire content of the mail to a third person so that the third person can compose the mail and send on the behalf of the visually challenged person [10]. But this approach does not take us to the solution of the problem. Every time finding a third person is not possible for a visually challenged person and also sometimes the content can be personal, for maintaining the Integrity of the Specifications. Therefore, for helping these people and developing society authors have come up with this idea that helps a visually challenged person by providing ability to send and receive emails throw voice commands without using any keyboard and visual thing.

**Artificial Intelligence for Speech Recognition:** Artificial intelligence (AI) is a technology used for creating intelligent systems and machines that simulate human intelligence. Some artificial intelligence applications include various expert systems, natural language processing (NLP), machine vision and speech recognition. To understand and analyse human language such as English by extracting metadata from keywords, emotions, relations and concepts is Natural Language Processing [13, 14, 15].

## II. EXISTING SYSTEM

According to Email Statistics Report, 2014-2018 by a technology market research firm Palo Alto, CA, USA, there are a total of 4.1 billion email accounts created until 2014 to over 5.2 billion accounts in end of 2018 and making it one of the most used form of communication[1].

The research, by the Vision Loss Expert Group (VLEG), shows that worldwide 253 million people are either blind or visually challenged that is, around 253 million people are not aware of how to use Internet or E-mail [6].

Existing systems of today are basically applications that provide accessing and managing of emails benefits to its users via web facilities [7]. Making email widely used communication form. The existing systems do not support any voice commands or audio facilities and therefore it is not suitable for visually challenged people.

Also, there are various existing search engine which take request in form of text from user and retrieve the relevant documents from server and responds by displaying it in the form of text which is not possible for visually challenged people [8].

Revised Manuscript Received on February 05, 2020.

\* Correspondence Author

**Rijwan Khan\***, Department of Computer Science and Engineering, ABES Institute of Technology, Ghaziabad, Affiliated to AKTU Lucknow, India. E-mail: rijwankhan786@gmail.com

**Pawan Kumar Sharma**, Department of Applied Science, Krishna Engineering College, Ghaziabad, Affiliated to AKTU Lucknow, India. E-mail: vashistha23@gmail.com

**Sumit Raj**, Department of Computer Science and Engineering, ABES Institute of Technology, Ghaziabad, Affiliated to AKTU Lucknow, India. E-mail: sumitraj20399@gmail.com

**Sushil Kr. Verma**, Department of Computer Science and Engineering, ABES Institute of Technology, Ghaziabad, Affiliated to AKTU Lucknow, India. E-mail: sushilssm92@gmail.com

**Sparsh Katiyar**, Department of Computer Science and Engineering, ABES Institute of Technology, Ghaziabad, Affiliated to AKTU Lucknow, India. E-mail: sparshktyr@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Although the existing web browsers can play audios and videos, but for that also the user has to request by typing some text to search after that the user will be able to play the audio and videos by using Graphical User Interfaces (GUI) [11].

The existing email system do not provide this facility and are not accessible by blind people therefore the system which is developing now completely different from exiting system [2].

III. PROPOSED SYSTEM

The Proposed system will make the email system very easily accessible to visually challenged people and also help society. Authors proposed the system keeping one idea in mind that it should be easily accessible for all kind of persons.

Any designed web application is claimed to be excellently accessible, if it is used by any person, visually challenged or not with efficient manner.

As critical the present system that prioritizes user-friendliness of traditional users, proposed system focuses on user-friendliness of all kinds' of individuals, together with traditional folks and visually weakened folks additionally as illiterate people.

In this system, the pc is going to be prompting the user to perform specific operations to avail various services and if the user has to access the various services then he/she has to perform that operation. Firstly, the user will have to register in application system through the registration form. The user goes to be assisted through voice commands whereas registering all the mandatory fields to be stuffed are going to be scan by website; once the user would speak it would get written automatically.

After successfully registering, the user can log in by speaking the Username and Password when prompted by the system, this username and password will then be converted from speech to text and then the user will be authenticated by verifying the credentials with the database. Users can access various sections like Compose, Inbox, and Sent Mail after successful login.

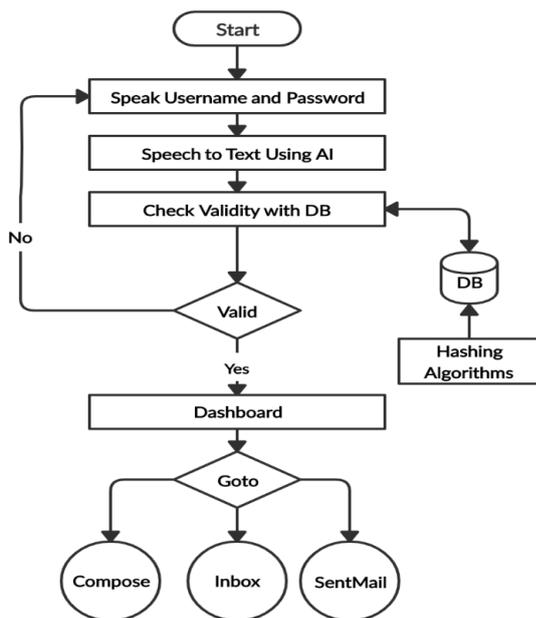


Fig. 1. Flow chart of User login and accessing system

The main advantage of this system is that the use of a keyboard is completely eliminated; the user will have to respond through voice only.

The application is totally voice-based allowing blind people to send and receive emails easily. It converts the user spoken voice into text and the text to voice and performs the action accordingly.

IV. DESIGN OF PROPOSED SYSTEM

This project is designed by dividing it into the following three phases:

1. **UI design:** In this phase, the UI or the user interface of the project is developed. That is the designing of the web pages in which the user will use to interact. The user interface is designed using HTML5 and CSS3.

2. **Database design:** Database is important in every project since it is responsible for storing of data and user credentials. That is, database mainly aims User authentication and storing all the user mails. The database design will include various tables' creation for storing emails.

3. **System design:** The system will consist all the modules such as: TTS (Text to Speech) and STT (Speech to Text) module, Mail programming module (Compose, Inbox, and Sent Mail).

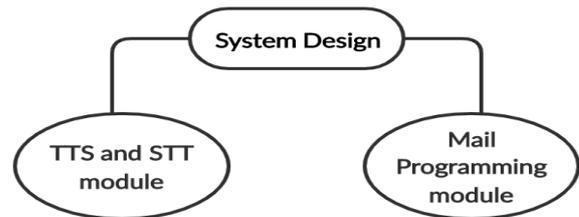


Fig. 2. System Design Modules

A. Mail Programming Module:

Email is appearing as one of the most valuable services on the internet today. Many of the internet systems use SMTP as a method to transfer mail from one user to another. SMTP is a sending protocol and is used to send the mail while POP (post office protocol) or IMAP (internet message access protocol) are used to retrieve those emails at the receiver's side.

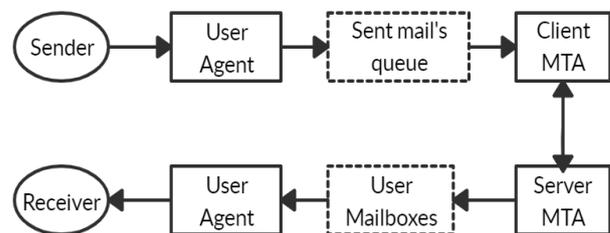


Fig. 3. Model of SMTP system

**Sending Email:** When an email is distributed across, it will contain certain things like header and its own body. A sequence of responses requested message is line up between shopper and server in sending an email.

How header and body are differing from each other, header will terminate when a null line is in it. The message body contain the particular data browse by receipt. In body every data is taken after null line.

**Receiving Email:** In a specific time interval user agent at server aspect checks mailboxes. If in this time during any data receive it alert the user immediately. When user opens its email he/she is availed to scan emails with some information as in its subject line.

## V. EXPERIMENTAL SETUP

### A. Speech to Text by using Artificial Intelligence

Audio to text can be converted by applying a simple to use API involving powerful neural network models provided to developers by Google Cloud Speech to Text. The API acknowledges a hundred and twenty languages & variants to support world user base. You'll alter voice command & management, transcribe of audio from decision centers, and more. It will method period streaming or pre-recorded audio, mistreatment Google's machine learning technology [4].

### B. Text-to-speech unification

This is an automated conversion of a text into speech. This technique resembles, as closely as doable, a verbalizer of the language reading that text. Text-to-speech synthesizer (TTS) is the technology that lets laptops speak to you. With some good mathematical models TTS engine availed to pre processes the text and synthesizes it. The TTS engine typically generates sound information in AN audio format due to the output.

### C. Text-To-Speech unification System Structure

Steps involved in text to speech are :-text is received by TTS and analyze for remodeling into an accent description, generate the prosody . Then in an exceedingly any step, it generates the prosody. From the data currently on the market, it will turn out speech signal. The text-to-speech's structure unification may be diminished into major modules:

### D. Processing of Natural Language module

It produces a written text of the scanned text, at the side of prosody.

### E. Processing of Digital Signal module

It transforms representative data it be subject from information science into hearable and comprehensible speech.

The operations of Processing of Natural Language module are as follows:

•**Text Review:** First, text is divided into tokens. The token-to-word transformation creates the writing kind of token. For token "Mr" the writing kind "Mister" is made by enlargement, the token "11" gets the writing kind "eleven" and "1979" is remodeled to "nineteen seventy-nine".

•**Application of Pronunciation Rules:** Pronunciation rules are applied after the completion of text analysis. Letters can't be reworked 1:1 into phonemes as a result of the correspondence isn't forever parallel. In sure environments, one letter will correspond to either no sound (for example, "" in "taught") or many phonemes ("n" in "known"). Additionally, many letters will correspond to one sound ("wh" in "which").

• **Prosody Generation** The prosody is created after the accent has been decided . The degree of openness of a TTS system relies on speech factors like inflection modelling (phrasing and accentuation), amplitude modelling and length modelling (including the length of sound and therefore the length of pauses, that determines the length of the linguistic unit and therefore the tempos of the speech) [3].

## VI. AUTHENTICATION & SECURITY

Authentication involves provide users the credentials, such as username and password, and verifying that user provides correct username and password every time whenever application needs to be accessed. Therefore, these credentials need to be stored in database for future comparisons. Authentication will be used in our system to verify the users.

**Storing Password directly can be Risky:** One simple method to store password is to simply map them by creating a table in database. Whenever a user login request arrives, the server will be called for authentication with a payload contain a username and a password. Then, these credentials will be matched with the password stored in database. If the match gives positive results then the user will be given access to the application.

When passwords are stored in clear text format it can be risky and will be open for attackers. If some attacker can steal the password then he can access the users account. One way to store password securely is to store it by transforming it into a form that cannot be converted into original password again. The process is referred as hashing.

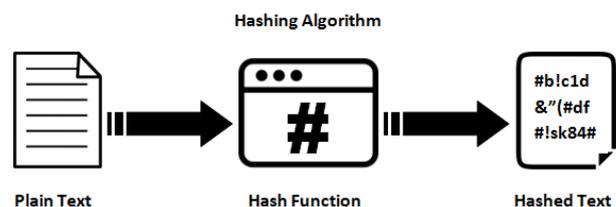


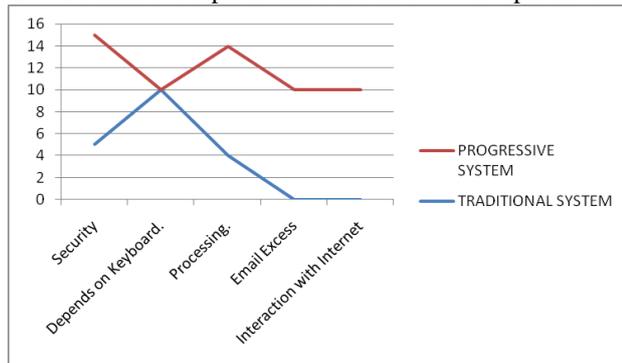
Fig. 4. Hashing Algorithm

There are various algorithms for hashing mostly used are Message Digest Algorithms (MD5) and Secure Hash Algorithms (SHA) [5].

## VII. COMPARISON & RESULT ANALYSIS

SR. NO.	TRADITIONAL SYSTEM	PROGRESSIVE SYSTEM
1.	Less security.	High Security provided.
2.	Depends on Keyboard.	The entire structure is based on IVR-interactive voice response.
3.	Slow processing.	Faster and more efficient
4.	The disables cannot use the normal mail system.	The disables can use the normal mail system.
5.	Blind people are not being able to interact with the web based email system.	They will be able to interact with the web based email system.

On 10 scales there parameters are taken in comparison.



**Fig. 5. Result with comparison study**

## VIII. FUTURE SCOPE

There is wide future scope of this system many enhancements can be done in the system such as including different languages, including functionality of accessing the deleted mails and spam mails. Also, this system can be enhanced such that it can also send attachments which are more beneficial for visually challenged people.

This system can be made available to all regional people who are not educated enough and inclusion of different languages will make this system easily accessible. Furthermore sign language system can also be integrated with the system to make the system more scalable and robust.

## IX. CONCLUSION

This paper proposes a system that will be beneficial for society by allowing disable people also to grow along with society. This project makes visually challenged people able enough to be part of growing digital India by allowing them to communicate via internet and also making life of such people much easier. This system overcomes many drawbacks that were faced by visually challenged people such as sending and receiving emails. Success of this project can make an impact on developers motivating them to make something useful that can help visually challenged or blind people

## REFERENCES

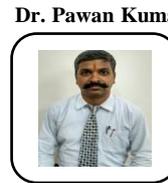
1. G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15-64.
2. The Radicati website Email Statistics Report, 2014-2018 Available:<http://www.radicati.com/wp/wpcontent/uploads/2014/01/EmailStatisticsReport-20142018-Executive-Summary.pdf>.
3. Ingle, Pranjali, Harshada Kanade, and Arti Lanke. "Voice based e-mail System for Blinds." *International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)* (2016): 25-30.
4. Isewon, Itunuoluwa, O. J. Oyelade, and O. O. Oladipupo. "Design and implementation of text to speech conversion for visually impaired people." *International Journal of Applied Information Systems* 7.2 (2012): 26-30.
5. Shakhovska, N., O. Basystiuk, and K. Shakhovska. "Development of the speech-to-text chatbot interface based on Google API." *CEUR Workshop Proceedings*. Vol. 2386. 2019.
6. [https://www.juniper.net/documentation/en\\_US/junos/topics/concept/ipsec-authentication-solutions.html](https://www.juniper.net/documentation/en_US/junos/topics/concept/ipsec-authentication-solutions.html).
7. Arlinghaus, Robert, et al. "Understanding the complexity of catch-and-release in recreational fishing: an integrative synthesis of global knowledge from historical, ethical, social, and biological perspectives." *Reviews in Fisheries Science* 15.1-2 (2007): 75-167.
8. Cole, Ron, et al. "The challenge of spoken language systems: Research directions for the nineties." *IEEE transactions on Speech and Audio processing* 3.1 (1995): 1-21.

9. Lazar, Jonathan, et al. "What frustrates screen reader users on the web: A study of 100 blind users." *International Journal of human-computer interaction* 22.3 (2007): 247-269.
10. December, John. "Units of analysis for Internet communication." *Journal of Computer-Mediated Communication* 1.4 (1996): JCMC143.
11. Pathan, Naziya, et al. "V-Mail (Voice Based E-Mail Application)." (2019).
12. Symons, Gary Mark, and Kirk David Symons. "Digital media editing interface using a supercursor for selecting media clips for editing." U.S. Patent No. 8,527,879. 3 Sep. 2013.
13. Shabana, T., et al. "Voice based email system for blinds." *International Journal of Advance Foundation And Research In Science & Engineering (IJAFRSE)* Volume 1 (2015).
14. Al Smadi, Takiaddin, et al. "Artificial intelligence for speech recognition based on neural networks." *Journal of Signal and Information Processing* 6.02 (2015): 66.
15. Senders, Joeky T., et al. "Natural and artificial intelligence in neurosurgery: a systematic review." *Neurosurgery* 83.2 (2018): 181-192.
16. Collobert, Ronan, et al. "Natural language processing (almost) from scratch." *Journal of machine learning research* 12.Aug (2011): 2493-2537.

## AUTHORS PROFILE



**Dr. Rijwan Khan** is B.Tech (CSE), M.Tech (CSE) and Ph.D in Computer Engineering. He has total 13 years of teaching experience. Now working as a Head of Department in ABES Institute of Technology. His area of research is software testing, soft computing, nature inspired algorithms. He published more than 25 research papers in different journals. He is reviewer of more than 6 different Scopus and SCI indexed journals like journals of Inderscience, Springer, and Elsevier etc. He is authors of 3 books on C Programming, Data Structure using C, and Operating System and one book chapter in springer book.



**Dr. Pawan Kumar Sharma** has completed his Ph.D. in Reliability Theory (Operations Research). He had done his research work on Operations Research and Graph Theory where he published research papers. He also published two research papers on Fuzzy graph & Image Processing. Currently he is actively involved in Birla Institute of Technology & Science, Pilani (Rajsthan) as a Guest Faculty for working professionals & INMANTEC Institutions Ghaziabad. He has more than 14 years of teaching learning & research experience of graduate & post graduate classes. Proven Expertise in designing/restructuring course outline, material and assessment criteria and implementing innovative instruction methodologies to match international standards.



**Sumit Raj** is an undergraduate computer science & engineering student pursuing B.Tech at ABES Institute of Technology, Ghaziabad. His Area of interest is Artificial Intelligence, Machine Learning and Soft Computing. He done some training in Machine learning.



**Sushil kr Verma** is an undergraduate computer science & engineering student pursuing B.Tech at ABES Institute of Technology, Ghaziabad. His Area of interest is Artificial Intelligence, Machine Learning and Soft Computing. He done some training in Machine learning.



**Sparsh Katiyar** is an undergraduate computer science & engineering student pursuing B.Tech at ABES Institute of Technology, Ghaziabad. His Area of interest is Artificial Intelligence, Machine Learning and Soft Computing. He done some training in Machine learning.

