

```

# len(files) > 0
# asking user if the current directory
# contains pdf file or not
# if file exist then
# print("File exist")
# else
# print("File don't exist")
# ask user which file he want to parse
# file = input("Type the id of file you want to parse: ")
# if len(file) == len(str(file)):
#     print("File " + file + " does not exist")
# else:
#     file = int(file)
#     line = input("From which line you want to parse the data? ")
#     line = int(line)
#     row_count = 0
#     row_max = 0
#     for i in range(0, len(files)):
#         if i == file:
#             row_count = 0
#             row_max = len(files[i])
#             break
#     if line > len(files[file]):
#         print("Line not exist")
#     else:
#         print("Starting Chrome...")
#         web = webdriver.Chrome()
#         full_path = path+r'\chromedriver.exe'
#         options = Options()
#         # options.add_argument('--log-level=2')
#         options.add_argument('--disable-gpu')
#         driver = webdriver.Chrome(executable_path=full_path, options=options)
#         book = openpyxl.load_workbook(files[int(file)])
#         sheet = book.active
#         row_count = sheet.max_row
#         row_max = sheet.max_col
#         for i in range(0, row_max):
#             print(files[file][line][i])
#             row_count += 1
#             if row_count == row_max:
#                 break
#             else:
#                 continue
#         driver.quit()
# 
```



# PDF READER PYTHON

<input checked="" type="checkbox"/> Class	INDEPENDENT PROJECT
<input checked="" type="checkbox"/> Materials	
<input checked="" type="checkbox"/> Property	PDF READER
<input checked="" type="checkbox"/> Reviewed	<input type="checkbox"/>
<input checked="" type="checkbox"/> Type	#by harshvardhan

## ABOUT IT

- It reads each and every pdf in minutes

## STEPS TO CREATE

- Install IDE and python in your pc .
- Here in this project we have used the following modules
  - pyttsx3 .
  - pyPDF2 .
  - speech\_recognition as sr .

- `datetime` .
- We have created an following functions .
  - `Engine` .
  - `speak`
  - `takecommand`
  - `wish`
  - `main program` .
- Now we will import the following modules by using the pip command .

```
import pytsxs3
import PyPDF2
import speech_recognition as sr
import datetime
```

- First we will create an engine with the code .

```
engine =pytsxs3.init('sapi5')
voices = engine.getProperty('voices')
print(voices[0].id)
engine.setProperty('voices',voices[0].id)
```

- Now we will use the engine function by creating speak function .

```
def speak(audio):
    engine.say(audio)
    print(audio)
    engine.runAndWait()
```

- Now we will create an take command function .

```
def takecommand():
    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        r.pause_threshold = 1
        audio = r.listen(source,timeout = 2,phrase_time_limit=5)

    try:
        print("Recognizing...")
        query = r.recognize_google(audio, language='en-in')
```

```

    print(f'user said: {query}')

except Exception as e:
    speak("Say that again please....")
    return "none"
return query

```

- Now we will create an wish function for greeting .

```

def wish():
    hour = int(datetime.datetime.now().hour)
    minute = datetime.datetime.now().strftime("%H:%M:%S")
    speak("Hello boss!")
    if hour>=0 and hour<=12:
        speak("Good morning Sir")
        speak(f'right now time is {minute}')
    elif hour>=12 and hour<=18:
        speak("Good afternoon")
        speak(f'right now time is {minute}')
    else:
        speak("Good evening")
        speak(f'right now time is {minute}')

```

- Now we will right the main program .

```

wish()
speak("Welcome to Pdf reader")
url = input("Please enter the Url of the pdf!\n")
speak("Which mode do you want to select")
quary = takecommand().lower()
while True:
    if "specific page" in quary:
        speak("please speak up the page which you want to read")
        quary2 = takecommand().lower()
        book = open('nameofthefile.pdf', 'rb')
        pdfReader = PyPDF2.PdfFileReader(book)
        page = pdfReader.getPage(int(quary2) - 1)
        text = page.extractText()
        speak(text)
    elif "book" in quary:
        book = open('nameofthefile.pdf', 'rb')
        pdfReader = PyPDF2.PdfFileReader(book)
        number = pdfReader.numPages
        for list in range(0,number):
            page = pdfReader.getPage(list)
            text = page.extractText()
            speak(text)

```

- After launching this code it will look like somewhat this .

```

import pyaudio
import pyPDF2
import speech_recognition as sr
import datetime

#engine setup
engine = pyaudio.PyAudio()
voices = engine.getProperty('voices')
print(voices[0].id)
engine.setProperty('voices', voices[0].id)

#creating speak function
def speak(audio):
    engine.say(audio)
    print(audio)
    engine.runAndWait()

#speaking code
def takecommand():
    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        r.pause_threshold = 1
        audio = r.listen(source, timeout = 2,phrase_time_limit=5)

    try:
        print("Recognizing...")
        query = r.recognize_google(audio, language='en-in')
        print("User said: " + query)
    except Exception as e:
        speak("Say that again please....")
        return "none"
    return query

#wish function
def wish():
    hour = int(datetime.datetime.now().hour)
    minute = datetime.datetime.now().strftime("%M:%S")
    speak("Hello boss!")
    if hour >= 6 and hour <= 12:
        speak("Good morning sir")
        speak("right now time is " + minute)
    elif hour >= 12 and hour <= 18:
        speak("Good afternoon")
        speak("right now time is " + minute)
    else:
        speak("Good evening")
        speak("right now time is " + minute)
    speak("I am legion , Online and ready to help")

#main program
if __name__ == "__main__":
    speak("Welcome to Pdf reader")
    print("Which mode do you want to select")
    query = takecommand().lower()
    while True:
        if "specific page" in query:
            speak("please speak up the page which you want to read")
            query2 = takecommand().lower()
            book = open('CHECKLIST.pdf', 'rb')
            pdfReader = PyPDF2.PdfFileReader(book)
            number = pdfReader.numPages
            for i in range(0, number):
                if query2 == str(i+1):
                    page = pdfReader.getPage(i)
                    text = page.extractText()
                    speak(text)
        elif "body" in query:
            book = open('CHECKLIST.pdf', 'rb')
            pdfReader = PyPDF2.PdfFileReader(book)
            number = pdfReader.numPages
            for i in range(0, number):
                page = pdfReader.getPage(i)
                text = page.extractText()
                speak(text)
        elif "language" in query:
            print("Positive Body Language")
            print("Smile")
            print("Eye Contact")
            print("Good Posture")

```

- This is the code output for this .

```

Hello boss!
Good afternoon
right now time is 18:54:59
I am legion , Online and ready to help
Welcome to Pdf reader
Which mode do you want to select
Listening...
Recognizing...
user said: book
CHECKLIST

BODY
-
LANGUAGE

Positive Body Language

Smile

Eye Contact

Good Posture

```