G H Raisoni College of Engineering and Management
Wagholi, Pune
Department of Artificial Intelligence
(An Autonomous Institute Affiliated to SPPU, Pune)

MINI PROJECT ON

"PDF TO AUDIO CONVERTER WITH LANGUAGE TRASLATION"

Submitted By

Name of Student : Devashri Nandkumar Bhosale
Roll No : A19
S.Y.BTech(AI)

Department of Artificial Intelligence

G.H.Raisoni College of Engineering and Managment
Wagholi, Pune – 412207
2021-22

1. INTRODUCTION

In this paper, the PDF to Audio Converter is proposed. It provides an alternative way to access the books and any pdf file for lazy, readers, and others. Using this PDF to Audio Converter the user will be able to listen to his\her favorite PDF and can do their daily routine. The application can be used to read any PDF which has[1].

The following application can be used to convert text from PDF to audio using PYSimpleGUI and python files, functions, and definitions. The main packages used in this audiobook converter are pytesseract and PyGame libraries. gTTs is a python library used for text to speech conversions. This is the reason which helps the machine to speak to us. Pytesseract extracts the text from the PDF. This is a python library built as a PDF toolkit. To overcome the(issues) the project PDF to Audio Converter has been developed to extract data from the pdf selected by the user, and to extract the data from the pdf, convert it to audio format to read out loud.

PDF to Audio Converter is a GUI application containing buttons for voice conversions, Destination Language and also has a feature to open a pdf, it will display the pdf while reading.

***Index Terms* - Python, pytesseract and PyMuPDF, Text to Speech, Converter, PYSimpleGUI .**

Text-to-speech and related read audio tools are being widely implemented in an attempt to assist students' reading comprehension skills.PDF to the audio system is a screen reader application designed and constructed for an effective audio communication system. PDFs were designed to present and exchange documents reliably, PDFs are an open standard document format used globally, maintained by the International Organization for Standardization (ISO). document format is one of the most convenient methods for electronic communication, and also for the exchange of information. Hence, there is a need to make it more accessible to readers on-screen through audio. PDF documents are designed and structured to contain links and buttons, form fields, audio or sounds, video, and business logic. The PDF to the audio system will power text on screens to read aloud (speak) with support for many languages . The PDF to Audio Converter project provides an alternative to access the PDF books for the blind, lazy, readers, and others. Using this PDF to Audio Converter the user will be able to listen to his\her favorite PDF and can do their daily routine. The following application can be used to convert text from PDF to audio using Python predefined libraries like pytesseract and PyMuPDF and also using some pygame files.

In the current busy routine people do not have time to take a book and spend time reading it, instead, everyone needs alternative access to read the content. If a person is traveling, he\she cannot read a book, instead of reading, they can listen to it. Reading stories or essays or any text can be arduous however an audiobook would make the task easy, by reading the text. However, an audio reading of the text is convenient and does not require much concentration as reading requires. When a person tends to read a book, it requires to invest his/her time in reading. Whereas the audiobook makes the task easy, and the user can perform their own task as well as listening to the audio. In this project, we have implemented a simple PDF to audio converter using python.

When we compare with the current features present in a normal audiobook converter, they convert PDF texts (or images) into speech, and they have volume controls with single voice conversion (either male or female). Only a single choice is given to the user in case of voice modification. They provide the play and pause options. The speed of voice is always fixed.

While seeing the Audiobook converter which we have built, it converts the PDF text (or images) into speech with differences in the same PDF document. Voice conversions are possible. The rate of speed of voices can also be changed (fast, normal, and slow) for better clarifications.

The user can also save, and modify the PDF document, while the audio of the converted PDF is running in the background. This exclusive feature reduces the time consumption of writing notes and reduces paper consumption.

We can also open the required PDF or any other PDF, while the audio is playing in the background.

The objective of our project is:
In the current generation students, researchers, authors don't find time to read a book on an electronic device as that might strain their eyes and might face other issues (headache, itchiness in the eye). So, to overcome those problems we have designed an application that extracts the text from the selected PDF and reads it out to the user.
Background Study of the project is:
Adobe Systems made the PDF specification available free of charge in 1993. In the early years PDF was popular mainly in desktop publishing workflows and competed with a variety of formats such as Common Ground Digital Paper, Farallon Replica, and even Adobe's own PostScript format.

PDF was a proprietary format controlled by Adobe until it was released as an open standard on July 1, 2008 and published by the International Organization for Standardization as ISO 32000-1:2008, at which time control of the specification passed to an ISO Committee of volunteer industry experts. In 2008, Adobe published a Public Patent License to ISO 32000-1 granting royalty-free rights for all patents owned by Adobe that are necessary to make, use, sell, and distribute PDF-compliant implementations [3].

Audiobooks allow students to hear explicit sounds of letters and letter patterns that form words. Audiobooks also help students engage in text and gain exposure to more words, ultimately improving vocabulary, comprehension, and critical thinking skills [4].

In this project, we present a simple way to combine different Python libraries for creating an Audiobook that takes a PDF file path as input and reads the text in the PDF file to the user via audio. Python programming language was used to create this project.

This code will ask the user to select any pdf file and will count the pages in the pdf and extract data from the pdf page by page and read the extracted data to the user.

1.1 Evolution

- **The creation of the audiobook medium has dramatically shaped the way that the world can access valuable information.**
- **For instance, the invention of audiobooks has opened the door for more and more people to benefit from the written word – including those with visual impairment, and those who would otherwise not have the time or ability to read.**
- **Whether the content is entertaining or educational, or anything in between, it's not uncommon for many of us to capitalize on 'downtime,' such as a lengthy commute, or a few hours spent doing household chores, by putting on an audiobook.**

1. The Precursor to Audiobooks, An Aluminum Disc and Battery-Powered Turntable

- **Audiobooks have a unique and fascinating history. In 1933, anthropologist J.P. Harrington, drove the length of North America to record oral histories of Native American tribes on aluminum discs using a car battery-powered turntable. Now, in the 21st Century, the definition of books and publishing is evolving as technology advances and the consumer demands change. Audiobooks allow avid readers to multi-task in today's hectic world. Consumers can listen to an audiobook as they**

commute, exercise, or cook. At the same time, audiobooks preserve the oral tradition of storytelling that J.P. Harrington pursued many years ago. Narration, sound effects, and music can complement the reading experience.

- **The Very First Audiobook Was Recorded in the 1950s**
- **In January 1952, Barbara Cohen and Marianne Roney, sat down with Dylan Thomas in the bar of the Chelsea Hotel and persuaded him to record some of his poetry. Spoken word records were almost unheard of at the time.**
- **Cohen and Roney knew that Thomas's poetry was shocking, moving and important, and that they wanted to record it to preserve the sounds. With the promise of $500, and much coaxing and cajoling, a recording session was arranged. Thomas selected the poems, writing the list in his tiny round letters in Miss Roney's appointment book.**
- **Caedmon Records was born the next week, named, appropriately enough, for the first poet to write in the native language of Old England. Then, on February 22, Peter Bartok, son of the composer Bela Bartok, had set up his equipment in Steinway Hall to do the recording.**
- **Thomas began the session with *Do Not Go Gentle Into That Good Night.***
- **Bartok had perhaps expected a quavery poet's voice, but instead he had to adjust the microphone for a symphonic recording to accommodate Thomas's sonorous voice.**
- **To fill the other side of the record, Thomas recorded a story he sold to Harper's Bazaar, *A Child's Christmas in Wales*. This recording established the story as a classic. Today, it is Dylan Thomas's most widely known work and, as a model of translucent prose, stands as an everlasting testament to his greatness as poet and bard.**

2. **Audio Format Reaches a Turning Point in the 1980s**
    - **Just like the history of music, the history of audiobooks closely follows the twists and turns of the recording industry.**
    - **One of the more remarkable turning points was the 1980s, with the longer audio format of cassette tapes and the inexpensive, portable Sony Walkman becoming prolific at the time.**
    - **With the accessibility and format of recordings now able to more adeptly serve the audiobook industry, by 1984, there were 11 audiobook publishing companies.**

- **However, it was Brilliance Audio who created waves in the industry after inventing a way to record twice as much on cassettes. This meant audiobook publishers could now produce affordable unabridged editions of their most popular books.**
- **Just one year later, there were twice as many audiobook publishers in the market. New major book publishers, such as Harper and Row, Random House, and Warner Communications joined in the distribution of audiobooks.**

3. **Audiobook Market Becomes Profitable in the Late '80s**
    - **Also in the '80s, a number of events happened that cemented audiobooks as a profitable market.**
    - **Leading the scene was the development of the Audio Publishers Association (APA), a professional non-profit trade association established by a group of publishers to promote awareness of spoken word audio and provide industry statistics to the public and its members.**
    - **Around the same time, the Book-of-the-Month, Time-Life, and the Literary Guild began offering audiobooks to their subscribers and other book clubs formed such as the History Book Club, Get Rich Club, Nostalgia Book Club, and Scholastic all offering audiobooks.**
    - **In 1987, Publishers Weekly began running a regular column to cover the audiobook industry. At the time, audiobooks were being sold in 75 percent of regional and independent bookstores and, as 1987 came to a close, the audiobook industry was estimated to be a worth $200 million.**
4. **The late 90s and Early 2000s Usher in the Age of Digital Audiobooks and Podcasts**
    - **In the late '90s to early 2000s, new compressed audio formats and portable media players furthered the popularity of audiobooks with consumers.**
    - **Perhaps unsurprisingly, by around the mid-'90s, the audiobook industry grew to a whooping 1.5 billion dollars per year in retail value. Not bad for something that started out as experimental curiosity. The audiobook industry was lifted to new heights when the APA introduced the Audie Awards, which would become known as the 'Oscars of spoken word entertainment.'**
    - **In 1997, Audible.com pioneered the world's first mass-market digital media player, named 'The Audible Player,' which sold for $200 and was advertised as being 'smaller and lighter than a Walkman.'**

- **As digital technology grew there was a movement in digital audiobooks which allowed consumers to access their audiobooks instantly from a growing number of online libraries. Audible.com was the first to establish a website (1998) where digital audiobooks could be purchased and downloaded.**

5. **The 2000s Leverage Podcasts to Bring Audiobooks to the Masses**
    - **In 2005, Montreal-based writer Hugh McGuire posed a question on his blog: "Can the net harness a bunch of volunteers to help bring books in the public domain to life through podcasting?"**
    - **From that notion he created of LibriVox, a website where volunteers select books in the public domain (royalty-free books) to narrate by themselves or as a group of narrators. By the time 2012 rolled around, LibriVox carried a catalogue of over 6,244 unabridged books and continues to produce dozens of titles per month.**
    - **You can still become a volunteer narrator at LibriVox. This platform provides a good way for up-and-comers to practice the art and skill of audiobook narration, as well as test drive the niche to see if it's right for your voice acting career.**
    - **In 2003 and 2004 cassettes were phased out and replaced by CDs as the dominant format for audiobooks, but the CD steadily declined as digital technology became more accessible and the popularity of audiobooks continued to grow.**
6. **Audiobooks in the 2010's**
    - **Did you know that one of the biggest, most diverse, and lucrative industries today is the consumption of audio books?**
    - **Audible, a company that sells audio books, has been profitable in this industry since 2003 and more than doubled its number of subscriptions (to 237,000) and its annual revenues (to an estimated $63 million) last year (source, Business 2.0 magazine).**
    - **Once Steve Jobs of Apple Computers downloaded "Lincoln at Gettysburg", he was hooked on Audible, introduced the service at a Macworld conference (2003), and played a recording of E.B. White narrating his beloved novel, "Charlotte's Web". Following Apple's lead, many Mac fans quickly adopted the new trend, downloading audio books to their iTunes, and most recently, to their iPods.**
    - **At the end of 2013, the Audio Publishers Association released a report indicating that CD revenue was down 7 percent, but still represented approximately 53 percent of the market, while download units were up 29 percent and represented approximately 61 percent of the market. Download revenue went up 24 percent,**

representing approximately 41 percent of the market. In the overview the report cited, "The greatest potential for growth exists in digital formats."

- **Michele Cobb, President of the APA, said that "The future of audiobooks is our industry continuing to do more of what we do best – make amazing performances from excellent content. We've been extremely good at evolving, embracing new formats and growing. I have no doubt that we will continue this and when the next big thing in digital makes its way to the surface will be riding the wave at the front."**

7. **Audiobook Publishing Industry in 2020**
    - **You may have noticed that audio books are entering into the mainstream regarding literary consumption. Books on tape, spoken word recordings, storytelling based upon oral tradition.**
    - **While it's lovely and whimsical to think about the endless options to satisfy us as consumers, the marriage of the spoken word and digital recording for audiobooks is currently a $900 million dollar industry. In 2019, Forbes reported that U.S. audiobook sales neared $1 billion, growing 25% year-over-year. The audiobook publishing industry has likely passed the $1 billion dollar mark in 2020.**
    - **The founder of Audible, Donald Katz, in an interview with Business 2.0, shared his insight on where this market is headed. The demand for audio books has ballooned, much aided by the presence of the iPod. People are taking their music with them wherever they go, are listening to podcasts, so the ability to take your favorite book with you in digital form via an MP3 player isn't a big stretch of the imagination.**
    - **Think of it as theatre of the mind. It's like floating off into an oasis of well crafted words flowing through the mouth of an excellent orator – it takes limited effort to listen, enjoy, and digest the plot, objective, or thematic material of an audio book.**
    - **As voice actors, narrators and storytellers, you too can have a piece of this pie. After all, the entire concept of this product is built upon using the human voice to breathe life into the written word, and now, this is an on-demand service.**
    - **Books, magazines, newspapers, speeches, interviews, how-to manuals, and also the localization of these is now the fastest growing market in the world regarding the use of voice-overs.**
    - **In 2020, you can take your favorite books anywhere now on an iPhone, Android or your preferred brand of smartphone. You'll also be able to take your pick between reading the printed word and listening to the spoken word in many cases. With**

Audible at the wheel and competitors creeping up quickly, there may be more audiobook titles available online than printed books stocked in libraries!

1.2 BENEFITS OF AUDIOBOOK OVER PDF:

Audiobooks have traditionally been used in schools by teachers of second-language learners, learning-disabled students, and struggling readers or nonreaders. In many cases, audiobooks have proven successful in providing a way for these students to access literature and enjoy books. But they have not been widely used with average, avid, or gifted readers. Varley (2002) writes, "Uncertain whether audiobooks belong to the respectable world of books or the more dubious world of entertainment, elementary and high-school teachers have often cast a fishy eye at them, and many have opted for the safe course of avoidance."

It might be appropriate, then, to list the benefits of audiobooks for *all* students. Audiobooks can be used to:

- **Introduce students to books above their reading level**
- **Model good interpretive reading**
- **Teach critical listening**
- **Highlight the humor in books**
- **Introduce new genres that students might not otherwise consider**
- **Introduce new vocabulary or difficult proper names or locales**
- **Sidestep unfamiliar dialects or accents, Old English, and old-fashioned literary styles**
- **Provide a read-aloud model**
- **Provide a bridge to important topics of discussion for parents and children who can listen together while commuting to sporting events, music lessons, or on vacations**
- **Recapture "the essence and the delights of hearing stories beautifully told by extraordinarily talented storytellers" (Baskin & Harris, 1995, p. 376)**

Additionally, many audiobooks are read by the author or include commentary by the author. A recording of *The Fighing Ground* by Avi, for example, includes an author interview in which he explains how he came up with the idea for the book. *Joey Pigza Swallowed the Key* is read by

author Jack Gantos and also includes commentary about why he wrote the book. This information can provide students with a connection to the author as well as insight into the author's thoughts and the writing process.

Even with all the benefits of audiobooks, however, they are not for all students. For some, the pace may be too fast or too slow. For others, the narrator's voice can be irritating or the use of cassette or CD players can be cumbersome when compared to the flexibility of the book. But the majority of students will find listening to well-narrated, quality literature to be a transformative experience. Varley (2002) states, "If one thing has struck me about the way people describe listening to audiobooks, it is the reported intensity of their absorption and the emotional grip of the experience. 'They go right to your soul,' says one listener."

One reason more audiobooks are not finding their way into classrooms is availability. Public libraries usually have a good quantity of audiobooks, but most school libraries have a limited number – audiobooks are expensive. The cost of cassette or CD players and headphones must also be taken into consideration, and though these costs have come down considerably in the last few years, schools typically do not budget funds for such purchases.

If money is available for purchasing audiobooks, it is important for librarians and teachers to do their homework before buying. Single-author unabridged audiobooks tend to be the best, though some dramatizations (such as Philip Pullman's *His Dark Materials* trilogy, read by the author with a cast of more than 40 British actors) can be excellent. There are many sources of audiobook reviews readily available online, including School Library Journal. [Note: The Association for Library Service for Children also publishes an annual Notable Children's Recordings list.]

Audiobooks can be a welcome addition to every classroom. Many students are avid readers while others are struggling to become readers and still others have given up hope. Audiobooks have something to offer all of them.

1.1.1 Image Processing

2.LITERATURE SURVEY

Sr. No.	Name	Method	Observations
1.	Dorian Miller., 2003	Microsoft's TTS engine. Microsoft Speech Application Programming Interface (SAPI), LAME software is used	The software is developed by python programming language and third-party components .
2.	Mentor Hamiti, Agni Dika., 2010	They have used Acoustic files of special letters and basic units albanian language [6].	In this paper, they presented that generating speech from written text in Albanian language, which helps the user those who have difficulties with their eye vision
3.	Itunuoluwa Isewon, AJelili Oyelade, Olufunke Oladipupo., 2014	Natural Language Processing (NLP) and Digital Signal Processing (DSP) technology are used	The software is called TextToSpeech Robot, it is a simple application that has text to speech functionality. The system was developed using Java programming language. Java programming language is used because it is an independent platform
4.	S. Venkateswarlu1 et. al, 2016	Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry pi are used.	The present paper has introduced a real time cost beneficial technique. It has the concept of optical character recognition and text to speech synthesizer. It basically has two modules, image processing module and voice processing .
5.	Mr. Manohar M.,2020	Python predefined libraries are used .	This application can read the text from a PDF that the user has selected which has page numbers .

This section is an overview of different computer solutions to provide audiotexts. Each solution has its unique navigation techniques between pages, sections, paragraphs etc. and its special features.

Books on tape is the oldest form of audio books started in 1970's and continued to be provided by the Recordings for Blind and Dyslexic (RFBD) [1]. Readings of a book are stored on tapes. A special tape recorder is used to listen to the four sides of the tape; besides the two sides of a normal tape, two more sides are available by reversing the play direction. The tape recorder also provides a dial to vary the playback speed; listeners familiar with the text or just skimming it can increase the tape speed. The inconvenience of using books on tape is that one book requires a large volume of tapes and it is tedious to flip and rewind/fast-forward the tape to find the desired section. However, to assist with the navigation, beeps are used while fast-forwarding or rewinding to indicate page transitions. The listener finds relevant information with an index card accompanying the tapes, which indicates which tape, side and direction pages are on. Digital talking book (DTB) is a standard for audio books developed by the Digital Accessible Information System (DASIY) consortium and standardized in 2002 by the National Information Standards Organization (NISO). The DTB standard describes how multimedia information, such as audio files, text files and images, are composed to create an audio book. The standard is flexible and combines variable amounts of audio and corresponding text, which enables text searches. Having complete audio and text is not needed, for example, in a dictionary, which might have the complete text and audio only for pronunciations. A DTB viewer program can use the text to display word definitions to a Braille display. Also with a DTB viewer the reader can efficiently navigate between or within sections, because of the hierarchical document structure defined by the DTB standard. Blind people rely on screen reader software, such as JAWS, to use a PC and read electronic documents. The screen reader reads all text that appears on a screen. Navigating the PC desktop and applications is possible with a series of keystrokes. Electronic documents, such as web pages, are also navigated with keystrokes, which enable moving between pages, lines, and words. The navigation, however, is limited Dorian Miller * Assistive technology* 4/26/2003 because there are no direct keystrokes to find the beginning of sections, paragraphs, or sentences. The original motivation for this project was to provide access to PDF files, which until recently were not accessible with screen readers. In the meanwhile, however, Adobe has

released Acrobat Reader 5.1, the standard PDF viewer, with screen reader accessibility. Regardless of Abode's recent development, the solution to listening to audiotexts on anMP3 player provides a unique and convenient access to text material.

3. EXISTING SYSTEM

The Blind and Dyslexic people would find it difficult to read, to support those people different computer solutions have provided many alternative ways to convert text to audio. . Readings of a book are stored on tapes. consortium and standardized in 2002 by the National Information Standards Organization (NISO). The DTB standard describes how multimedia information, such as audio files, text files and images, are composed to create an audio book. Many other solutions are to convert PDF files to MP3 players using third party applications or web applications.

ISSUES

In this section we are discussing the issues faced in the above section, they are .Readings of a book are stored on tapes, the inconvenience of using books on tape is that one book requires a large volume of tapes and it is tedious to flip and rewind/fast-forward the tape to find the desired section.

4. PROPOSED SYSTEM

In this current busy routine people do not find time to read a book, or to convert the PDF file into MP3 player using third party applications or web application. In this system I am developing an application using python to convert the PDF file into audio format and read out to the user. The application is more used friendly as it not requires any audio file or MP3 player. The user will have to select the PDF file which user wants to listen.

In this system, we are developing a GUI application using python to convert the PDF file into audio format and read it out to the user. The application is more user-friendly as it does not require any audio file or MP3 player. The user will have to select the PDF file which the user wants to listen to. The problem Statement of this project is: To create a PDF to Audiobook Converter using different Tkinter and python files, functions, and definitions. The main

packages used in this audiobook converter are PyMuPDf,Pytesseract, fitz and Pygame. PyMuPDF is a Python binding for MuPDF – a lightweight PDF, XPS, and E-book viewer, renderer, and toolkit, which is maintained and developed by Artifex Software, Inc MuPDF can access files in PDF, XPS, OpenXPS, CBZ, EPUB and FB2 (e-books) formats, and it is known for its top performance and high rendering quality. MuPDF stands out among all similar products for its top rendering capability and unsurpassed processing speed. At the same time, its "light weight" makes it an excellent choice for platforms where resources are typically limited, like smartphones.

Python-tesseract is a wrapper for Google's Tesseract-OCR Engine. It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if used as a script, Python-tesseract will print the recognized text instead of writing it to a file.

The pygame library is an open-source module for the Python programming language specifically intended to help you make games and other multimedia applications. Built on top of the highly portable SDL (Simple DirectMedia Layer) development library, pygame can run across many platforms and operating systems. By using the pygame module, you can control the logic and graphics of your games without worrying about the backend complexities required for working with video and audio.

Googletrans is a free and unlimited python library that implemented Google Translate API. This uses the Google Translate Ajax API to make calls to such methods as detect and translate. PDF to Audio Converter is a GUI application containing play, pause(buttons), and label to display text which allows the user to select the PDF files and the user has to click on the play button in order to extract text from the PDF file and will read the text. The application has been developed in a way that until the speaker reads out the extracted text the player cannot be paused. The GUI is also provided with a label to display the text, the text will only be displayed only after reading the extracted text. With PDF being the most used document format globally, there is a need to convert the text in PDF formats into Audio signals. These can be utilized for various purposes, e.g., in the educational system, car navigation, announcements in railway stations, response services in telecommunications, and email reading. Furthermore, people with vision disabilities cannot view or read PDF files and this is a major setback. This research addresses the problems in converting PDF text into speech. One is how to improve the naturalness of synthetic speech in PDF-based text into an Audio system.

The Workflow of the project is:

• In this PDF to Audio Converter the user needs to select any PDF file from the desired location by pressing the open pdf.

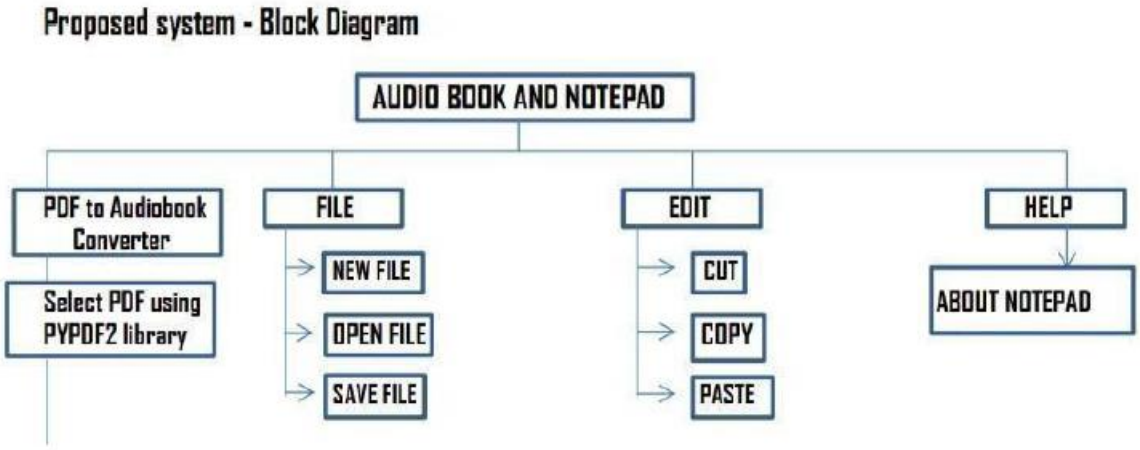
• After selecting the PDF file, we have to select the Range of pages we want to convert in the audio.

• After selecting the Range, we have to select in which language we want the pdf to be converted .

• After pressing enter key ,Program will display the desired output.

• To exit the program, we press the cancel button. The objective of our project is: In the current generation students, researchers, authors don't find time to read a book on an electronic device as that might strain their eyes and might face other issues (headache, itchiness in the eye).

So, to overcome those problems we have designed an application that extracts the text from the selected PDF and reads it out to the user. Audiobooks allow students to hear explicit sounds of letters and letter patterns that form words. Audiobooks also help students engage in text and gain exposure to more words, ultimately improving vocabulary, comprehension, and critical thinking skills. In this project, we present a simple way to combine different Python libraries for creating an Audiobook that takes a PDF file path as input and reads the text in the PDF file to the user via audio. Python programming language was used to create this project



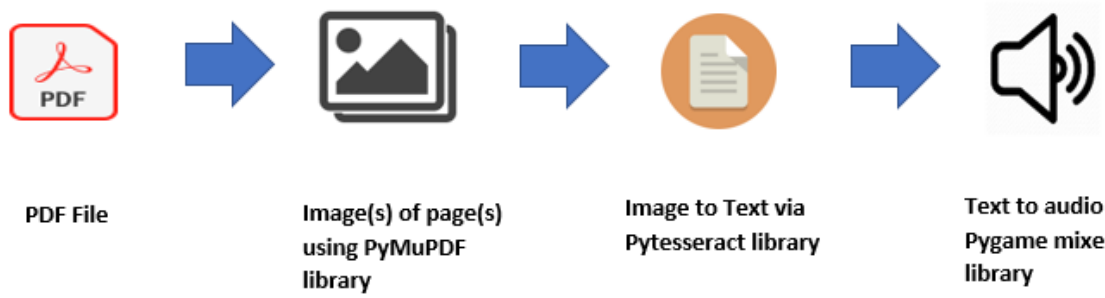
5. PROBLEM STATEMENT & ARCHITECTURE

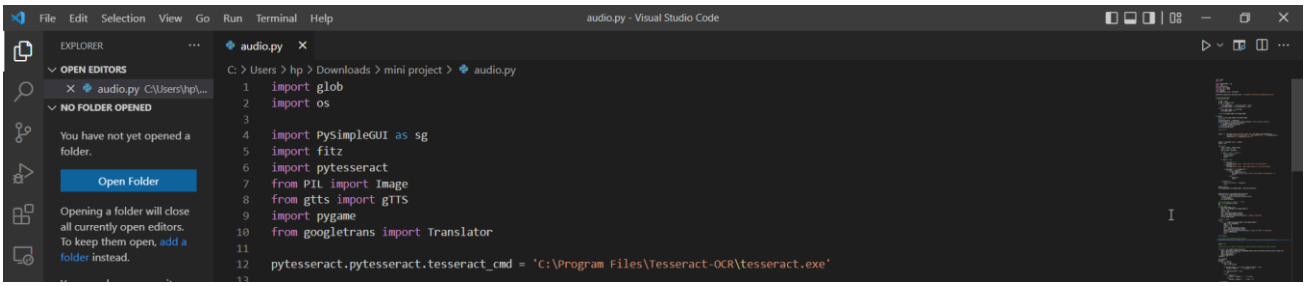
The above section (related work) we have discussed the way audio book was used. This section(issues) describes the drawback of the older versions of audio book. To overcome the(issues) the project PDF to Audio Converter has been developed to extract data from the pdf selected by the user, and to extract the data from the pdf, convert it to audio format to read out loud with language Translation.

ARCHITECTURE

PDF to Audio Converter is a GUI application contains play, pause(buttons) and label to display text which allows the user to select the PDF files and the user has to click on the play button in order to extract text from PDF file and will read the text. The application has been developed in a way that the until the speaker read's out the extracted text the player cannot be paused. The GUI is also, provided with a label to display the text, the text will only be displayed only after reading the extracted text.

6. WORK FLOW





I. **The glob module finds all the pathnames matching a specified pattern according to the rules used by the Unix shell, although results are returned in arbitrary order. No tilde expansion is done, but *, ?, and character ranges expressed with [] will be correctly**

matched. This is done by using the os.scandir() and fnmatch.fnmatch() functions in concert, and not by actually invoking a subshell.

II. **This module provides a portable way of using operating system dependent functionality. If you just want to read or write a file see open(), if you want to manipulate paths, see the os.path module, and if you want to read all the lines in all the files on the command line see the fileinput module. For creating temporary files and directories see the tempfile module, and for high-level file and directory handling see the shutil module.**

III. **PySimpleGUI is a Python package that enables Python programmers of all levels to create GUIs. You specify your GUI window using a "layout" which contains widgets (they're called "Elements" in PySimpleGUI). Your layout is used to create a window using one of the 4 supported frameworks to display and interact with your window. Supported frameworks include tkinter, Qt, WxPython, or Remi. The term "wrapper" is sometimes used for these kinds of packages.**

IV. PyMuPDF **is a Python binding for MuPDF – a lightweight PDF, XPS, and E-book viewer, renderer, and toolkit, which is maintained and developed by Artifex Software, Inc MuPDF can access files in PDF, XPS, OpenXPS, CBZ, EPUB and FB2 (e-books) formats, and it is known for its top performance and high rendering quality.**

V. **The top level Python import name for this library is** "fitz"**. This has historical reasons: The original rendering library for MuPDF was called *Libart.*"*After Artifex Software acquired the MuPDF project, the development focus shifted on writing a new modern graphics library called "Fitz". Fitz was originally intended as an R&D project to replace the aging Ghostscript graphics library, but has instead become the rendering engine powering MuPDF."* (Quoted from Wikipedia).So PyMuPDF** cannot coexist **with packages named "fitz" in the same Python environment.**

VI. **Python-tesseract is an optical character recognition (OCR) tool for python. That is, it will recognize and "read" the text embedded in images.Python-tesseract is a wrapper for Google's Tesseract-OCR Engine. It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if used as a script, Python-tesseract will print the recognized text instead of writing it to a file.**

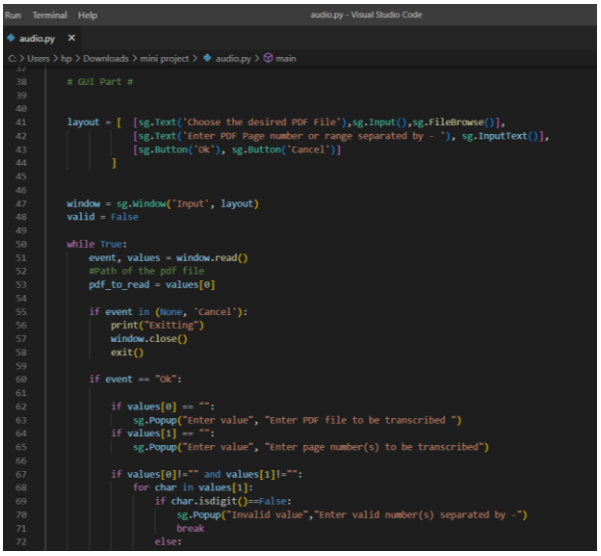
VII. **Pillow is the friendly PIL fork by Alex Clark and Contributors. PIL is the Python Imaging Library by Fredrik Lundh and Contributors.The Python Imaging Library**

adds image processing capabilities to your Python interpreter.This library provides extensive file format support, an efficient internal representation, and fairly powerful image processing capabilities.The core image library is designed for fast access to data stored in a few basic pixel formats. It should provide a solid foundation for a general image processing tool.

VIII. gTTS (*Google Text-to-Speech*), a Python library and CLI tool to interface with Google Translate's text-to-speech API. Write spoken `mp3` data to a file, a file-like object (bytestring) for further audio manipulation, or `stdout`. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

IX. The pygame library is an open-source module for the Python programming language specifically intended to help you make games and other multimedia applications. Built on top of the highly portable SDL (Simple DirectMedia Layer) development library, pygame can run across many platforms and operating systems.By using the pygame module, you can control the logic and graphics of your games without worrying about the backend complexities required for working with video and audio. This module contains classes for loading Sound objects and controlling playback. The mixer module is optional and depends on SDL_mixer. Your program should test that pygame.mixerpygame module for loading and playing sounds is available and initialized before using it.

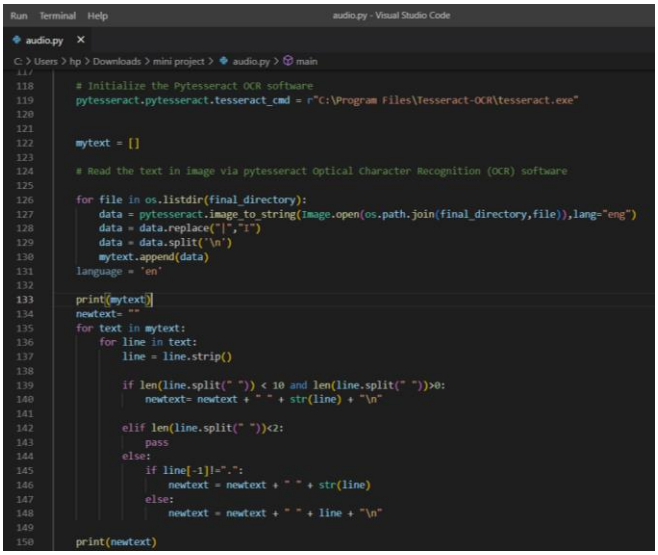
X. Googletrans is a free and unlimited python library that implemented Google Translate API. This uses the Google Translate Ajax API to make calls to such methods as detect and translate.



SimpleGUI Layout for input

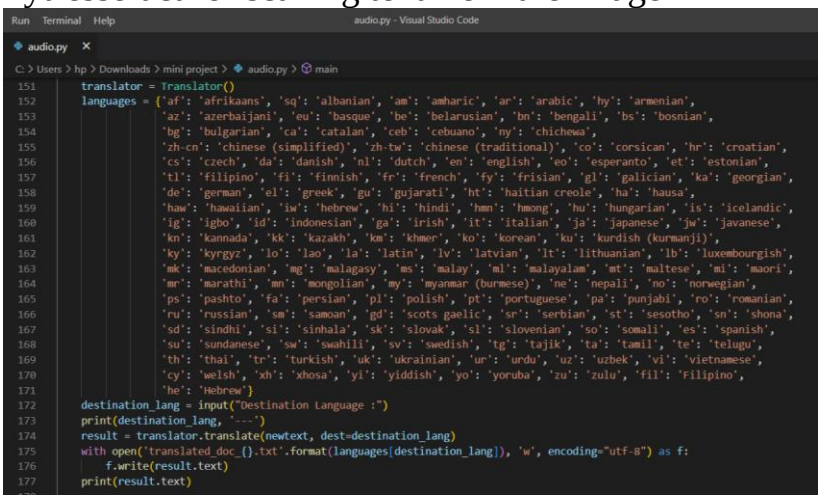
```
                        valid=True
                        break

            if valid==True:
                print('You entered ', values[1])
                break

    window.close()
    first_page_number,last_page_number = get_text(values[1])



    image_directory = glob.glob(final_directory)
    for file in os.listdir(final_directory):
        filepath = os.path.join(final_directory,file)
        print(filepath)
        os.remove(filepath)

    # Store PDF pages as images in a folder
    doc = fitz.open(pdf_to_read)
    k=1
    # Single page
    if last_page_number == 0:
        page = doc.loadPage(first_page_number-1)
        zoom_x = 2.0
        zoom_y = 2.0
        mat = fitz.Matrix(zoom_x,zoom_y)
        pix = page.getPixmap(matrix=mat)
        output = os.path.join(final_directory, r"image_to_read.png")
        pix.writePNG(output)

    # Range of pages
    else:
        for i in range(first_page_number-1,last_page_number):
            page = doc.loadPage(i)
            zoom_x = 2.0
            zoom_y = 2.0
```

Use of Fitz for opening image and extracting text



```
    # Initialize the Pytesseract OCR software
    pytesseract.pytesseract.tesseract_cmd = r"C:\Program Files\Tesseract-OCR\tesseract.exe"


    mytext = []

    # Read the text in image via pytesseract Optical Character Recognition (OCR) software

    for file in os.listdir(final_directory):
        data = pytesseract.image_to_string(Image.open(os.path.join(final_directory,file)),lang="eng")
        data = data.replace("|","I")
        data = data.split("\n")
        mytext.append(data)
    language = 'en'

    print(mytext)
    newtext= ""
    for text in mytext:
        for line in text:
            line = line.strip()

            if len(line.split(" ")) < 10 and len(line.split(" "))>0:
                newtext= newtext + " " + str(line) + "\n"

            elif len(line.split(" "))<2:
                pass
            else:
                if line[-1]!=".":
                    newtext = newtext + " " + str(line)
                else:
                    newtext = newtext + " " + line + "\n"

    print(newtext)
```

Pytresseract for scaning text from the image



```
    translator = Translator()
    languages = {'af': 'afrikaans', 'sq': 'albanian', 'am': 'amharic', 'ar': 'arabic', 'hy': 'armenian',
                 'az': 'azerbaijani', 'eu': 'basque', 'be': 'belarusian', 'bn': 'bengali', 'bs': 'bosnian',
                 'bg': 'bulgarian', 'ca': 'catalan', 'ceb': 'cebuano', 'ny': 'chichewa',
                 'zh-cn': 'chinese (simplified)', 'zh-tw': 'chinese (traditional)', 'co': 'corsican', 'hr': 'croatian',
                 'cs': 'czech', 'da': 'danish', 'nl': 'dutch', 'en': 'english', 'eo': 'esperanto', 'et': 'estonian',
                 'tl': 'filipino', 'fi': 'finnish', 'fr': 'french', 'fy': 'frisian', 'gl': 'galician', 'ka': 'georgian',
                 'de': 'german', 'el': 'greek', 'gu': 'gujarati', 'ht': 'haitian creole', 'ha': 'hausa',
                 'haw': 'hawaiian', 'iw': 'hebrew', 'hi': 'hindi', 'hmn': 'hmong', 'hu': 'hungarian', 'is': 'icelandic',
                 'ig': 'igbo', 'id': 'indonesian', 'ga': 'irish', 'it': 'italian', 'ja': 'japanese', 'jw': 'javanese',
                 'kn': 'kannada', 'kk': 'kazakh', 'km': 'khmer', 'ko': 'korean', 'ku': 'kurdish (kurmanji)',
                 'ky': 'kyrgyz', 'lo': 'lao', 'la': 'latin', 'lv': 'latvian', 'lt': 'lithuanian', 'lb': 'luxembourgish',
                 'mk': 'macedonian', 'mg': 'malagasy', 'ms': 'malay', 'ml': 'malayalam', 'mt': 'maltese', 'mi': 'maori',
                 'mr': 'marathi', 'mn': 'mongolian', 'my': 'myanmar (burmese)', 'ne': 'nepali', 'no': 'norwegian',
                 'ps': 'pashto', 'fa': 'persian', 'pl': 'polish', 'pt': 'portuguese', 'pa': 'punjabi', 'ro': 'romanian',
                 'ru': 'russian', 'sm': 'samoan', 'gd': 'scots gaelic', 'sr': 'serbian', 'st': 'sesotho', 'sn': 'shona',
                 'sd': 'sindhi', 'si': 'sinhala', 'sk': 'slovak', 'sl': 'slovenian', 'so': 'somali', 'es': 'spanish',
                 'su': 'sundanese', 'sw': 'swahili', 'sv': 'swedish', 'tg': 'tajik', 'ta': 'tamil', 'te': 'telugu',
                 'th': 'thai', 'tr': 'turkish', 'uk': 'ukrainian', 'ur': 'urdu', 'uz': 'uzbek', 'vi': 'vietnamese',
                 'cy': 'welsh', 'xh': 'xhosa', 'yi': 'yiddish', 'yo': 'yoruba', 'zu': 'zulu', 'fil': 'Filipino',
                 'he': 'Hebrew'}
    destination_lang = input("Destination Language :")
    print(destination_lang, '---')
    result = translator.translate(newtext, dest=destination_lang)
    with open('translated_doc_{}.txt'.format(languages[destination_lang]), 'w', encoding="utf-8") as f:
        f.write(result.text)
    print(result.text)
```

Translation Part

```
    myobj = gTTS(text=result.text, lang=destination_lang, slow=False)

    # Audio in a mp3 file
    myobj.save(os.path.join(final_directory,"pdf_audio.mp3"))

    # Play the audio file
    pygame.init()
    pygame.display.set_mode((200,100))
    pygame.mixer.init()
    pygame.mixer.music.load(os.path.join(final_directory,"pdf_audio.mp3"))
    pygame.mixer.music.play()
    pygame.event.wait()

    clock = pygame.time.Clock()
    clock.tick(10)
    while pygame.mixer.music.get_busy():
        pygame.event.poll()
        clock.tick(10)


  # GUI END #

if __name__ == '__main__':
    main()
```

gTTs to convert google text to speech
pygame for playing audio.

7. ALGORITHM

1.Take the PDF file and convert each page into image using PyMuPDF library.

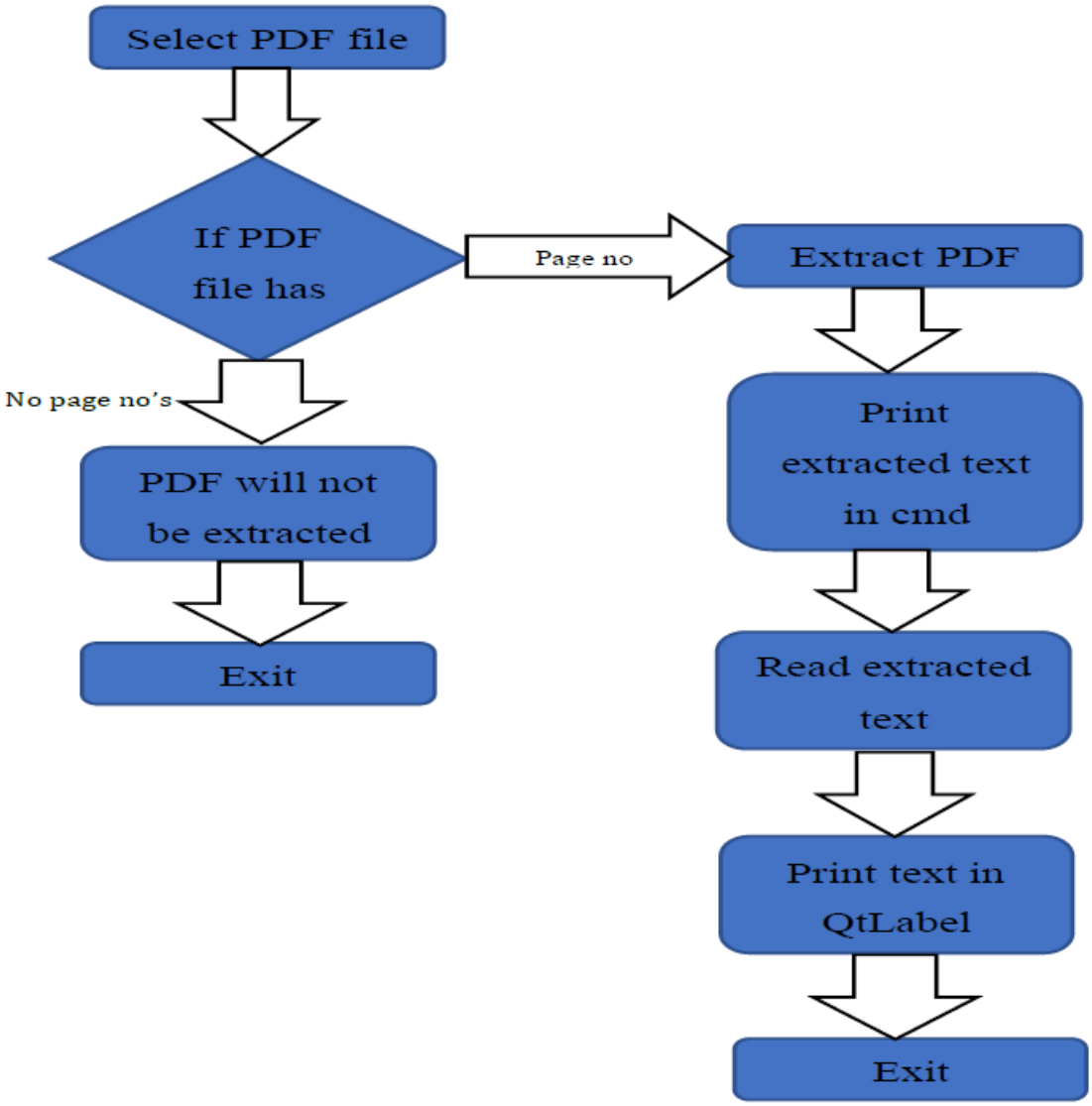
2.Take the image(s) and scan the text in the image using Pytesseract OCR library.

3.Use Google Text to Speech (gTTS) library to convert text to audio file.

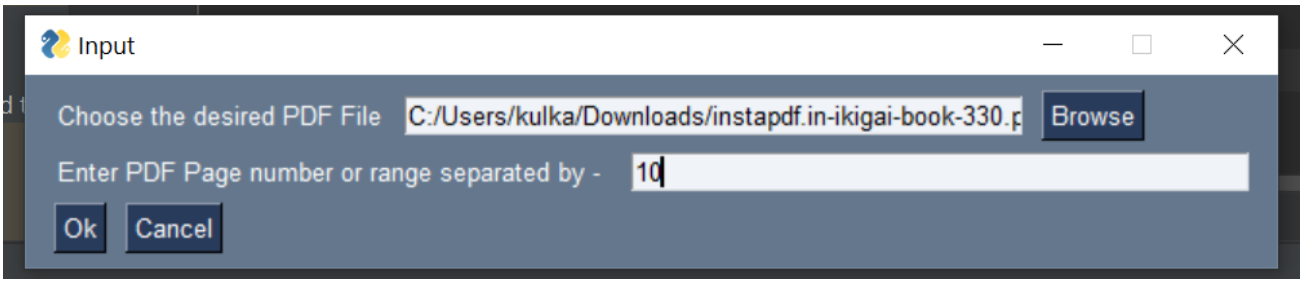
4.Get the Pygame mixer to play the audio file loud.

5.Convert text in selected language and create a .txt file for storing the converted text.

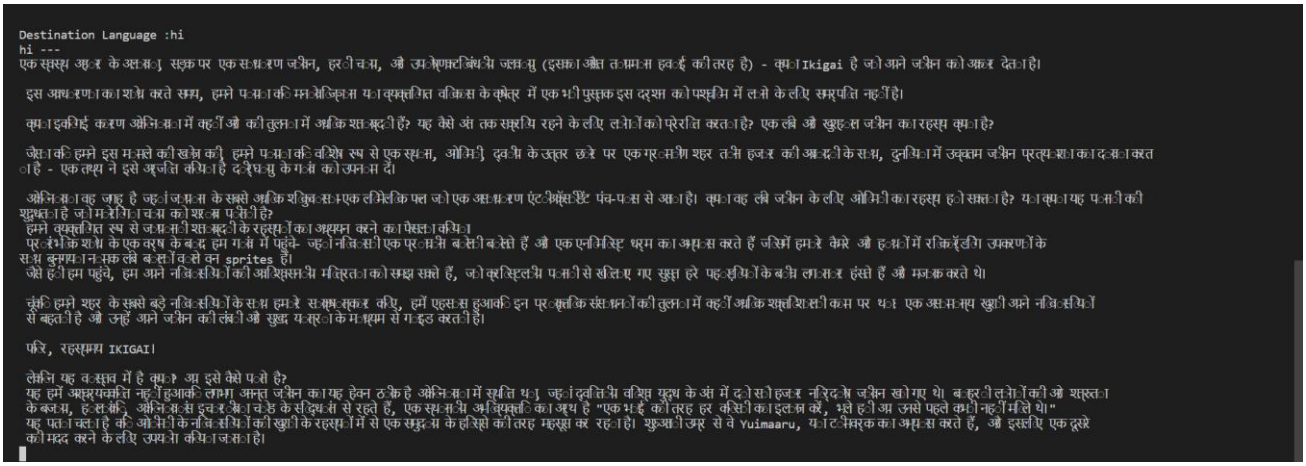
8. FLOWCHART



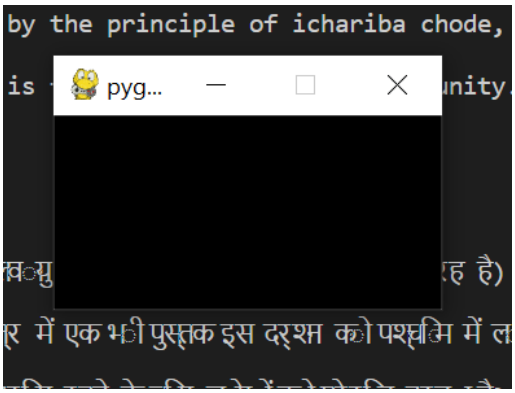
9. RESULTS



Screenshot Of PDF to audio converter interface



TRANSLATED TEXT



AUDIO FILE GENERATED

10.ADVANTAGES & DISADVANTAGES

- **It avoids eye-strain from too much reading.**
- **It helps seniors or those having vision problems.**
- **It can help in reading large paragraphs and offers range of different languages.**
- **Pronunciation analysis from written text is a major concern.**
- **It is difficult to build a perfect system.**
- **Filtering background noise is a task which can even be difficult for humans to accomplish**

11.CONCLUSION AND FUTURE WORK

- **It was seen that this code performs really well in reading straightforward PDF text files.**
- **Should enable users to select the desired PDF and convert it to audio and display text in, so the user can understand that particular text has been read.**
- **Should enable students with reading disabilities.**
- **The success of this research project is significant given the broad use of audiobooks in literacy and library programs across the United States.**
- **Teachers and school librarians may also use these findings as a rationale for adding audiobooks to the list of reading strategies used successfully with struggling readers.**
- **We are interested in future research on the use of audiobooks with struggling readers who are younger and older than those who participated in the study and on audiobook usage with English Language Learners [10].**
- **At this point, the code does not have a stop feature, I intend to add those and do more interesting things with the application of Machine Learning in the audiobook. With the help of machine learning, we can add the features that will recognize the voice of the user and implement the function as the user wants [10]. This feature will help mostly for the disabled users like the blind, handicap.**

12.REFERENCES

• **AJINKYA DOMALE; BHIMSEN PADALKAR; RAJ PAREKH; M.A. JOSHI PRINTED BOOK TO AUDIO BOOK CONVERTER FOR VISUALLY IMPAIRED https://ieeexplore.ieee.org/document/6757125**

• • **Bernard Conan Pobiak-ADJUSTABLE ACCESS ELECTRONIC BOOKS, The Reader Project of Washington, D.C., 1992 IEEE**

• • **R. M. K. Sinha, "A Syntactic pattern analysis system and its application to Devanagari script recognitio," Ph.D. Thesis, Dept. Elect. Eng., Indian Institute of Technology, Kanpur, India, 1973.**

• • **R. M. K. Sinha and H. Mahabala, "Machine recognition of Devanagari script" IEEE Trans. Syst. Man Cybern., vol. 9, no. 8, pp. 435-441, Aug.1979.**

• • **S. Palit and B. B. Chaudhuri, "A feature-based scheme for the machine recognition of printed Devanagari script" in Pattern Recognition, Image Processing and Computer Vision, P. P. Das and B. N. Chatterjee, Eds. New Delhi, India: Narosa Publishing House, 1995, pp. 163-168.**

• • **P. V. S. Rao and R. B. Thosar, "A programming system for studies in speech synthesis," IEEE Trans. Acoust., Speech and Signal Processing, vol. 22, no. 3, pp. 217-225, 1974.**

• • **B. Yegnanarayana, S. Rajendran, V. R. Ramachandran, and A. S. Madhukumar, "Significance of knowledge sources for a text-tospeech system for Indian languages," Sadhana, pp. 147-169,1994**

• • **S. S. Agrawal, "Synthesizing Hindi speech using Klsyn and Hlsyn for natural sounding," in Workshop on Spoken Language Processing, TIFR, Mumbai, 2003.**

• • **X. A. Furtado and A. Sen, "Synthesis of unlimited speech in Indian languages using formant-based rules," Sadhana, vol. 21, no. 3, pp. 345-362, 1996.**

• • **Kishore Prahallad, E.Naresh Kumar, Venkatesh Keri, S.Rajendran, Alan W Black," The IIIT-H Indic Speech Databases**

• • **Jan Mennens "Optical recognition of Braille writing", IEEE 1993. pp 428-431.**

• • **Rajamani M. Kulkarni; Preetam R. Kalburgimath An Intelligent Text Reader based on Python An Intelligent Text Reader based on Python | IEEE Conference Publication | IEEE Xplore**

• • **Design and Implementation of Text To Speech Conversion for Visually Impaired People Itunuoluwa Isewon*, Jelili Oyelade, Olufunke Oladipupo Design and Implementation of Text To Speech Conversion for Visually Impaired People (core.ac.uk)**

• **Chithra Selvaraj, Bhalaji Natarajan Enhanced portable text to speech converter for visually impaired. o (PDF) Enhanced portable text to speech converter for visually impaired (researchgate.net)**