



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 12, December 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.165



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Voice Based E-Mail for the Visually Impaired

Aadil Ashraf, Prince Sood

PG Student, Dept. of CSE, Swami Vivekananda Institute of Engineering and Technology, Chandigarh-Patialia
Highway (Banur), India

Assistant Professor, Dept. of CSE, Swami Vivekananda Institute of Engineering and Technology, Chandigarh-
Patialia Highway (Banur), India

ABSTRACT: Internet is one of the basic luxuries for daily living. Every person is using the facts and information on internet. On the other hand, blind people face difficulty in accessing the text resources. The advancement in computer based accessible systems has opened up many avenues for the visually impaired across a wide majority of the globe. Audio feedback based virtual environment like, the screen readers have helped blind people to access internet applications immensely. However, the visually challenged people find it very difficult to utilize this technology because of the fact that using them requires visual perception. Even though many new advancements have been implemented to help them use the computers efficiently no naïve user who is visually challenged can use this technology as efficiently as a normal naïve user can do that is unlike normal users, they require some practice for using the available technologies. In this project, the voicemail system architecture that can be used by a blind person to access e-mails easily and efficiently. The contribution made by project has enabled the blind people to send and receive voice-based e-mail message. The proposed system GUI has been evaluated against the GUI of a traditional mail server and found that the proposed architecture performs much better than that of the existing GUIs. In this project, the use of voice to text and text to voice technique access for blind people. Also, this system can be used by any normal person also for example the one who is not able to read. The system is completely based on interactive voice response which will make it efficient.

I. INTRODUCTION

We have seen that the introduction of Internet has revolutionized many fields. Internet has made life of people so easy that people today have access to any information they want easily. Communication is one of the main fields highly changed by Internet. E-mails are the most dependable way of communication over Internet, for sending and receiving some important information. But there is a certain norm for humans to access the Internet and the norm is you must be able to see. But

there are also differently able people in our society who are not gifted with what you have. There are some visually impaired people or blind people who can't see things and thus can't see the computer screen or keyboard. A survey has shown that there are more than 240 million visually impaired people around the globe. That is, around 240 million people are unaware of how to use Internet or E-mail. The only way by which a visually challenged person can send an E-mail is, they have to speak the entire content of the mail to another person (not visually challenged) and then that third person will compose the mail and send on the behalf of the visually challenged person. But this is not a right way to deal with the problem. It is very unlikely that every time a visually impaired person can find someone for help. Although for these reasons the visually impaired people are criticized by our society].

II. PURPOSE OF THE PROJECT

This project proposes a python-based application, designed specifically for visually impaired people. This application provide a voice based mailing service where they could read and send mail on their own, without any guidance through their Email accounts. The VMAIL system can be used by a blind person to access mails easily and adeptly. Hence dependence of visually challenged on other individual for their activities associated to mail can be condensed.

The application will be a python-based application for visually challenged persons using IVR- Interactive voice response, thus sanctioning everyone to control their mail accounts using their voice only and to be able to read, send, and perform all the other useful tasks. The system will ask the user with voice commands to perform certain action and the user will respond to it. The main advantage of this system is that use of keyboard is completely eliminated, the user will have to respond through voice only.

III.SYSTEM DEVELOPMENT

There are a complete number of 4.1 billion email accounts made until 2014 and a there will be evaluated 5.2 billion records by end of 2018. This makes messages the most utilized type of correspondence. The most generally perceived mail benefits that we use in our regular day to day existence can't be used by ostensibly tried people. This is on the grounds that they don't give any office so the individual in front can hear out the substance of the screen. As they can't imagine what is now present on screen they can't make out where to click so as to play out the necessary tasks. For an outwardly tested individual utilizing a PC just because isn't that helpful for what it's worth for an ordinary client despite the fact that it is easy to understand. In spite of the fact that there are many screen readers accessible then likewise these individuals face some minor troubles.

Screen readers read out whatever substance is there on the screen and to play out those activities the individual should utilize console alternate routes as mouse area can't be followed by the screen reader. This implies two things; one that the client can't utilize mouse pointer as it is totally awkward if the pointer area can't be followed and second that client ought to be knowledgeable with the console concerning where every single key is found. A client is new to PC can accordingly not utilize this administration as they don't know about the key areas. Another disadvantage that sets in is that screen reader read out the substance in successive way and subsequently client can make out the substance of the screen just on the off chance that they are in essential HTML position.

Therefore the new propelled pages which don't follow this worldview so as to make the site more easy to use just make additional issues for these individuals. Moreover the systems that do use only voice for interaction between the user and the system don't have good voice transcription. All these are a few downsides of the present framework which we will defeat in the framework we are creating.

IV.PROPOSED SYSTEM

The proposed system is based on a completely novel idea and is nowhere like the existing mail systems. The most important aspect that has been kept in mind while developing the proposed system is accessibility. A web system is said to be perfectly accessible only if it can be used efficiently by all types of people whether able or disable. The current systems do not provide this accessibility. Thus the system we are developing is completely different from the current system. Unlike current system which emphasizes more on user friendliness of normal users, our system focuses more on user friendliness of all types of people including normal people visually impaired people as well as illiterate people. The complete system is based on IVR- interactive voice response.

When using this system the computer will be prompting the user to perform specific operations to avail respective services and if the user needs to access the respective services then he/she needs to perform that operation. One of the major advantages of this system is that user won't require to use the keyboard. All operations will be based on voice commands. This system will be perfectly accessible to all types of users as it is just based on simple speech inputs and there is no need to remember keyboard shortcuts. Also because of IVR facility those who cannot read need not worry as they can listen to the prompting done by the system and perform respective actions.

V. COMPONENT DESCRIPTION

- Detailed description of IVR.
- Increase first contact resolution
- Reduce operational costs
- Increase professionalism
- Increase customer satisfaction

VI. WORKING OF SPEECH RECOGNITION

Speech recognition is the inter-disciplinary sub-field of computational linguistics that develops methodologies and technologies that enables the recognition and translation of spoken language into text by computers. It is also known as & automatic speech recognition (ASR), & computer speech recognition or just "speech to text" (STT). It incorporates knowledge and research in the linguistics, computer science, and electrical engineering fields. Some speech recognition systems require & training (also called & enrollment ;) where an individual speaker reads text or

isolated vocabulary into the system. The system analyzes the person's specific voice and uses it to fine-tune the recognition of that person's speech, resulting in increased accuracy. Systems that do not use training are called & speaker independent & systems.

Systems that use training are called & speaker dependent & Speech recognition applications include voice user interfaces such as voice dialing (e.g. Call home), call routing (e.g. I would like to make a collect call), domestic appliance control, search (e.g. find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g. a radiology report), speech-to-text processing (e.g., word processors or emails), and aircraft (usually termed Direct Voice Input).

VII. SPEECH TO TEXT CONVERTER

The process of converting spoken speech or audio into text is called speech to text converter. The process is usually called speech recognition. The Speech recognition is used to characterize the broader operation of deriving content from speech which is known as speech understanding. We often associate the process of identifying a person from their voice, which is voice recognition or speaker recognition so it is wrong to use this term for it. As shown in the above block diagram speech to text converters depends mostly on two models 1. Acoustic model and 2. Language model. Systems generally use the pronunciation model. It is really imperative to learn that there is nothing like a universal speech recognizer. If you want to get the best quality of transcription, you can specialize the above models for the any given language communication channel.

Likewise another pattern recognition technology, speech recognition can also not be without error. Accuracy of speech transcript deeply relies on the voice of the speaker, the characteristic of speech and the environmental conditions. Speech recognition is a tougher method than what folks unremarkably assume, for a personality's being. Humans are born for understanding speech, not to transcribing it, and solely speech that's well developed will be transcribed unequivocally. From the user's purpose of read, a speech to text system will be categorized based in its use. To convert speech to on-screen text or a computer command, a computer has to go through several complex steps. When you speak, you create vibrations in the air. The analog-to-digital converter (ADC) translates this analog wave into digital data that the computer can understand. To do this, it samples, or digitizes, the sound by taking precise measurements of the wave at frequent intervals. The system filters the digitized sound to remove unwanted noise, and sometimes to separate it into different bands of frequency (frequency is the wavelength of the sound waves, heard by humans as differences in pitch). It also normalizes the sound, or adjusts it to a constant volume level. It may also have to be temporally aligned. People don't always speak at the same speed, so the sound must be adjusted to match the speed of the template sound samples already stored in the system's memory.

Next the signal is divided into small segments as short as a few hundredths of a second, or even thousandths in the case of plosive consonant sounds -- consonant stops produced by obstructing airflow in the vocal tract -- like p or t; The program then matches these segments to known phonemes in the appropriate language. A phoneme is the smallest element of a language -- a representation of the sounds we make and put together to form meaningful expressions. There are roughly 40 phonemes in the English language (different linguists have different opinions on the exact number), while other languages have more or fewer phonemes. The next step seems simple, but it is actually the most difficult to accomplish and is the focus of most speech recognition research. The program examines phonemes in the context of the other phonemes around them. It runs the contextual phoneme plot through a complex statistical model and compares them to a large library of known words, phrases and sentences. The program then determines what the user was probably saying and either outputs it as text or issues a computer command.

VIII. SPEECH SYNTHESIS

Speech synthesis is the synthetic production of speech. An automatic data handing out system used for this purpose is called as speech synthesizer, and may be enforced in software package and hardware product. A text-to-speech (TTS) system converts language text into speech, alternative systems render symbolic linguistic representations. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diaphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output.

Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output. The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood clearly. An intelligible text to speech program permits individual with ocular wreckage or reading disabilities to concentrate to written words on a computing device. Several computer operational systems have enclosed speech synthesizers since the first nineteen nineties years.

The text to speech system is consisting of 2 parts:-front-end and a back-end. The front- end consist of 2 major tasks. Firstly, it disciple unprocessed text containing symbols like numbers and abstraction into the equivalent of written out words. This method is commonly known as text, standardization, or processing. Front end then assigns spoken transcriptions to every word, and divides and marks the text into speech units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end—often referred to as the synthesizer—then converts the symbolic linguistic representation into sound. In certain systems, this part includes the computation of the target prosody (pitch contour, phoneme durations), which is then imposed on the output speech. Text-to-speech (TTS) is a type of speech synthesis application that is used to create a spoken sound version of the text in a computer document, such as a help file or a Web page. TTS can enable the reading of computer display information for the visually challenged person, or may simply be used to augment the reading of a text message. Current TTS applications include voice-enabled e-mail and spoken prompts in voice response systems. TTS is often used with voice recognition programs. There are numerous TTS products available, including Read Please 2000, Pro verbe Speech Unit, and Next Up Technology Text Aloud. Lucent, Elan, and AT&T each have products called “Text-to-Speech”.

In addition to TTS software, a number of vendors offer products involving hardware, including the Quick Link Pen from WizCom Technologies, a pen-shaped device that can scan and read words; the Road Runner from Ostrich Software, a handheld device that reads ASCII text; and DecTalk TTS from Digital Equipment, an external hardware device that substitutes for a sound card and which includes an internal software device that works in conjunction with the PC's own sound card.

IX. DESIGN

Phase I: The tasks that can be performed using the program developed will be prompted using the voice prompt. In background python module pyttsx3 is used for text to speech conversion. User will be asked to provide input for the following tasks written below.

The input is expected in the form of speech by the user which will be converted to text by the Google speech application interface in python and accordingly tasks will be performed.

- Login to their Gmail account.
- Send e-mail through Gmail.
- Read e-mail through Gmail.

Phase II: In phase-2 of our program the user will give speech input to the system.

This speech input will be handled by speech_recognition module. It is a python library which is used to handle the voice requests and it converts speech into text. Now after receiving input from the user speech to text converter will save the response in respective variables used in the script and based on their value it will further enter into respective modules.

Phase III: In this phase our program will handle the requests by the user. Based on the speech input given by the user it will launch the modules. Login to G-mail account:- This module will handle the request by user to login in their g-mail account. This module will make the connection with the user's gmail account based on the credentials provided through voice input. This module's script designed as such it will prompt user to enter their g-mail username and password and then it will use selenium web-driver to automate the task for the user and as a result connection will be made. Send E-mail through G-mail:- This module will handle the request by user to send email through their g-mail account. The python script for this module will prompt the user to enter their credentials and then it will make connection with their account. After the connection has been done it will further prompt the user to enter the receiver's account e- mail id and it will then allow the user to speak their message and it will repeat it for them and by saying ok it will send the mail. SMTP library in python is used for the above task. Read E-mail through G-mail:- This module will handle the request by user to read email through their g-mail account. The python script for this module will prompt the user to enter their credentials and then it will make connection with their account. After the connection has been done it will start fetching



the unread mails for the user and will speak it for them with the help of pyttsx3 or GTTS library in python for text to speech conversion.

X. IMPLEMENTATION

- Speech Recognition in Python
- Required Installation
- Text to Speech in Python
- Simple Mail Transfer Protocol
- Sending Email in Python using SMTPLIB
- Reading Email from Gmail using Python

XI. SOFTWARE REQUIREMENTS

- Python IDLE.
- Interpreters for scripts.
- Selenium Web driver in python.
- Google Speech-to-text and text-to-speech Converters.
- Pyttsx text to speech api in python.

XII. HARDWARE REQUIREMENTS

- Windows Desktop

XIII. CODE

```
import smtplib
from bs4 import BeautifulSoup
import email
import imaplib
from gtts import gTTS
import pygame
import os, time
tts = gTTS(text="Project: Voice based Email for blind", lang='en')
ttsname = ("name.mp3")
tts.save(ttsname)
music = pygame.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
login = os.getlogin()
print("You are logging from : " + login())
print("1. composed a mail.")
tts = gTTS(text="option 1. composed a mail.", lang='en')
ttsname = ("hello.mp3")
tts.save(ttsname)
music = pygame.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
print("2. Check your inbox")
tts = gTTS(text="option 2. Check your inbox", lang='en')
ttsname = ("second.mp3")
tts.save(ttsname)
music = pygame.media.load(ttsname, streaming=False)
music.play()
```

```
time.sleep(music.duration)
os.remove(ttsname)
tts = gTTS(text="Your choice ", lang='en')
ttsname = ("hello.mp3")
tts.save(ttsname)
music = pygamelet.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
r = sr.Recognizer()
with sr.Microphone() as source:
    print("Your choice:")
    audio = r.listen(source)
    print("ok done!!")
try:
    text = r.recognize_google(audio)
    print("You said : " + text)
except sr.UnknownValueError:
    print("Google Speech Recognition could not understand audio.")
except sr.RequestError as e:
    print("Could not request results from Google Speech Recognition service; {0}".format(e))
if text == '1' or text == 'One' or text == 'one':
    r = sr.Recognizer() # recognize
    with sr.Microphone() as source:
        print("Your message :")
        audio = r.listen(source)
        print("ok done!!")
    try:
        text1 = r.recognize_google(audio)
        print("You said : " + text1)
        msg = text1
    except sr.UnknownValueError:
        print("Google Speech Recognition could not understand audio.")
    except sr.RequestError as e:
        print("Could not request results from Google Speech Recognition service; {0}".format(e))
    mail = smtplib.SMTP('smtp.gmail.com', 587)
    mail.ehlo()
    mail.starttls()
    mail.login('aadilashraftestmail@gmail.com', 'TestPassword123')
    mail.sendmail('aadilashraftestmail@gmail.com', 'aadilashraftomyfirend@gmail.com', msg)
    print("Congrates! Your mail has send. ")
    tts = gTTS(text="Congrates! Your mail has send. ", lang='en')
    ttsname = ("send.mp3")
    tts.save(ttsname)
    music = pygamelet.media.load(ttsname, streaming=False)
    music.play()
    time.sleep(music.duration)
    os.remove(ttsname)
    mail.close()
if text == '2' or text == 'tu' or text == 'two' or text == 'Tu' or text == 'to' or text == 'To':
    mail = imaplib.IMAP4_SSL('imap.gmail.com', 993)
    unm = ('aadilashraftestmail@gmail.com')
    psw = ('TestPassword123')
    mail.login(unm, psw)
    stat, total = mail.select('Inbox')
    print("Number of mails in your inbox : " + str(total))
```

```
tts = gTTS(text="Total mails are :" + str(total), lang='en')
ttsname = ("total.mp3")
tts.save(ttsname)
music = pyglet.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
unseen = mail.search(None, 'UnSeen')
print("Number of UnSeen mails :" + str(unseen))
tts = gTTS(text="Your Unseen mail :" + str(unseen), lang='en')
ttsname = ("unseen.mp3")
tts.save(ttsname)
music = pyglet.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
result, data = mail.uid('search', None, "ALL")
inbox_item_list = data[0].split()
new = inbox_item_list[-1]
old = inbox_item_list[0]
result2, email_data = mail.uid('fetch', new, '(RFC822)')
raw_email = email_data[0][1].decode("utf-8")
email_message = email.message_from_string(raw_email)
print("From: " + email_message['From'])
print("Subject: " + str(email_message['Subject']))
tts = gTTS(text="From: " + email_message['From'] + " And Your subject: " + str(email_message['Subject']),
lang='en')
ttsname = ("mail.mp3")
tts.save(ttsname)
music = pyglet.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
stat, total1 = mail.select('Inbox')
stat, data1 = mail.fetch(total1[0], "(UID BODY[TEXT])")
msg = data1[0][1]
soup = BeautifulSoup(msg, "html.parser")
txt = soup.get_text()
print("Body :" + txt)
tts = gTTS(text="Body: " + txt, lang='en')
ttsname = ("body.mp3")
tts.save(ttsname)
music = pyglet.media.load(ttsname, streaming=False)
music.play()
time.sleep(music.duration)
os.remove(ttsname)
mail.close()
mail.logout().
```

XIV. ADVANTAGES

- The disabilities of visually impaired folks are thrashed.
- This method makes the disabled folks desire a standard user.
- Completely voice based, wiped out the use of keyboard and mouse.
- Efficient and robust

- This design also scales back psychological feature load taken by blind to recollect and kind characters mistreatment keyboard.
- User friendly

XV. CONCLUSION

This e-mail system can be used by any user of any age group with ease of access. It has highlight of speech to content just as content to speech with discourse reader which makes planned framework to be taken care of by out wardly hindered individual too. Now the visually impaired people can send and receive mails with a lot of ease only through voice commands without making any use of a keyboard or any mouse. It has helped eradicate the difficulties that the blind people face and made them more the normal individuals.

It has wiped out the idea of utilizing console easy routes alongside screen readers which will help decreasing the intellectual heap of recollecting console alternate ways. Also any non-sophisticated user who does not know the position of keys on the keyboard need not bother as keyboard usage is eliminated. Instructions given by the IVR accordingly to get the respective services offered.

XVI. FUTURE SCOPE

It is a observation that about 70% of total blind population across the world is present in INDIA. This depict the voice message engineering utilized by daze individuals to get to E-mail and multimedia elements of working framework effectively and efficiently. Separated from this the uneducated, crippled and daze individuals will too be able to send sends in their local dialects. This design will likewise decrease intellectual burden taken by blinds to recall and type characters utilizing console. Advances in technology will allow consumers and business to implement speech recognition systems at a relatively low cost and efficiently. Apart from this the system can be enhanced to help the illiterate people by making speech recognition possible in their native languages.

ACKNOWLEDGMENT

Thanks to my Guide and my family member who always support, help and guide me during my work. Special thanks to my father who always support my innovative ideas and also to my spouse.

REFERENCES

1. Jagtap Nilesh, Pawan Alai, Chavhan Swapnil and Bendre M.R.. "Voice Based System in Desktop and Mobile Devices for Blind People". In International Journal of Emerging Technology and Advanced Engineering (IJETA), 2014 on Pages 404-407 (Volume 4, issue 2).
2. Ummuhanyifa U., Nizar Banu P K , "Voice Based Search Engine and Web page Reader". In International Journal of Computational Engineering Research (IJCER). Pages 1-5.
3. The Radicati website. [Online]. Available: <http://www.radicati.com/wp/wp-content/uploads/2014/01/EmailStatistics-Report-2014-2018-Executive-Summary.pdf>.
4. Geeks for geeks - <https://www.geeksforgeeks.org/project-idea-voice-based-email-visually-challenged/>
5. K. Jayachandran and P. Anbumani "Voice Based Email for Blind People" in International Journal of Advance Research, Ideas and Innovations in Technology (IJARIIT), 2017 on Pages 1065-1071
6. Pranjali Ingle, Harshada Kanade and Arti Lanke "Voice based e-mail system for blinds" in International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 2016 on Pages 25-30 (Volume 3, issue 1)
7. G. Broll, S. Keck, P. Holleis and A. Butz, "Improving the Accessibility of NFC/RFID-based Mobile Interaction through Learnability and Guidance", International Conference on Human-Computer Interaction with Mobile devices and services, vol. 11, (2009).

BIOGRAPHY

Aadil Ashraf is currently a Research Scholar at Swami Vivekanand Institute Of Engineering And Technology. He is currently doing masters in computer science and engineering and is involved in many research projects. His choicest areas of research are web designing, networking, image processing, data mining etc.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi[®]
cross **ref**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details