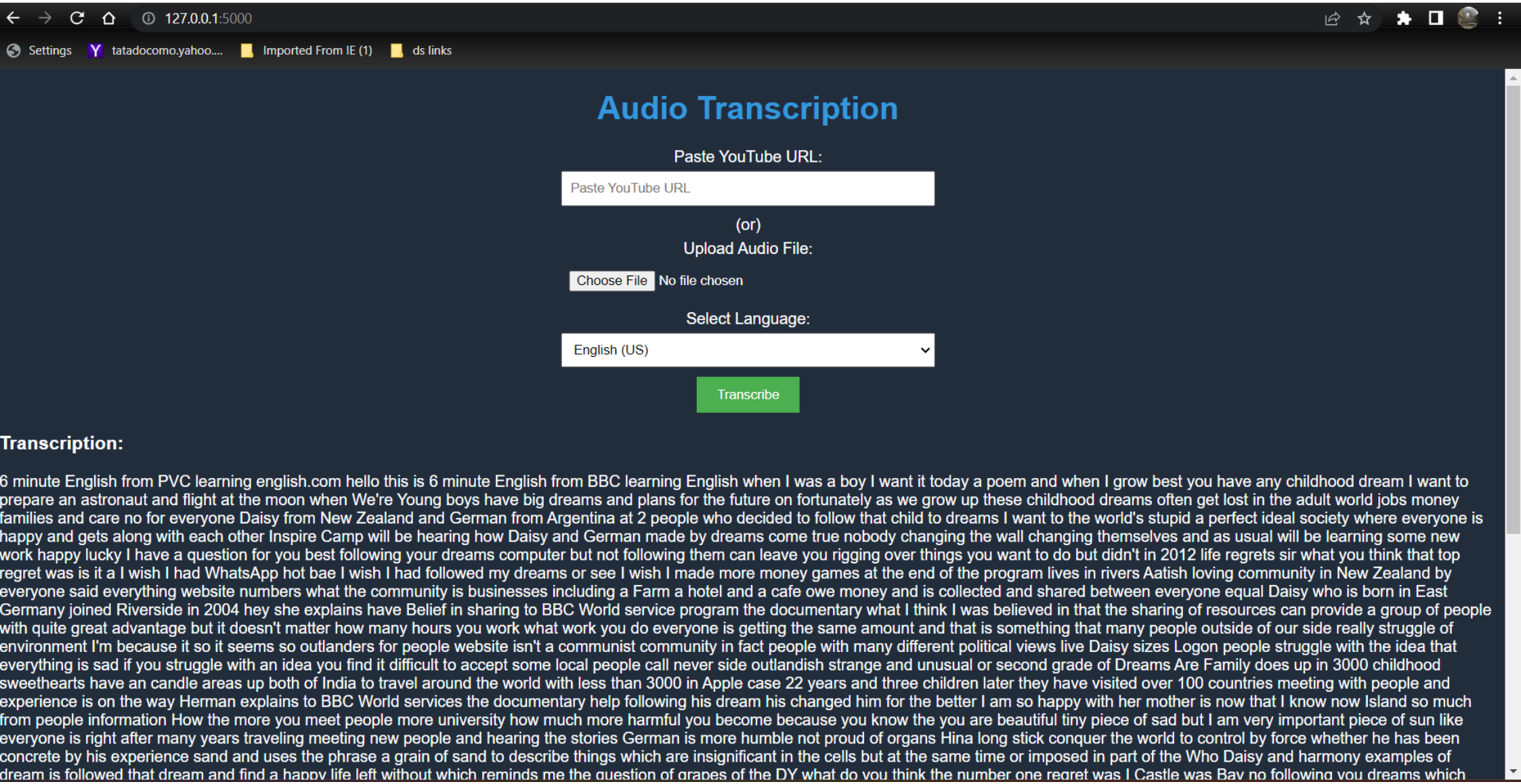

127.0.0.1:5000
Settings
tatadocomo.yahoo....
Imported From IE (1)
ds links
Audio Transcription
Paste YouTube URL:
Paste YouTube URL
(or)
Upload Audio File:
Choose File
No file chosen
Select Language:
English (US)
Transcribe
Transcription:
6 minute English from PVC learning english.com hello this is 6 minute English from BBC learning English when I was a boy I want it today a poem and when I grow best you have any childhood dream I want to prepare an astronaut and flight at the moon when We're Young boys have big dreams and plans for the future on fortunately as we grow up these childhood dreams often get lost in the adult world jobs money families and care no for everyone Daisy from New Zealand and German from Argentina at 2 people who decided to follow that child to dreams I want to the world's stupid a perfect ideal society where everyone is happy and gets along with each other Inspire Camp will be hearing how Daisy and German made by dreams come true nobody changing the wall changing themselves and as usual will be learning some new work happy lucky I have a question for you best following your dreams computer but not following them can leave you rigging over things you want to do but didn't in 2012 life regrets sir what you think that top regret was is it a I wish I had WhatsApp hot bae I wish I had followed my dreams or see I wish I made more money games at the end of the program lives in rivers Aatish loving community in New Zealand by everyone said everything website numbers what the community is businesses including a Farm a hotel and a cafe owe money and is collected and shared between everyone equal Daisy who is born in East Germany joined Riverside in 2004 hey she explains have Belief in sharing to BBC World service program the documentary what I think I was believed in that the sharing of resources can provide a group of people with quite great advantage but it doesn't matter how many hours you work what work you do everyone is getting the same amount and that is something that many people outside of our side really struggle of environment I'm because it so it seems so outlanders for people website isn't a communist community in fact people with many different political views live Daisy sizes Logon people struggle with the idea that everything is sad if you struggle with an idea you find it difficult to accept some local people call never side outlandish strange and unusual or second grade of Dreams Are Family does up in 3000 childhood sweethearts have an candle areas up both of India to travel around the world with less than 3000 in Apple case 22 years and three children later they have visited over 100 countries meeting with people and experience is on the way Herman explains to BBC World services the documentary help following his dream his changed him for the better I am so happy with her mother is now that I know now Island so much from people information How the more you meet people more university how much more harmful you become because you know the you are beautiful tiny piece of sad but I am very important piece of sun like everyone is right after many years traveling meeting new people and hearing the stories German is more humble not proud of organs Hina long stick conquer the world to control by force whether he has been concrete by his experience sand and uses the phrase a grain of sand to describe things which are insignificant in the cells but at the same time or imposed in part of the Who Daisy and harmony examples of dream is followed that dream and find a happy life left without which reminds me the question of grapes of the DY what do you think the number one regret was I Castle was Bay no following you dreams which

Audio Transcription

In [ ]:

```
- audio transcription can be considered as a part of artificial intelligence (AI). AI in
volves developing computer systems that can perform tasks that typically require human i
ntelligence, such as understanding and interpreting audio or speech. Audio transcription
involves converting spoken language into written text, which requires AI techniques such
as speech recognition and natural language processing. These techniques enable computers
to analyze and understand audio data, recognize words and phrases, and convert them into
written form. Therefore, audio transcription is an application of AI technology in the f
ield of language processing.
```

- **Task:**
- **Download video/audio from youtube**
- **transcribe the audio to Text**
- **preprocessing the audio file**
- **visualize the audio file**

Using yt-dlp module downloading the video from youtube

In [6]:

```
import subprocess

def download_video(video_url, output_path):
    command = f"yt-dlp {video_url} --output {output_path}"
    subprocess.call(command, shell=True)

video_url='https://youtu.be/Jl3GdMd843Q'
output_path = r"D:\pytube\video\titanic_sub.mp4"
download_video(video_url, output_path)
```

Using yt-dlp module downloading the audio from youtube

In [7]:

```
def download_audio(video_url, output_path):
    command = f"yt-dlp {video_url} --extract-audio --audio-format mp3 --output {output_p
ath}"
    subprocess.call(command, shell=True)

video_url='https://youtu.be/Jl3GdMd843Q'
output_path = r"D:\pytube\audio\titanic_sub.mp3"
download_audio(video_url, output_path)
```

using FFmpeg software package,i converted the .webm audio format to .wav format

In [8]:

```
input_file = r"D:\pytube\audio\titanic_sub.webm"
output_file = r"D:\pytube\audio\titanic_sub.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_
file, output_file]
subprocess.run(command, capture_output=True)
```

In [1]:

```
pip install whisper
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/publi
c/simple/
Collecting whisper
  Downloading whisper-1.1.10.tar.gz (42 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 42.8/42.8 kB 1.9 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from whisp
er) (1.16.0)
Building wheels for collected packages: whisper
  Building wheel for whisper (setup.py) ... done
  Created wheel for whisper: filename=whisper-1.1.10-py3-none-any.whl size=41121 sha256=9
50f9cfa8c51e3e7d5ebbe8a0f1a30752c82a60d590b49d61787cbe909009c0a
  Stored in directory: /root/.cache/pip/wheels/aa/7c/1d/015619716e2facae6631312503baf3c32
20e6a9a3508cb14b6
Successfully built whisper
Installing collected packages: whisper
Successfully installed whisper-1.1.10
```

by using the whisper module which is multilingual speech recognition module,we converted the audio files which we downlaoded from youtube using yt-dlp module now we convert the downloaded audio .wav format to text

We are transcribing the english audio to text as excecuted below

In [1]:

```
import whisper
input_file = '/content/drive/MyDrive/Colab Notebooks/titanic_sub.wav'
```

In [2]:

```
model = whisper.load_model('medium')
```

```
100%|████████████████████████████████████████| 1.42G/1.42G [00:11<00:00, 134MiB/s]
```

In [3]:

```
input_file = '/content/drive/MyDrive/Colab Notebooks/titanic_sub.wav'
```

In [4]:

```
transcript = model.transcribe(input_file)
text = transcript["text"]
print(text)
```

```
/usr/local/lib/python3.10/dist-packages/whisper/transcribe.py:114: UserWarning: FP16 is n
ot supported on CPU; using FP32 instead
  warnings.warn("FP16 is not supported on CPU; using FP32 instead")
```

```
 It is a race against time as the search for the missing Titanic intensifies. The rescuer
s are working all night to save the Titanic sub-crew. The oxygen level in the submersible
is depleting as we speak. There are five passengers on board. What are the challenges tha
t lie ahead for the rescuers and what are the survival chances of the explorers? To answe
r all of this, we have with us Greg Swenson, founder of Brigg Macadam. He is also an inte
rnational affairs expert. Thank you so much for joining us and beyond, sir. Good to be her
e. Thank you. Now, Mr. Swenson, the search and rescue crew have detected underwater sound
s. They have described them as banging noises. Do you think that this will help the crew
members in locating the exact position of the missing Titan? I'm not sure about exact pos
ition, but I think it's a good sign. And the fact that they were coming in 30-minute inte
rvals could indicate that the people on the vessel are taking turns or purposely making t
hese noises in intervals, which would make it a lot more attractive to the search vessels
. So look, the U.S. and Canadian P3 and P8 have been dropping sonobuoys, and that's the C
anadian sonobuoy is what picked up the noise. So it's a good sign. As for pinpointing exa
ctly where it is, that's going to require, I think, some more luck, for sure, as well as
the French Victor 6000 remotely operated vessel and as well as the U.S. Navy vessel that
has probably arrived overnight in the region. Right, Mr. Swenson, I was just coming to th
at. What are the challenges that lie ahead for the crew members and what are the protocol
s that they are following in search of the missing passengers? Well, I mean, the good new
s now is they're still considering it an active search. You know, granted, there's real d
```

ifficulty locating a 22-foot vessel in the vast ocean, you know, and the area they're sea
rching is two times the state of Connecticut. It's a real challenge just for that difficu
lty. But also how to bring it to the surface if they do find it. That's another challenge
, given how deep the vessel probably is, you know, 2,000 feet below the surface or 2 mile
s below the surface. And also, as you pointed out, the oxygen supply is due to run out th
is morning, this morning East Coast time. So then there's also the possibility of limited
rations. So, you know, this is a race against time. It is still an active search. There i
s probably a point, and I don't know when, but in the next few days where it becomes wher
e it moves into the recovery phase if they give up hope that the passengers are still ali
ve. Absolutely. I was just coming to that. The sub was deployed on Sunday with 96 hours w
orth of oxygen, and it is expected to run out any moment. Talk to us about the severity of
the situation. Yeah, it really can't be more severe. In many ways reminds me of the Chile
an mine 10 years ago, which ended well, but not without great risk. So, look, there are f
ive surface vessels and two remotely operated vessels. We've got the Fado system, which i
s the flyway deep ocean salvage system. These are the best search and rescue teams in the
world, the U.S. Navy, the Canadian Navy, and then you have some assistance from other cou
ntries like the French. You know, there is a all hands on deck, obviously, but it is a re
al challenge. It is an unbelievable challenge given that, you know, the vessel was suppos
ed to surface on Sunday. It was supposed to be, you know, just a half a day trip. And, yo
u know, for safety, they had, you know, 96 hours of oxygen, but it's looking difficult at
this point because it is now almost Thursday morning on the East Coast. Absolutely. The o
xygen level is depleting as we speak, and apparently there were reports that claim that a
former employee of the sub in 2018 voiced concerns regarding the titan citing potential d
anger to the passengers. However, no action was taken. Is it true? Well, it's difficult t
o say if it's true or not. I think there will be, you know, a lot of, you know, repercuss
ions from this.

- **audio to text converted using whisper module,where the audio file references to recently occurred titanic submarine accident**

Whisper supports around 99 languages,now trying with Tamil Indian language

In [16]:

```
video_url = 'https://youtu.be/592mNGkpYig'
output_path = "/content/drive/MyDrive/Colab Notebooks/audio/tamil_song.mp3"
download_audio(video_url, output_path)
```

By using ffmpeg we are converting audio format .webm to .wav format

In [18]:

```
!ffmpeg -i '/content/drive/MyDrive/Colab Notebooks/audio/tamil_song.webm' '/content/driv
e/MyDrive/Colab Notebooks/audio/tamil_song.wav'
```

```
ffmpeg version 4.2.7-0ubuntu0.1 Copyright (c) 2000-2022 the FFmpeg developers
  built with gcc 9 (Ubuntu 9.4.0-1ubuntu1~20.04.1)
  configuration: --prefix=/usr --extra-version=0ubuntu0.1 --toolchain=hardened --libdir=/
usr/lib/x86_64-linux-gnu --incdir=/usr/include/x86_64-linux-gnu --arch=amd64 --enable-gpl
--disable-stripping --enable-avresample --disable-filter=resample --enable-avisynth --ena
ble-gnutls --enable-ladspa --enable-libaom --enable-libass --enable-libbluray --enable-l
ibbs2b --enable-libcaca --enable-libcdio --enable-libcodec2 --enable-libflite --enable-l
ibfontconfig --enable-libfreetype --enable-libfribidi --enable-libgme --enable-libgsm --
enable-libjack --enable-libmp3lame --enable-libmysofa --enable-libopenjpeg --enable-libo
penmpt --enable-libopus --enable-libpulse --enable-librsvg --enable-librubberband --enab
le-libshine --enable-libsnappy --enable-libsoxr --enable-libspeex --enable-libssh --enabl
e-libtheora --enable-libtwolame --enable-libvidstab --enable-libvorbis --enable-libvpx -
-enable-libwavpack --enable-libwebp --enable-libx265 --enable-libxml2 --enable-libxvid -
-enable-libzmq --enable-libzvbi --enable-lv2 --enable-omx --enable-openal --enable-openc
l --enable-opengl --enable-sdl2 --enable-libdc1394 --enable-libdrm --enable-libiec61883
--enable-nvenc --enable-chromaprint --enable-frei0r --enable-libx264 --enable-shared
  libavutil      56. 31.100 / 56. 31.100
  libavcodec     58. 54.100 / 58. 54.100
  libavformat    58. 29.100 / 58. 29.100
  libavdevice    58.  8.100 / 58.  8.100
  libavfilter     7. 57.100 /  7. 57.100
  libavresample   4.  0.  0 /  4.  0.  0
```

```
  libswscale      5.  5.100 /  5.  5.100
  libswresample   3.  5.100 /  3.  5.100
  libpostproc    55.  5.100 / 55.  5.100
Input #0, matroska,webm, from '/content/drive/MyDrive/Colab Notebooks/audio/tamil_song.we
bm':
  Metadata:
    encoder         : google/video-file
  Duration: 00:03:14.78, start: -0.007000, bitrate: 124 kb/s
    Stream #0:0(eng): Audio: opus, 48000 Hz, stereo, fltp (default)
Stream mapping:
  Stream #0:0 -> #0:0 (opus (native) -> pcm_s16le (native))
Press [q] to stop, [?] for help
Output #0, wav, to '/content/drive/MyDrive/Colab Notebooks/audio/tamil_song.wav':
  Metadata:
    ISFT            : Lavf58.29.100
    Stream #0:0(eng): Audio: pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 15
36 kb/s (default)
    Metadata:
      encoder         : Lavc58.54.100 pcm_s16le
size=   36517kB time=00:03:14.76 bitrate=1535.9kbits/s speed= 323x
video:0kB audio:36517kB subtitle:0kB other streams:0kB global headers:0kB muxing overhead
: 0.000209%
```

In [19]:

```
input_file='/content/drive/MyDrive/Colab Notebooks/audio/tamil_song.wav'
transcript = model.transcribe(input_file)
text = transcript["text"]
print(text)
```

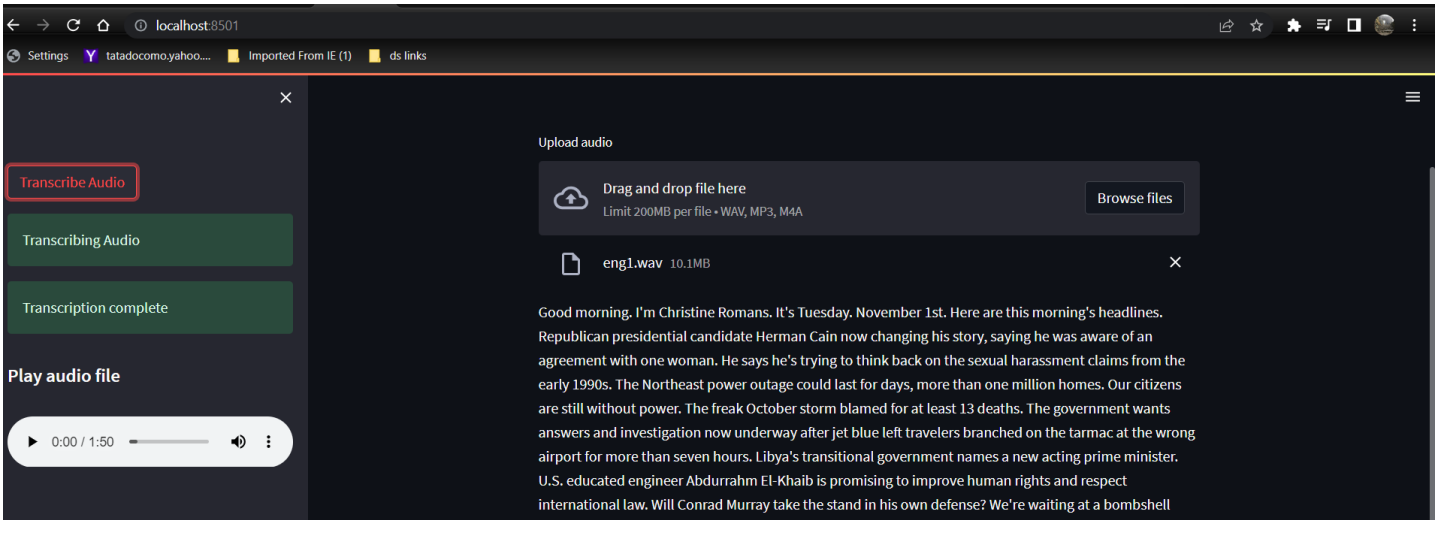
```
/usr/local/lib/python3.10/dist-packages/whisper/transcribe.py:114: UserWarning: FP16 is n
ot supported on CPU; using FP32 instead
  warnings.warn("FP16 is not supported on CPU; using FP32 instead")
```

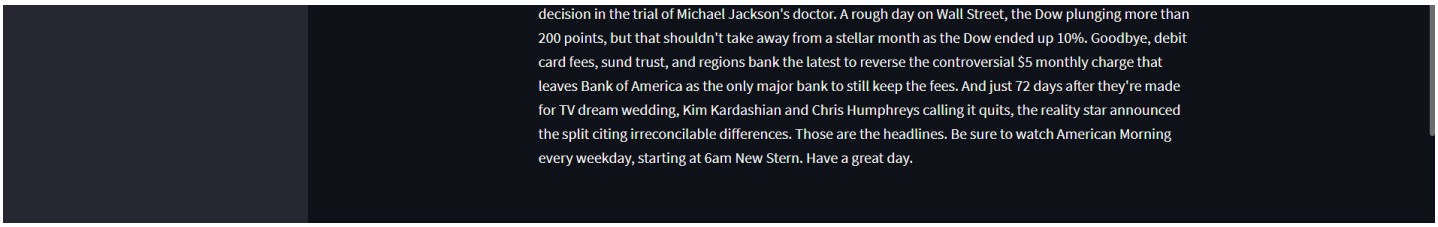
ஓரு தலக் காதலை தந்தே... இந்த தருதல மனஸுக்குள் வந்தே... ஓரு தலக் காதலை தந்தே... இந்த தருதல மனஸுக்குள் vந்தே... காதலிக் ககைடு, இல்ல... சொல்லித் தரlar, வா ?IZ. சேட்டு வ?சamiento ஆசை, எல்லா அள்ளி தர, வா fitin' siturar athar athar Switzerland இன்னாூட்டு டேயில் description டேயில் descriptionVa, advertisement இது என்ன? me had a spell உன்னால்band புடிக்காமல் ஓட்டு நரிலdate. இனிமேண நாம்மொன்று ஓள். அடி ஆத்தி... இது என்ன hm? உன் தாளட்டாம் நாம் Cherry. புடிக்காமல் ஓட்டு நரிலdate. இனிமேண நாம்மொன்று ஓள். கூkபமாகiser.. குணுகுரா... தனி FC உண்மைதாக பாருங்கள், தீட் mercy சர உன்னுடைய பெணானே கொஞ்சம் பாக்கணம் கைகள் கூற்கணம் ஜோடி சேர்ந்து லவ்வசாக உரசுத்தணம் பெண்டு வாசிச்சு கிரண்டா மெரிஜு கெட்டி மேலோ எங்கு கேக்கணும் இது என் ஜாதகம் நீங்க வீட்டில் உங்க அம்மா கிட்ட பேசித்து சொன்னீங்கன்னா காப்பிட்டு கொடுக்கும்

I was able to successfully deploy my application locally using the Streamlit library in the PyCharm IDE. The deployment process went smoothly, and I encountered no issues. I utilized the Streamlit library to create an interactive and user-friendly interface for my application.

In [3]:

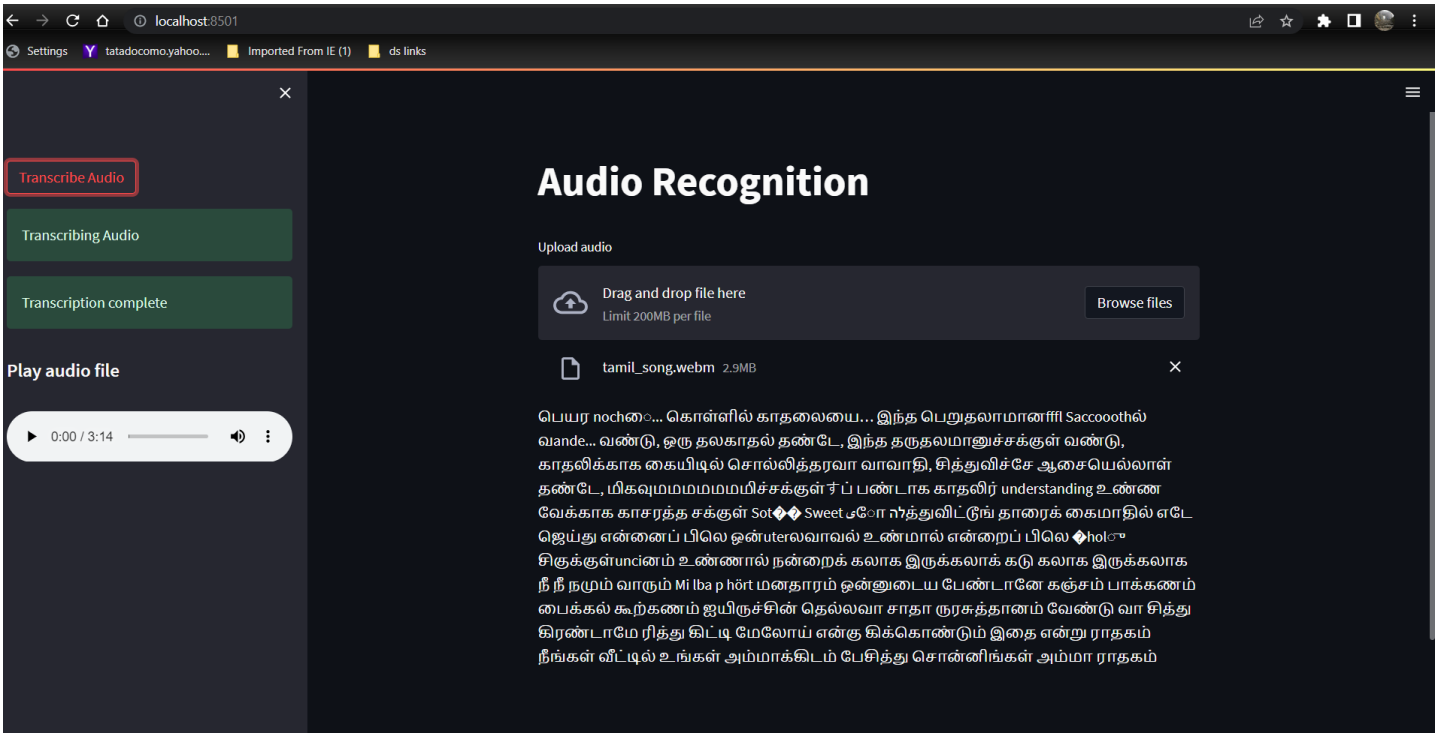
```
from IPython.display import Image, display
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\Screenshot 2023-06-25 190836.png"
image = Image(image_path)
display(image)
```





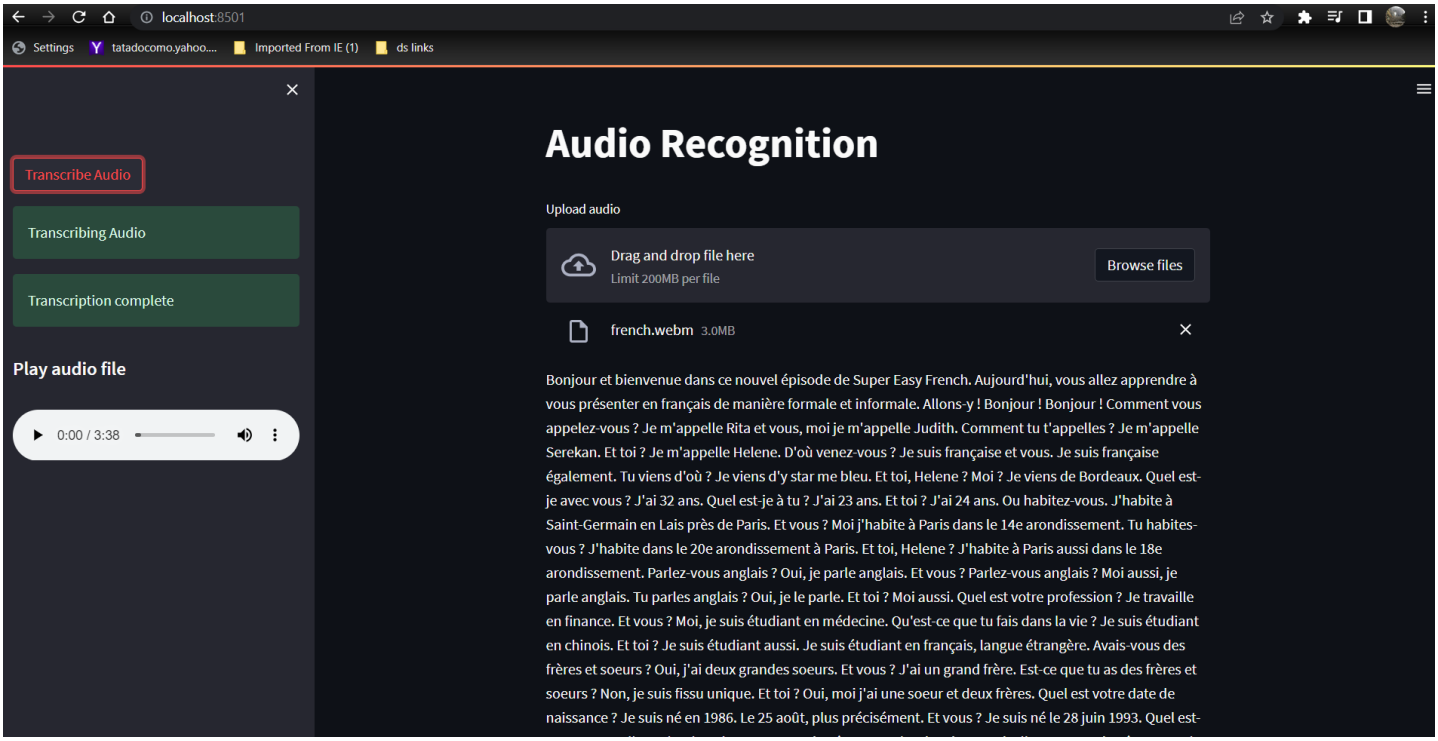
In [4]:

```
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\tamil.png"
image = Image(image_path)
display(image)
```



In [5]:

```
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\french.png"
image = Image(image_path)
display(image)
```



Audio preprocessing and visualization

In [10]:

```
import librosa
import numpy as np
import matplotlib.pyplot as plt
from IPython.display import Audio
```

audio file we using in here where we downlaoded previously from youtube yt-dlp module

In [22]:

```
audio_path = r"D:\pytube\audio\titanic_sub.wav"
audio, sample_rate = librosa.load(audio_path)
```

In [23]:

```
audio
```

Out[23]:

```
array([-3.6996323e-07,  5.3847708e-07, -6.6838982e-07, ...,
        0.0000000e+00,  0.0000000e+00,  0.0000000e+00], dtype=float32)
```

- **Sample rate, refers to the number of samples of audio that are captured per second. It is a crucial parameter in digital audio processing and represents the frequency at which the audio signal is sampled.**

In [24]:

```
sample_rate
```

Out[24]:

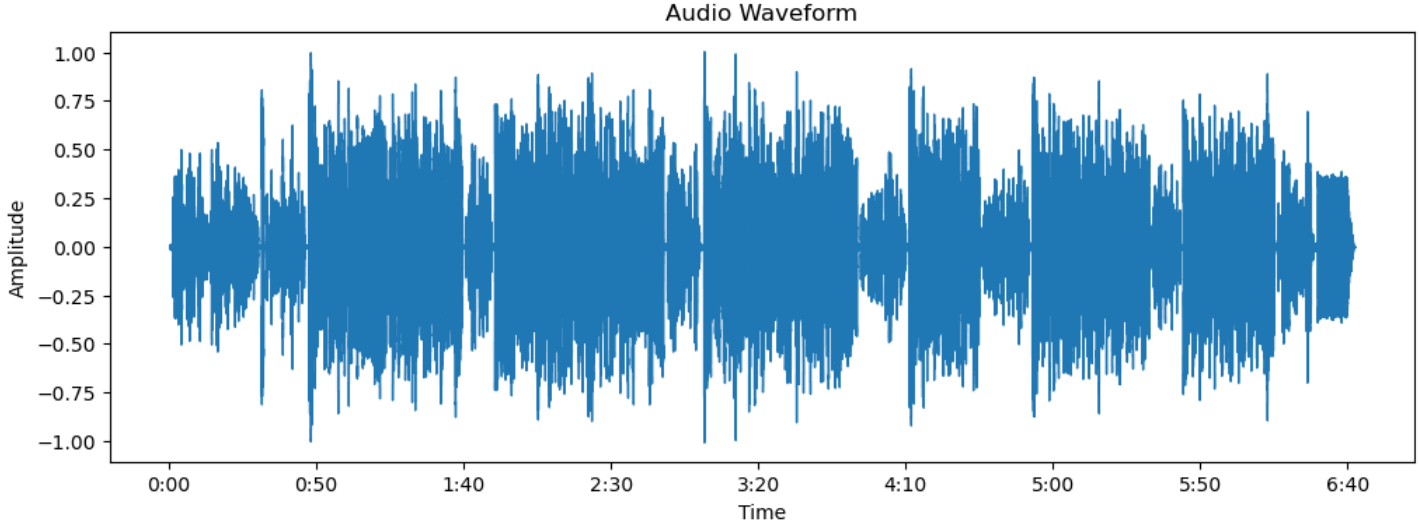
```
22050
```

Since the audio already have sample_rate 22050 so we dont have to resample it

Visualize the audio file to waveform

In [26]:

```
plt.figure(figsize=(12, 4))
librosa.display.waveshow(audio, sr=sample_rate)
plt.title('Audio Waveform')
plt.xlabel('Time')
plt.ylabel('Amplitude')
plt.show()
```

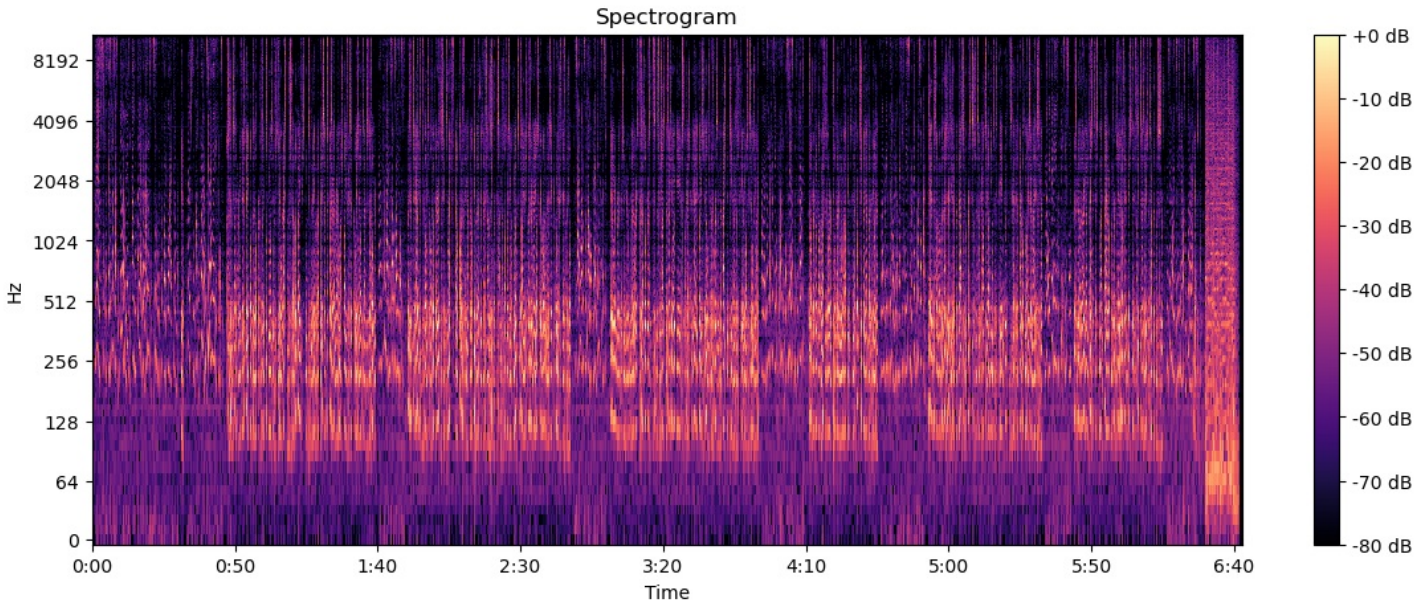


spectrogram

In [29]:

```
spectrogram = librosa.amplitude_to_db(librosa.stft(audio), ref=np.max)
plt.figure(figsize=(14, 5))
librosa.display.specshow(spectrogram,sr=target_sample_rate,x_axis='time',y_axis='log')
plt.colorbar(format='%+2.0f dB')
plt.title('Spectrogram')
plt.show()
```

```
C:\Users\lenovo\AppData\Local\Temp\ipykernel_23668\4070593631.py:1: UserWarning: amplitud
e_to_db was called on complex input so phase information will be discarded. To suppress t
his warning, call amplitude_to_db(np.abs(S)) instead.
  spectrogram = librosa.amplitude_to_db(librosa.stft(audio), ref=np.max)
```



- **A spectrogram is a visual representation of the frequencies and their intensities present in an audio signal over time. It provides a way to analyze the frequency content of an audio signal and is commonly used in audio processing tasks and audio classification.**
- **Additional preprocessing steps, such as scaling or normalization, may be applied to the spectrogram before feeding it into the classifier. Machine learning algorithms, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or other classification models**

In [4]:

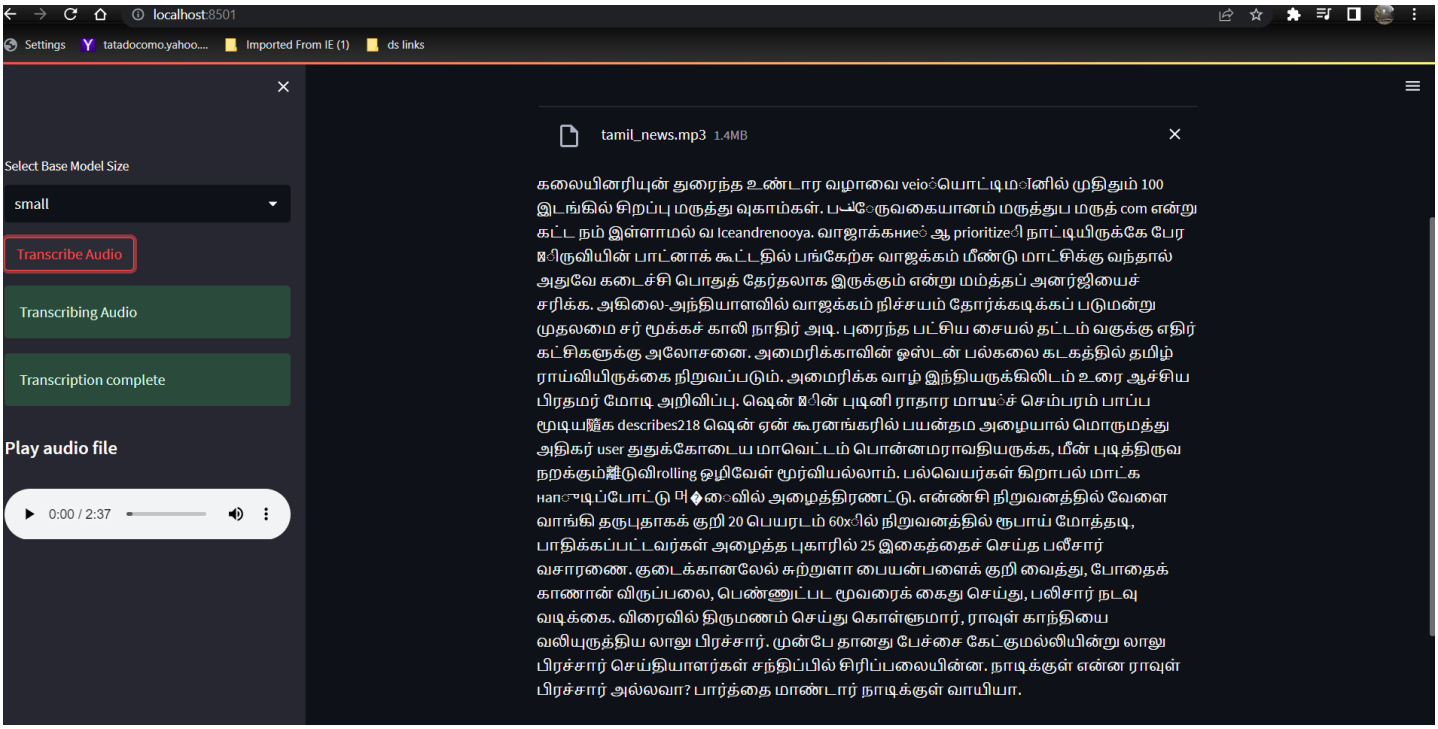
```
import subprocess

def download_audio(video_url, output_path):
    command = f"yt-dlp {video_url} --extract-audio --audio-format mp3 --output {output_p
ath}"
    try:
        subprocess.call(command, shell=True)
    except Exception as e:
        print(f"An error occurred: {e}")

video_url = 'https://youtu.be/_y8tbPCVU-0'
output_path = r'D:\pytube\audio\tamil_news.mp3'
download_audio(video_url, output_path)
```

Whisper module

In [6]:

```
from IPython.display import Image, display
```

```
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\tamilnews.png"
image = Image(image_path)
display(image)
```



In [8]:

```
input_file = r"D:\pytube\audio\tamil_news.mp3"
output_file = r"D:\pytube\audio\tamil_new.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_
file, output_file]
subprocess.run(command, capture_output=True)
```

Out[8]:

```
CompletedProcess(args=['C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe', '
-i', 'D:\\pytube\\audio\\tamil_news.mp3', 'D:\\pytube\\audio\\tamil_new.wav'], returncode
=0, stdout=b'', stderr=b"ffmpeg version 6.0-full_build-www.gyan.dev Copyright (c) 2000-20
23 the FFmpeg developers\r\n  built with gcc 12.2.0 (Rev10, Built by MSYS2 project)\r\n
configuration: --enable-gpl --enable-version3 --enable-static --disable-w32threads --disa
ble-autodetect --enable-fontconfig --enable-iconv --enable-gnutls --enable-libxml2 --ena
ble-gmp --enable-bzlib --enable-lzma --enable-libsnappy --enable-zlib --enable-librist -
-enable-libsrt --enable-libssh --enable-libzmq --enable-avisynth --enable-libbluray --en
able-libcaca --enable-sdl2 --enable-libaribb24 --enable-libdav1d --enable-libdavs2 --enab
le-libuavs3d --enable-libzvbi --enable-librav1e --enable-libsvtav1 --enable-libwebp --ena
ble-libx264 --enable-libx265 --enable-libxavs2 --enable-libxvid --enable-libaom --enable
-libjxl --enable-libopenjpeg --enable-libvpx --enable-mediafoundation --enable-libass --
enable-frei0r --enable-libfreetype --enable-libfribidi --enable-liblensfun --enable-libv
idstab --enable-libvmaf --enable-libzimg --enable-amf --enable-cuda-llvm --enable-cuvid
--enable-ffnvcodec --enable-nvdec --enable-nvenc --enable-d3d11va --enable-dxva2 --enabl
e-libmfx --enable-libshaderc --enable-vulkan --enable-libplacebo --enable-opencl --enabl
e-libcdio --enable-libgme --enable-libmodplug --enable-libopenmpt --enable-libopencore-a
mrwb --enable-libmp3lame --enable-libshine --enable-libtheora --enable-libtwolame --enabl
e-libvo-amrwbenc --enable-libilbc --enable-libgsm --enable-libopencore-amrnb --enable-li
bopus --enable-libspeex --enable-libvorbis --enable-ladspa --enable-libbs2b --enable-lib
flite --enable-libmysofa --enable-librubberband --enable-libsoxr --enable-chromaprint\r\n
libavutil      58.  2.100 / 58.  2.100\r\n  libavcodec     60.  3.100 / 60.  3.100\r\n  l
ibavformat    60.  3.100 / 60.  3.100\r\n  libavdevice    60.  1.100 / 60.  1.100\r\n  li
bavfilter     9.  3.100 /  9.  3.100\r\n  libswscale     7.  1.100 /  7.  1.100\r\n  lib
swresample    4. 10.100 /  4. 10.100\r\n  libpostproc    57.  1.100 / 57.  1.100\r\nInput
#0, mp3, from 'D:\\pytube\\audio\\tamil_news.mp3':\r\n  Metadata:\r\n    encoder
: Lavf60.3.100\r\n  Duration: 00:02:37.90, start: 0.023021, bitrate: 75 kb/s\r\n  Stream
#0:0: Audio: mp3, 48000 Hz, stereo, fltp, 75 kb/s\r\n    Metadata:\r\n      encoder
: Lavc60.3.\r\nStream mapping:\r\n  Stream #0:0 -> #0:0 (mp3 (mp3float) -> pcm_s16le (nat
ive))\r\nPress [q] to stop, [?] for help\r\nOutput #0, wav, to 'D:\\pytube\\audio\\tamil_
new.wav':\r\n  Metadata:\r\n    ISFT            : Lavf60.3.100\r\n  Stream #0:0: Audio:
pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 1536 kb/s\r\n    Metadata:\r\n
encoder         : Lavc60.3.100 pcm_s16le\r\nsize=       0kB time=00:00:00.00 bitrate=N/A
```

```
speed=N/A    \rsize=     256kB time=00:00:01.58 bitrate=1323.1kbits/s speed=   1x    \rsi
ze=   29599kB time=00:02:37.84 bitrate=1536.1kbits/s speed=  89x    \r\nvideo:0kB audio:2
9599kB subtitle:0kB other streams:0kB global headers:0kB muxing overhead: 0.000257%\r\n")
```

- **Due to memory allocation,i could use base and small pretrained whisper model when i make use of medium,large module im getiing memory error so i started use speech recognition model**

Speech recognition module

In [19]:

```
import speech_recognition as sr
from pydub import AudioSegment
from pydub.utils import make_chunks
import os
```

In [21]:

```
def process_audio(filename, language='en-US'):
    transcriptions = []
    myaudio = AudioSegment.from_wav(filename)
    chunks_length = 8000
    chunks = make_chunks(myaudio, chunks_length)
    for i, chunk in enumerate(chunks):
        chunkName = f"./chunked/{os.path.basename(filename)}_{i}.wav"
        print(f"I am exporting {chunkName}")
        chunk.export(chunkName, format="wav")
        file = chunkName
        r = sr.Recognizer()
        with sr.AudioFile(file) as source:
            audio_listened = r.listen(source)
            try:
                rec = r.recognize_google(audio_listened, language=language)
                transcriptions.append(rec)
            except sr.UnknownValueError:
                print("I don't recognize your audio")
            except sr.RequestError as e:
                print("Could not get the result.")
    return ' '.join(transcriptions)

try:
    os.makedirs("chunked")
except:
    pass
```

In [23]:

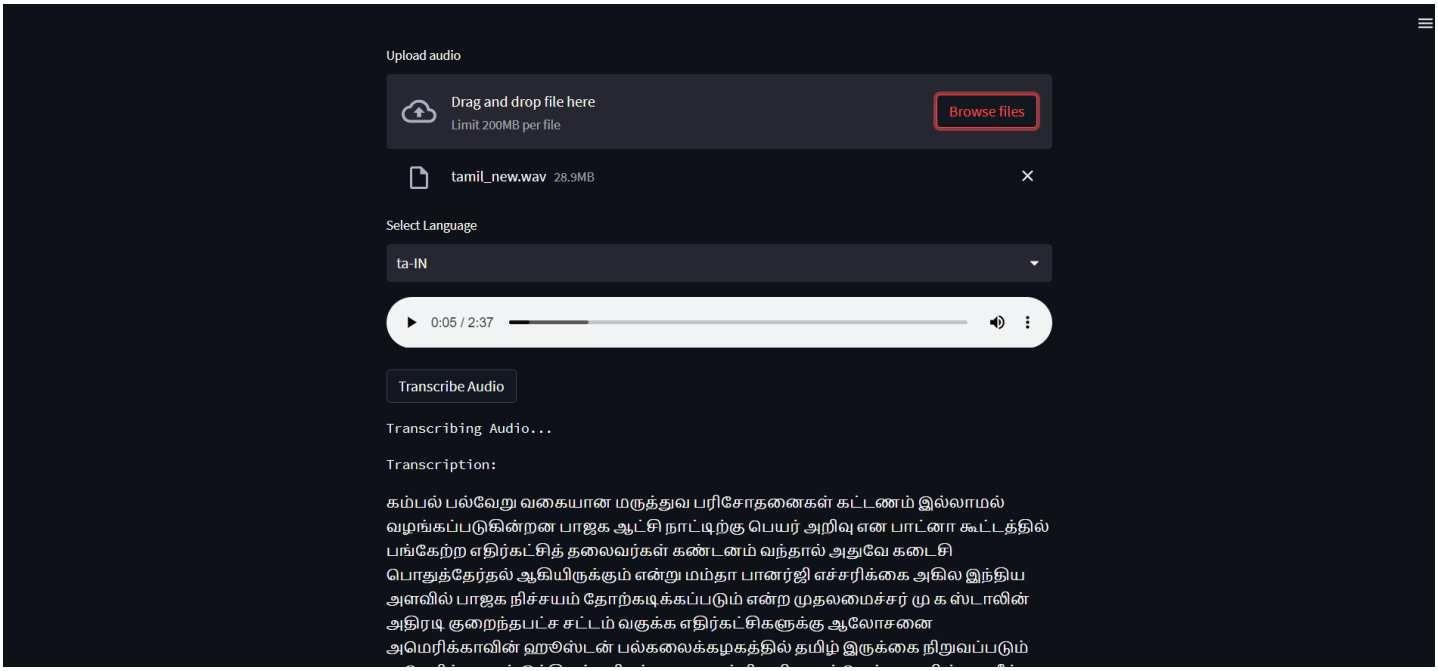
```
filename = "D:\\pytube\\audio\\tamil_new.wav"
language = 'ta-IN'
transcription_paragraph = process_audio(filename, language)

print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/tamil_new.wav_0.wav
I am exporting ./chunked/tamil_new.wav_1.wav
I am exporting ./chunked/tamil_new.wav_2.wav
I am exporting ./chunked/tamil_new.wav_3.wav
I am exporting ./chunked/tamil_new.wav_4.wav
I am exporting ./chunked/tamil_new.wav_5.wav
I am exporting ./chunked/tamil_new.wav_6.wav
I am exporting ./chunked/tamil_new.wav_7.wav
I am exporting ./chunked/tamil_new.wav_8.wav
I am exporting ./chunked/tamil_new.wav_9.wav
I am exporting ./chunked/tamil_new.wav_10.wav
I am exporting ./chunked/tamil_new.wav_11.wav
I don't recognize your audio
I am exporting ./chunked/tamil_new.wav_12.wav
```

```
I am exporting ./chunked/tamil_new.wav_13.wav
I am exporting ./chunked/tamil_new.wav_14.wav
I am exporting ./chunked/tamil_new.wav_15.wav
I am exporting ./chunked/tamil_new.wav_16.wav
I am exporting ./chunked/tamil_new.wav_17.wav
I am exporting ./chunked/tamil_new.wav_18.wav
I am exporting ./chunked/tamil_new.wav_19.wav
I don't recognize your audio
Transcription:
```
கம்பல் பல்வேறு வகையான மருத்துவ பரிசோதனைகள் கட்டணம் இல்லாமல் வழங்கப்படுகின்றன பாஜக ஆட்சி நாட்டிற்கு பெயர் அறிவு என பாட்னா கூட்டத்தில் பங்கேற்ற எதிர்கட்சித் தலைவர்கள் கண்டனம் வந்தால் அதுவே கடைசி பொதுத்தேர்தல் ஆகியிருக்கும் என்று மம்தா பானர்ஜி எச்சரிக்கை அகில இந்திய அளவில் பாஜக நிச்சயம் தோற்கடிக்கப்படும் என்ற முதலமைச்சர் மு க ஸ்டாலின் அதிரடி குறைந்தபட்ச சட்டம் வகுக்க எதிர்கட்சிகளுக்கு ஆலோசனை அமெரிக்காவின் ஹூஸ்டன் பல்கலைக்கழகத்தில் தமிழ் இருக்கை நிறுவப்படும் அமெரிக்க வாழ் இந்தியர்களிடம் உரையாற்றிய பிரதமர் சென்னையின் குடிநீர் ஆதாரமான பாக்கம் ஏரி 20 அடியை எட்டியது சென்னையின் புறநகர் களில் பெய்த மழையால் ஏரிக்கு நீர்வரத்து அதிகரிப்பு புதுக்கோட்டை மாவட்டம் பொன்னமராவதி அருகே மீன்பிடித்து வருகிறார்கள் பல்வேறு கிராம மக்கள் போட்டி போட்டு மீன்களை அள்ளிச் சென்றனர் என்எல்சி நிறுவனத்தில் வேலை வாங்கி தருவதாக கூறி 20 பெயர் இடம் 60 லட்சம் ரூபாய் மோசடி பாதிக்கப்பட்டவர்கள் அளித்த அவரை கைது செய்து போலீசார் விசாரணை கொடைக்கானலில் சுற்றுலா பயணிகளை குறிவைத்து கொடைக்கானல் விற்பனை பெண் உட்பட மூவரை கைது செய்த பொலிசார் விரைவில் திருமணம் செய்து கொள்ளும் ராகுல் காந்தியை வலியுறுத்திய லாலு பிரசாத் தனது பேச்சை கேட்கவில்லை என்று கூறியதால் செய்தியாளர்கள் சந்திப்பு பலை உடனுக்குடன் செய்திகளை தெரிந்து கொள்ள சன் நியூஸ் பிள்ளை கணக்கில் பண்ணுங்க

In [29]:

```
from IPython.display import Image, display
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\t1.png"
image = Image(image_path)
display(image)
```



Telugu language

In [17]:

```
input_file = r"D:\pytube\audio\telugu_news.mp3"
output_file = r"D:\pytube\audio\telugu_news.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_file, output_file]
subprocess.run(command, capture_output=True)
```

Out[17]:

```
CompletedProcess(args=['C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe', '-i', 'D:\\pytube\\audio\\telugu_news.mp3', 'D:\\pytube\\audio\\telugu_news.wav'], returncode=0, stdout=b'', stderr=b"ffmpeg version 6.0-full_build-www.gyan.dev Copyright (c) 2000-2023 the FFmpeg developers\r\n  built with gcc 12.2.0 (Rev10, Built by MSYS2 project)\r\n  configuration: --enable-gpl --enable-version3 --enable-static --disable-w32threads --disable-autodetect --enable-fontconfig --enable-iconv --enable-gnutls --enable-libxml2 --
```

```
isable autodetect   enable fontconfig   enable iconv   enable gnutls   enable libxml2
enable-gmp --enable-bzlib --enable-lzma --enable-libsnappy --enable-zlib --enable-libris
t --enable-libsrt --enable-libssh --enable-libzmq --enable-avisynth --enable-libbluray -
-enable-libcaca --enable-sdl2 --enable-libaribb24 --enable-libdav1d --enable-libdavs2 --
enable-libuavs3d --enable-libzvbi --enable-librav1e --enable-libsvtav1 --enable-libwebp
--enable-libx264 --enable-libx265 --enable-libxavs2 --enable-libxvid --enable-libaom --e
nable-libjxl --enable-libopenjpeg --enable-libvpx --enable-mediafoundation --enable-liba
ss --enable-frei0r --enable-libfreetype --enable-libfribidi --enable-liblensfun --enable
-libvidstab --enable-libvmaf --enable-libzimg --enable-amf --enable-cuda-llvm --enable-c
uvid --enable-ffnvcodec --enable-nvdec --enable-nvenc --enable-d3d11va --enable-dxva2 --
enable-libmfx --enable-libshaderc --enable-vulkan --enable-libplacebo --enable-opencl --
enable-libcdio --enable-libgme --enable-libmodplug --enable-libopenmpt --enable-libopenc
ore-amrwb --enable-libmp3lame --enable-libshine --enable-libtheora --enable-libtwolame --
enable-libvo-amrwbenc --enable-libilbc --enable-libgsm --enable-libopencore-amrnb --enab
le-libopus --enable-libspeex --enable-libvorbis --enable-ladspa --enable-libbs2b --enabl
e-libflite --enable-libmysofa --enable-librubberband --enable-libsoxr --enable-chromaprin
t\r\n  libavutil      58.  2.100 / 58.  2.100\r\n  libavcodec     60.  3.100 / 60.  3.100
\r\n  libavformat    60.  3.100 / 60.  3.100\r\n  libavdevice    60.  1.100 / 60.  1.100\
r\n  libavfilter     9.  3.100 /  9.  3.100\r\n  libswscale      7.  1.100 /  7.  1.100\r
\n  libswresample   4. 10.100 /  4. 10.100\r\n  libpostproc    57.  1.100 / 57.  1.100\r\
nInput #0, mp3, from 'D:\\pytube\\audio\\telugu_news.mp3':\r\n  Metadata:\r\n    encoder
: Lavf60.3.100\r\n  Duration: 00:01:41.52, start: 0.023021, bitrate: 78 kb/s\r\n  Stream
#0:0: Audio: mp3, 48000 Hz, stereo, fltp, 78 kb/s\r\n    Metadata:\r\n      encoder
: Lavc60.3.\r\nStream mapping:\r\n  Stream #0:0 -> #0:0 (mp3 (mp3float) -> pcm_s16le (nat
ive))\r\nPress [q] to stop, [?] for help\r\nOutput #0, wav, to 'D:\\pytube\\audio\\telugu
_news.wav':\r\n  Metadata:\r\n    ISFT            : Lavf60.3.100\r\n  Stream #0:0: Audio
: pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 1536 kb/s\r\n    Metadata:\r\
n      encoder         : Lavc60.3.100 pcm_s16le\r\nsize=       0kB time=00:00:00.00 bitra
te=N/A speed=N/A    \rsize=   19028kB time=00:01:41.47 bitrate=1536.2kbits/s speed= 857x
\r\nvideo:0kB audio:19028kB subtitle:0kB other streams:0kB global headers:0kB muxing over
head: 0.000400%\r\n")
```

In [25]:

```
filename = "D:\\pytube\\audio\\telugu_news.wav"
language = 'te-IN'
transcription_paragraph = process_audio(filename, language)

print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/telugu_news.wav_0.wav
I am exporting ./chunked/telugu_news.wav_1.wav
I am exporting ./chunked/telugu_news.wav_2.wav
I am exporting ./chunked/telugu_news.wav_3.wav
I am exporting ./chunked/telugu_news.wav_4.wav
I am exporting ./chunked/telugu_news.wav_5.wav
I am exporting ./chunked/telugu_news.wav_6.wav
I am exporting ./chunked/telugu_news.wav_7.wav
I am exporting ./chunked/telugu_news.wav_8.wav
I am exporting ./chunked/telugu_news.wav_9.wav
I am exporting ./chunked/telugu_news.wav_10.wav
I am exporting ./chunked/telugu_news.wav_11.wav
I am exporting ./chunked/telugu_news.wav_12.wav
I don't recognize your audio
Transcription:
జులై 10 వరకు వర్షాలు ఉండకపోవచ్చు ఎదుర్కొంటామని మంత్రి హరీష్ రావు టిఆర్ఎస్ ఎమ్మెల్యే ధర్మారెడ్డి పై కొండా మురళి ఫైర్ వచ్చే ఎ
న్నికల్లో పరకాలలో కానీ ప్రకటన ఆంధ్రప్రదేశ్లా ఏపీ వైసిపి పాలనలో క్షీణించాయని టిడిపి నారాయణ లేఖ వీశాఖ ఎంపీ కుటుంబం కిడ్నాప్ పై
సిబిఐ విచారణకు డిమాండ్ ఏపీ ఉద్యోగ సంఘం నేత విచారణ వాణిజ్యపన్నుల కుంభకోణంలో ఆరా తీస్తున్న పోలీసులు మణికొండ కిడ్స్
ప్లే స్కూల్ లో అగ్నిప్రమాదం మంటలు చూసి పరుగులు తీసిన చిన్నారులు వివేకా హత్య కేసు అప్రూవర్ దస్తగిరి కొత్త దండ అధిక వడ్డీకి
అప్పులు ఇస్తున్నట్లు ఫిర్యాదు జగిత్యాల జిల్లాలో స్కూల్ బస్సులపై ఆత్మీయ ఫిట్నెస్ లేని వాహనాలు సీజ్ ఆడబిడ్డ మాకు అపురూపం ఇంటిల్లి
పాదీ చాలా సంతోషంగా ఉన్నా మన మెగాస్టార్ సేవల్లో జగన్నాథ రథయాత్ర భక్తులతో కిక్కిరిసిన పూడి వీధులు
```

English language

In [26]:

```
filename="D:\\pytube\\audio\\titanic_sub.wav"
language = 'en-US'
transcription_paragraph = process_audio(filename, language)
```

```
print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/titanic_sub.wav_0.wav
I am exporting ./chunked/titanic_sub.wav_1.wav
I am exporting ./chunked/titanic_sub.wav_2.wav
I am exporting ./chunked/titanic_sub.wav_3.wav
I am exporting ./chunked/titanic_sub.wav_4.wav
I am exporting ./chunked/titanic_sub.wav_5.wav
I am exporting ./chunked/titanic_sub.wav_6.wav
I am exporting ./chunked/titanic_sub.wav_7.wav
I am exporting ./chunked/titanic_sub.wav_8.wav
I am exporting ./chunked/titanic_sub.wav_9.wav
I am exporting ./chunked/titanic_sub.wav_10.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_11.wav
I am exporting ./chunked/titanic_sub.wav_12.wav
I am exporting ./chunked/titanic_sub.wav_13.wav
I am exporting ./chunked/titanic_sub.wav_14.wav
I am exporting ./chunked/titanic_sub.wav_15.wav
I am exporting ./chunked/titanic_sub.wav_16.wav
I am exporting ./chunked/titanic_sub.wav_17.wav
I am exporting ./chunked/titanic_sub.wav_18.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_19.wav
I am exporting ./chunked/titanic_sub.wav_20.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_21.wav
I am exporting ./chunked/titanic_sub.wav_22.wav
I am exporting ./chunked/titanic_sub.wav_23.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_24.wav
I am exporting ./chunked/titanic_sub.wav_25.wav
I am exporting ./chunked/titanic_sub.wav_26.wav
I am exporting ./chunked/titanic_sub.wav_27.wav
I am exporting ./chunked/titanic_sub.wav_28.wav
I am exporting ./chunked/titanic_sub.wav_29.wav
I am exporting ./chunked/titanic_sub.wav_30.wav
I am exporting ./chunked/titanic_sub.wav_31.wav
I am exporting ./chunked/titanic_sub.wav_32.wav
I am exporting ./chunked/titanic_sub.wav_33.wav
I am exporting ./chunked/titanic_sub.wav_34.wav
I am exporting ./chunked/titanic_sub.wav_35.wav
I am exporting ./chunked/titanic_sub.wav_36.wav
I am exporting ./chunked/titanic_sub.wav_37.wav
I am exporting ./chunked/titanic_sub.wav_38.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_39.wav
I am exporting ./chunked/titanic_sub.wav_40.wav
I am exporting ./chunked/titanic_sub.wav_41.wav
I am exporting ./chunked/titanic_sub.wav_42.wav
I am exporting ./chunked/titanic_sub.wav_43.wav
I am exporting ./chunked/titanic_sub.wav_44.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_45.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_46.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_47.wav
I am exporting ./chunked/titanic_sub.wav_48.wav
I am exporting ./chunked/titanic_sub.wav_49.wav
I don't recognize your audio
I am exporting ./chunked/titanic_sub.wav_50.wav
I don't recognize your audio
Transcription:
is the race against time as the search for the missing Titanic intensifies the rescuers a
re working or not the oxygen level in the submersible is depleting as we speak the five p
assengers on board and what are the survival chances of the explorers to answer of this w
e have without grax vencen founder of brick Academy is also international affairs expert
thank you so much for joining after the Search and rescue crew have detected underwater s
ound that describe them as banging do you think that this will help the crew members in l
```

ocating the exact position of the missing Titan people making is noises in what what as w
ell as the French sector 6000 and as well as the US Navy Faisal in written crew members a
nd what are the protocols that they are following in search of the missing passengers bes
t just from from product Wes and also as you put it search there is problem Apps without
coming to that that is stop was deployed on Sunday with 96 hours worth of Oxygen and Spec
tre to run out any moment talk to Sabha the variety of the situation permission US Navy t
he Canadian navy and you have some assistant real challenge and after it's working just c
all at this point because it is there is no Anushka employee of the Sab in 2018 voice conc
erns regarding the Titan exciting potential danger to the passengers you know interest in
using the absol question that husband wife and for the missing a Sab a crafts and ships h
ave been the property locate the missing the band the explorers what are the survival cha
nces of the explorers at the moment because the oxygen area and find also say even if the
Titan is located in the North Atlantic it could be nearly impossible to reach if it is st
uck on the ocean floor that's right because you can other American Navy Divas contact be
only hope have a search in Rescue Team are able to find out the 5% on board thank you

Hindi language

In [35]:

```
input_file = r"D:\pytube\audio\hindi_news.mp3"
output_file = r"D:\pytube\audio\hindi_news1.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_
file, output_file]
subprocess.run(command, capture_output=True)
```

Out[35]:

```
CompletedProcess(args=['C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe', '
-i', 'D:\\pytube\\audio\\hindi_news.mp3', 'D:\\pytube\\audio\\hindi_news1.wav'], returnco
de=0, stdout=b'', stderr=b"ffmpeg version 6.0-full_build-www.gyan.dev Copyright (c) 2000-
2023 the FFmpeg developers\r\n  built with gcc 12.2.0 (Rev10, Built by MSYS2 project)\r\n
configuration: --enable-gpl --enable-version3 --enable-static --disable-w32threads --disa
ble-autodetect --enable-fontconfig --enable-iconv --enable-gnutls --enable-libxml2 --ena
ble-gmp --enable-bzlib --enable-lzma --enable-libsnappy --enable-zlib --enable-librist -
-enable-libsrt --enable-libssh --enable-libzmq --enable-avisynth --enable-libbluray --en
able-libcaca --enable-sdl2 --enable-libaribb24 --enable-libdav1d --enable-libdavs2 --enab
le-libuavs3d --enable-libzvbi --enable-librav1e --enable-libsvtav1 --enable-libwebp --ena
ble-libx264 --enable-libx265 --enable-libxavs2 --enable-libxvid --enable-libaom --enable
-libjxl --enable-libopenjpeg --enable-libvpx --enable-mediafoundation --enable-libass --
enable-frei0r --enable-libfreetype --enable-libfribidi --enable-liblensfun --enable-libv
idstab --enable-libvmaf --enable-libzimg --enable-amf --enable-cuda-llvm --enable-cuvid
--enable-ffnvcodec --enable-nvdec --enable-nvenc --enable-d3d11va --enable-dxva2 --enabl
e-libmfx --enable-libshaderc --enable-vulkan --enable-libplacebo --enable-opencl --enabl
e-libcdio --enable-libgme --enable-libmodplug --enable-libopenmpt --enable-libopencore-a
mrwb --enable-libmp3lame --enable-libshine --enable-libtheora --enable-libtwolame --enabl
e-libvo-amrwbenc --enable-libilbc --enable-libgsm --enable-libopencore-amrnb --enable-li
bopus --enable-libspeex --enable-libvorbis --enable-ladspa --enable-libbs2b --enable-lib
flite --enable-libmysofa --enable-librubberband --enable-libsoxr --enable-chromaprint\r\n
libavutil      58.  2.100 / 58.  2.100\r\n  libavcodec     60.  3.100 / 60.  3.100\r\n  l
ibavformat    60.  3.100 / 60.  3.100\r\n  libavdevice    60.  1.100 / 60.  1.100\r\n  li
bavfilter     9.  3.100 /  9.  3.100\r\n  libswscale     7.  1.100 /  7.  1.100\r\n  lib
swresample   4. 10.100 /  4. 10.100\r\n  libpostproc    57.  1.100 / 57.  1.100\r\nInput
#0, mp3, from 'D:\\pytube\\audio\\hindi_news.mp3':\r\n  Metadata:\r\n    encoder
: Lavf60.3.100\r\n  Duration: 00:02:31.56, start: 0.023021, bitrate: 77 kb/s\r\n  Stream
#0:0: Audio: mp3, 48000 Hz, stereo, fltp, 77 kb/s\r\n    Metadata:\r\n      encoder
: Lavc60.3.\r\nStream mapping:\r\n  Stream #0:0 -> #0:0 (mp3 (mp3float) -> pcm_s16le (nat
ive))\r\nPress [q] to stop, [?] for help\r\nOutput #0, wav, to 'D:\\pytube\\audio\\hindi_
news1.wav':\r\n  Metadata:\r\n    ISFT            : Lavf60.3.100\r\n  Stream #0:0: Audio
: pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 1536 kb/s\r\n    Metadata:\r\
n      encoder         : Lavc60.3.100 pcm_s16le\r\nsize=       0kB time=00:00:00.00 bitra
te=N/A speed=N/A    \rsize=     512kB time=00:00:03.07 bitrate=1364.9kbits/s speed=1.94x
\rsize=   16896kB time=00:01:31.51 bitrate=1512.5kbits/s speed=  44x    \rsize=   28412kB
time=00:02:31.51 bitrate=1536.2kbits/s speed=66.3x    \r\nvideo:0kB audio:28412kB subtitl
e:0kB other streams:0kB global headers:0kB muxing overhead: 0.000268%\r\n")
```

In [36]:

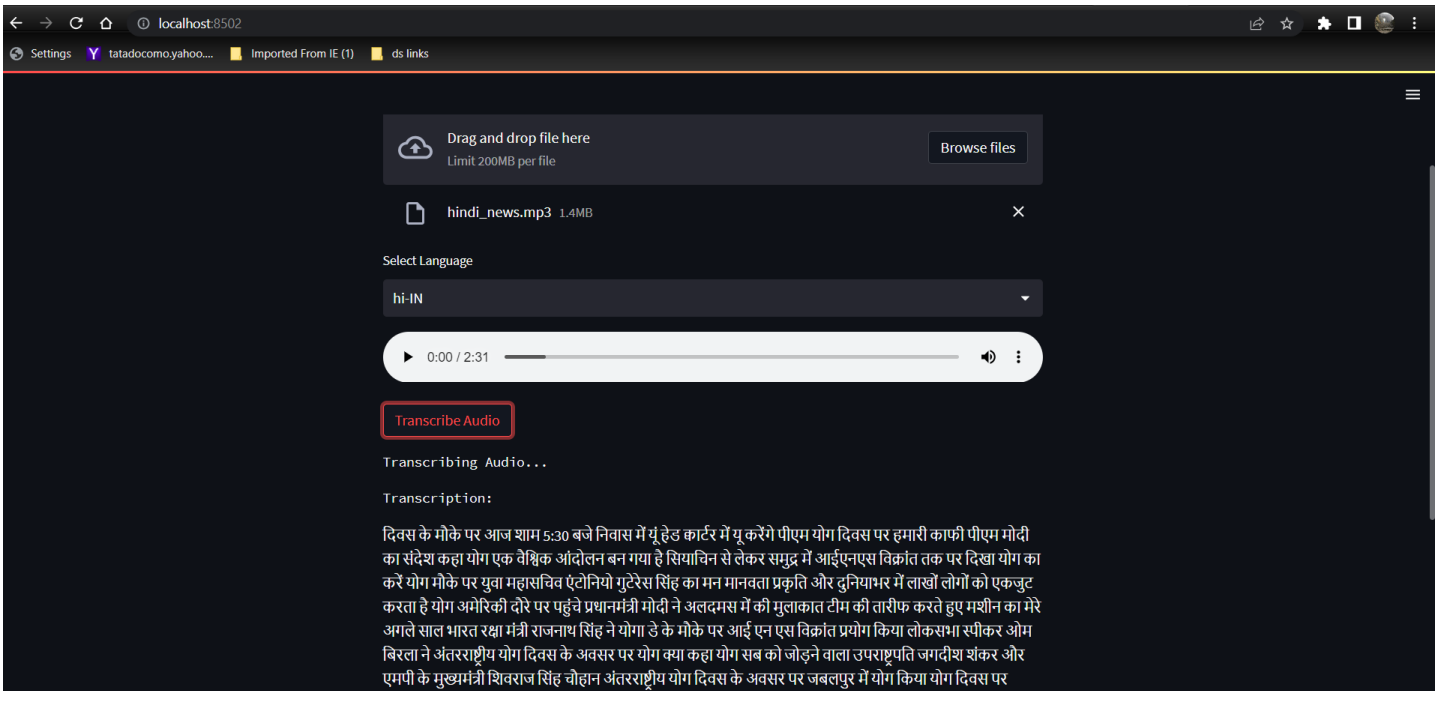
```
filename="D:\\pytube\\audio\\hindi_news1.wav"
```

```
language = 'hi-IN'
transcription_paragraph = process_audio(filename, language)

print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/hindi_news1.wav_0.wav
I am exporting ./chunked/hindi_news1.wav_1.wav
I am exporting ./chunked/hindi_news1.wav_2.wav
I am exporting ./chunked/hindi_news1.wav_3.wav
I am exporting ./chunked/hindi_news1.wav_4.wav
I am exporting ./chunked/hindi_news1.wav_5.wav
I am exporting ./chunked/hindi_news1.wav_6.wav
I am exporting ./chunked/hindi_news1.wav_7.wav
I am exporting ./chunked/hindi_news1.wav_8.wav
I am exporting ./chunked/hindi_news1.wav_9.wav
I am exporting ./chunked/hindi_news1.wav_10.wav
I am exporting ./chunked/hindi_news1.wav_11.wav
I am exporting ./chunked/hindi_news1.wav_12.wav
I am exporting ./chunked/hindi_news1.wav_13.wav
I am exporting ./chunked/hindi_news1.wav_14.wav
I am exporting ./chunked/hindi_news1.wav_15.wav
I am exporting ./chunked/hindi_news1.wav_16.wav
I am exporting ./chunked/hindi_news1.wav_17.wav
I am exporting ./chunked/hindi_news1.wav_18.wav
Transcription:
दिवस के मौके पर आज शाम 5:30 बजे निवास में यूं हेड क्वार्टर में यू करेंगे पीएम योग दिवस पर हमारी काफी पीएम मोदी का सं देश कहा योग एक वैश्विक आंदोलन बन गया है सियाचिन से लेकर समुद्र में आई एन एस विक्रांत तक पर दिखाओ युवा का क्रेज मौ के पर युवा महासचिव एंटोनियो गुटेरेस सिंह का मन मानवता प्रकृति और दुनियाभर में लाखों लोगों को एकजुट करता है योग अमेरि की दौरे पर पहुंचे प्रधानमंत्री मोदी से अलदमस में की मुलाकात टीम की तारीफ करते हुए मशीन का मेरे अगले साल भारत रक्षा मंत्र ी राजनाथ सिंह ने योगा डे के मौके पर आई एन एस विक्रांत प्रयोग किया लोकसभा स्पीकर ओम बिरला ने अंतरराष्ट्रीय योग दिवस के अवसर पर योग क्या कहा योग सब को जोड़ने वाला उपराष्ट्रपति जगदीश शंकर और एमपी के मुख्यमंत्री शिवराज सिंह चौहान अंत रराष्ट्रीय योग दिवस के अवसर पर जबलपुर में योग किया योग दिवस पर पुलिस में तैनात सेना के जवानों नियुक्त किया वहीं दूसरी ओर लद्दाख में भारतीय जवानों ने बैंकों में योग किया योग दिवस के अवसर पर गुजरात के राजकोट में लोगों ने योग किया तमिलना डु के रामेश्वरम में लोगों ने पानी के अंदर जो केजरीवाल ने विपक्षी नेताओं को चिट्ठी लिखी चिट्ठी में अध्यादेश को संसद में आरानी पर सबसे पहले चर्चा की मांग उठाई अभी पुलिस पर दायर याचिका पर दिल्ली हाईकोर्ट ने जल्द सुनवाई से इनकार किया बीजेपी रा हुल गांधी की विदेश यात्रा पर सवाल उठाए ट्विटर का विदेशी एजेंसियों के विरोधी समूहों के साथ राहुल ने कई गुप्त बैठते हैं मुंबई और पुणे में संजय रावत के करीबी के ठिकानों पर छापेमारी पर कर्नाटक के सीएम सिद्धारमैया जो भी मंत्री से करेंगे मुलाकात भी सर कार नहीं सीएनजी पर चलने वाली टैक्सी परमिट वैलिडिटी 15 साल तक प्रदूषण सिंह मामले में कोर्ट में आज सुनवाई होगी बंगाल में पंचायत चुनाव से पहले फिर इंसाफ प्रदर्शनकारियों ने कई दुकानों में आग लगाई में बुजुर्ग की हत्या का दिल दहला देने वाला किस ी भी बदमाशों ने चाकू मारकर बुजुर्ग की हत्या के गुरुग्राम में भारी बारिश के बाद दिल्ली जॉब पेपर फाइल भारी बारिश के बाद डूबी शहर की सड़कें
```

In [37]:

```
from IPython.display import Image, display
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\t2.png"
image = Image(image_path)
display(image)
```



पुलिस में तैनात सेना के जवानों नियुक्त किया वहीं दूसरी ओर लद्दाख में भारतीय जवानों ने बैंकों में योग किया योग दिवस के
अवसर पर गुजरात के राजकोट में लोगों ने योग किया तमिलनाडु के रामेश्वरम में लोगों ने पानी के अंदर जो केजरीवाल ने

Kannada language

In [32]:

```
video_url='https://youtu.be/Byeqrm9TWko'
output_path = r"D:\pytube\audio\kannada_news.mp3"
download_audio(video_url, output_path)
```

In [38]:

```
input_file = r"D:\pytube\audio\kannada_news.mp3"
output_file = r"D:\pytube\audio\kannada_news1.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_
file, output_file]
subprocess.run(command, capture_output=True)
```

Out[38]:

```
CompletedProcess(args=['C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe', '
-i', 'D:\\pytube\\audio\\kannada_news.mp3', 'D:\\pytube\\audio\\kannada_news1.wav'], retu
rncode=0, stdout=b'', stderr=b"ffmpeg version 6.0-full_build-www.gyan.dev Copyright (c) 2
000-2023 the FFmpeg developers\r\n  built with gcc 12.2.0 (Rev10, Built by MSYS2 project)
\r\n  configuration: --enable-gpl --enable-version3 --enable-static --disable-w32threads
--disable-autodetect --enable-fontconfig --enable-iconv --enable-gnutls --enable-libxml2
--enable-gmp --enable-bzlib --enable-lzma --enable-libsnappy --enable-zlib --enable-libr
ist --enable-libsrt --enable-libssh --enable-libzmq --enable-avisynth --enable-libbluray
--enable-libcaca --enable-sdl2 --enable-libaribb24 --enable-libdav1d --enable-libdavs2 -
-enable-libuavs3d --enable-libzvbi --enable-librav1e --enable-libsvtav1 --enable-libwebp
--enable-libx264 --enable-libx265 --enable-libxavs2 --enable-libxvid --enable-libaom --e
nable-libjxl --enable-libopenjpeg --enable-libvpx --enable-mediafoundation --enable-liba
ss --enable-frei0r --enable-libfreetype --enable-libfribidi --enable-liblensfun --enable
-libvidstab --enable-libvmaf --enable-libzimg --enable-amf --enable-cuda-llvm --enable-c
uvid --enable-ffnvcodec --enable-nvdec --enable-nvenc --enable-d3d11va --enable-dxva2 --
enable-libmfx --enable-libshaderc --enable-vulkan --enable-libplacebo --enable-opencl --
enable-libcdio --enable-libgme --enable-libmodplug --enable-libopenmpt --enable-libopenc
ore-amrwb --enable-libmp3lame --enable-libshine --enable-libtheora --enable-libtwolame --
enable-libvo-amrwbenc --enable-libilbc --enable-libgsm --enable-libopencore-amrnb --enab
le-libopus --enable-libspeex --enable-libvorbis --enable-ladspa --enable-libbs2b --enabl
e-libflite --enable-libmysofa --enable-librubberband --enable-libsoxr --enable-chromaprin
t\r\n  libavutil      58.  2.100 / 58.  2.100\r\n  libavcodec     60.  3.100 / 60.  3.100
\r\n  libavformat    60.  3.100 / 60.  3.100\r\n  libavdevice    60.  1.100 / 60.  1.100\
r\n  libavfilter     9.  3.100 /  9.  3.100\r\n  libswscale      7.  1.100 /  7.  1.100\r
\n  libswresample   4. 10.100 /  4. 10.100\r\n  libpostproc    57.  1.100 / 57.  1.100\r\
nInput #0, mp3, from 'D:\\pytube\\audio\\kannada_news.mp3':\r\n  Metadata:\r\n    encoder
: Lavf60.3.100\r\n  Duration: 00:01:49.97, start: 0.023021, bitrate: 81 kb/s\r\n  Stream
#0:0: Audio: mp3, 48000 Hz, stereo, fltp, 81 kb/s\r\n    Metadata:\r\n      encoder
: Lavc60.3.\r\nStream mapping:\r\n  Stream #0:0 -> #0:0 (mp3 (mp3float) -> pcm_s16le (nat
ive))\r\nPress [q] to stop, [?] for help\r\nOutput #0, wav, to 'D:\\pytube\\audio\\kannad
a_news1.wav':\r\n  Metadata:\r\n    ISFT            : Lavf60.3.100\r\n  Stream #0:0: Aud
io: pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 1536 kb/s\r\n    Metadata:\
r\n      encoder         : Lavc60.3.100 pcm_s16le\r\nsize=       0kB time=00:00:00.00 bit
rate=N/A speed=N/A    \rsize=   20611kB time=00:01:49.92 bitrate=1536.0kbits/s speed= 349
x    \r\nvideo:0kB audio:20611kB subtitle:0kB other streams:0kB global headers:0kB muxing
overhead: 0.000370%\r\n")
```

In [39]:

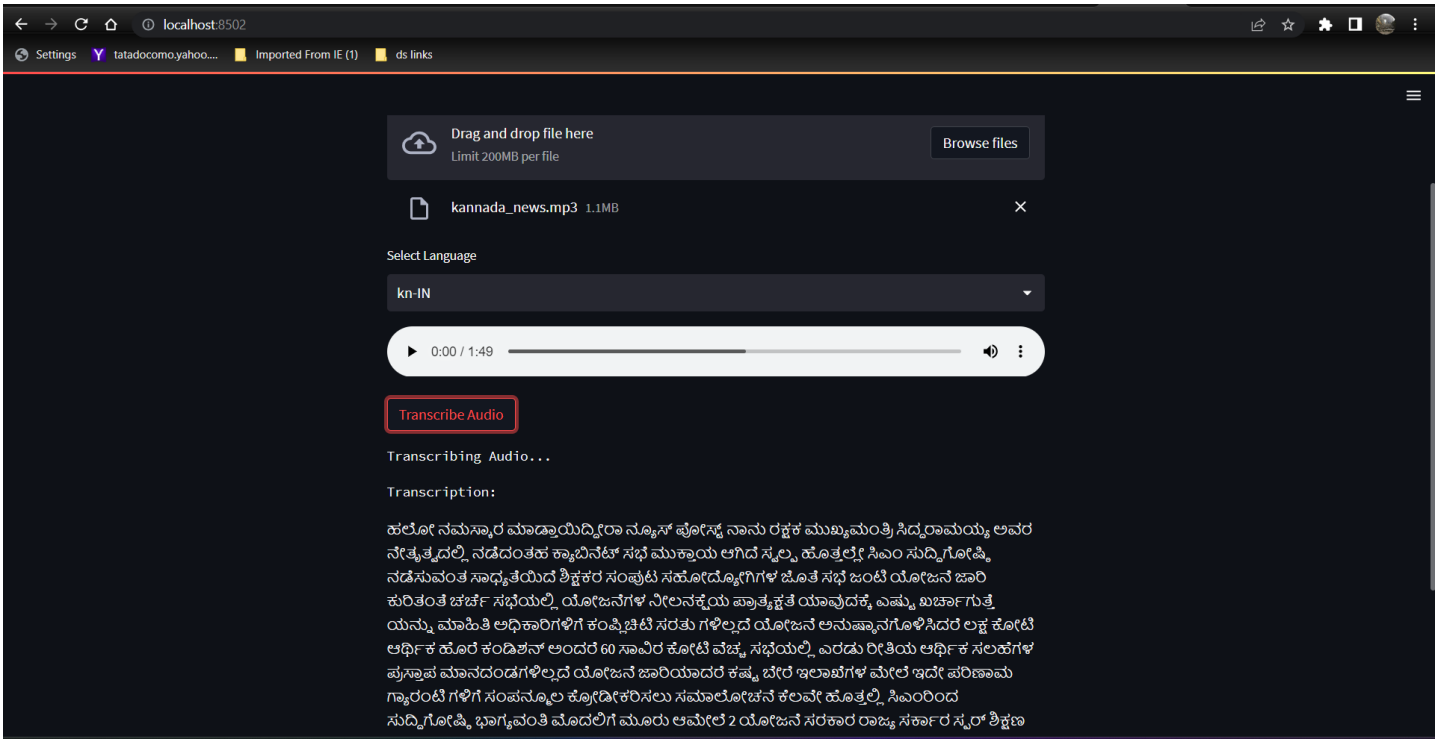
```
filename="D:\\pytube\\audio\\kannada_news1.wav"
language = 'kn-IN'
transcription_paragraph = process_audio(filename, language)

print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/kannada_news1.wav_0.wav
I am exporting ./chunked/kannada_news1.wav_1.wav
I am exporting ./chunked/kannada_news1.wav_2.wav
```

```
I am exporting ./chunked/kannada_news1.wav_3.wav
I am exporting ./chunked/kannada_news1.wav_4.wav
I am exporting ./chunked/kannada_news1.wav_5.wav
I am exporting ./chunked/kannada_news1.wav_6.wav
I am exporting ./chunked/kannada_news1.wav_7.wav
I am exporting ./chunked/kannada_news1.wav_8.wav
I am exporting ./chunked/kannada_news1.wav_9.wav
I am exporting ./chunked/kannada_news1.wav_10.wav
I am exporting ./chunked/kannada_news1.wav_11.wav
I am exporting ./chunked/kannada_news1.wav_12.wav
I am exporting ./chunked/kannada_news1.wav_13.wav
Transcription:
ಹಲೋ ನಮಸ್ಕಾರ ಮಾಡ್ತಾಯಿದ್ದೀರಾ ನ್ಯೂಸ್ ಪೋಸ್ಟ್ ನಾನು ರಕ್ಷಕ ಮುಖ್ಯಮಂತ್ರಿ ಸಿದ್ದರಾಮಯ್ಯ ಅವರ ನೇತೃತ್ವದಲ್ಲಿ ನಡೆದಂತಹ ಕ್ಯಾಬಿನೆಟ್ ಸಭೆ ಮುಕ್ತಾಯ ಆಗಿದೆ ಸ್ವಲ್ಪ ಹೊತ್ತಲ್ಲೇ ಸಿಎಂ ಸುದ್ದಿಗೋಷ್ಠಿ ನಡೆಸುವಂತ ಸಾಧ್ಯತೆಯಿದೆ ಶಿಕ್ಷಕರ ಸಂಪುಟ ಸಹೋದ್ಯೋಗಿಗಳ ಜೊತೆ ಸಭೆ ಜಂಟಿ ಯೋಜನೆ ಜಾರಿ ಕುರಿತಂತೆ ಚರ್ಚೆ ಸಭೆಯಲ್ಲಿ ಯೋಜನೆಗಳ ನೀಲನಕ್ಷೆಯ ಪ್ರಾತ್ಯಕ್ಷತೆ ಯಾವುದಕ್ಕೆ ಎಷ್ಟು ಖರ್ಚಾಗುತ್ತೆ ಯನ್ನು ಮಾಹಿತಿ ಅಧಿಕಾರಿಗಳಿಗೆ ಕಂಪ್ಲೀಟಿ ಸರತು ಗಳಿಲ್ಲದೆ ಯೋಜನೆ ಅನುಷ್ಠಾನಗೊಳಿಸಿದರೆ ಲಕ್ಷ ಕೋಟಿ ಆರ್ಥಿಕ ಹೊರೆ ಕಂಡಿಶನ್ ಅಂದರೆ 60 ಸಾವಿರ ಕೋಟಿ ವೆಚ್ಚ ಸಭೆಯಲ್ಲಿ ಎರಡು ರೀತಿಯ ಆರ್ಥಿಕ ಸಲಹೆಗಳ ಪ್ರಸ್ತಾಪ ಮಾನದಂಡಗಳಿಲ್ಲದೆ ಯೋಜನೆ ಜಾರಿಯಾದರೆ ಕಷ್ಟ ಬೇರೆ ಇಲಾಖೆಗಳ ಮೇಲೆ ಇದೇ ಪರಿಣಾಮ ಗ್ಯಾರಂಟಿ ಗಳಿಗೆ ಸಂಪನ್ಮೂಲ ಕ್ರೂಢೀಕರಿಸಲು ಸಮಾಲೋಚನೆ ಕೆಲವೇ ಹೊತ್ತಲ್ಲಿ ಸಿಎಂರಿಂದ ಸುದ್ದಿಗೋಷ್ಠಿ ಭಾಗ್ಯವಂತಿ ಮೊದಲಿಗೆ ಮೂರು ಆಮೇಲೆ 2 ಯೋಜನೆ ಸರಕಾರ ರಾಜ್ಯ ಸರ್ಕಾರ ಸ್ಪರ್ ಶಿಕ್ಷಣ ನೀತಿಗೆ
```

In [40]:

```
from IPython.display import Image, display
image_path = "C:\\Users\\lenovo\\OneDrive\\Desktop\\t3.png"
image = Image(image_path)
display(image)
```



- **Download the video or audio from YouTube using the yt-dlp library.**
- **Convert the downloaded file to the WAV format using ffmpeg.**
- **Use the Whisper module for transcription. However, note that only the tiny and small pretrained models are supported due to memory allocation limitations and getting some errors in output.**
- **If attempting to use the large or medium models, i was facing memory allocation errors.**
- **As an alternative,i utilized the SpeechRecognition library, which provides good transcription results.**
- **Deployed the application using Streamlit, allowing users to upload audio files and transcribe them. -The Streamlit app provides a user interface to upload audio files, select the desired language, and initiate the transcription process.**
- **The audio file is processed by splitting it into chunks and transcribing each chunk using the selected speech recognition engine.**
- **The resulting transcriptions are then joined together to form the complete transcription of the audio file.**
- **The transcription is displayed in the Streamlit app for users to view.**

In [46]:

```
pip install insta-scrape
```

```
Collecting insta-scrapeNote: you may need to restart the kernel to use updated packages.

  Using cached insta_scrape-2.1.2-py3-none-any.whl (26 kB)
Requirement already satisfied: beautifulsoup4 in c:\users\lenovo\anaconda3\lib\site-packa
ges (from insta-scrape) (4.11.1)
Requirement already satisfied: requests in c:\users\lenovo\anaconda3\lib\site-packages (f
rom insta-scrape) (2.28.1)
Requirement already satisfied: soupsieve>1.2 in c:\users\lenovo\anaconda3\lib\site-packag
es (from beautifulsoup4->insta-scrape) (2.3.1)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\users\lenovo\anaconda3\lib\
site-packages (from requests->insta-scrape) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\lenovo\anaconda3\lib\site-p
ackages (from requests->insta-scrape) (2022.12.7)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\lenovo\anaconda3\lib\sit
e-packages (from requests->insta-scrape) (1.26.11)
Requirement already satisfied: idna<4,>=2.5 in c:\users\lenovo\anaconda3\lib\site-package
s (from requests->insta-scrape) (3.3)
Installing collected packages: insta-scrape
Successfully installed insta-scrape-2.1.2
```

```
[notice] A new release of pip is available: 23.0.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [51]:

```
from instascrape import Reel
import time
import datetime
import math

# Session id
SESSIONID = "4201753765%3AqhRYt8ozT7aEMW%3A2%3AAYcc9AcLYK1Pl5Il1wM6B-8Y310Dj9fVoR5QFgIG_Q
"

# Header with session id
headers = {
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, l
ike Gecko) Chrome/79.0.3945.74 Safari/537.36 Edg/79.0.309.43",
    "cookie": f'sessionid={SESSIONID};'
}

# Passing Instagram reel link as an argument in the Reel module
insta_reel = Reel('https://www.instagram.com/reel/Ctt205It3xZ/?utm_source=ig_web_copy_lin
k&igshid=MzRlODBiNWFlZA==')

try:
    # Using scrape function and passing the headers
    insta_reel.scrape(headers=headers)

    # Check if the reel has a valid video URL
    if insta_reel.is_video and isinstance(insta_reel.video_url, str) and not math.isnan(f
loat(insta_reel.video_url)):
        # Giving the path where we want to download the reel to
        # the download function
        insta_reel.download(fp=f".\\Desktop\\reel{int(time.time())}.mp4")

        # Printing success message
        print('Downloaded successfully.')
    else:
        print('Failed to retrieve a valid video URL for the reel.')
except Exception as e:
    print(f'Error occurred: {e}')
```

```
Error occurred: Invalid value NaN (not a number)
```

Instagram video transcription

In [55]:

```
import instaloader
loader = instaloader.Instaloader()
url = 'https://www.instagram.com/reel/CrqhSpTuujA/?utm_source=ig_web_copy_link&igshid=MzRlODBiNWFlZA=='
try:
    post = instaloader.Post.from_shortcode(loader.context, url.split("/")[-2])
    loader.download_post(post, target='#target_directory')
    print('Downloaded successfully.')
except Exception as e:
    print(f'Error occurred: {e}')
```

```
#target_directory\2023-04-30_15-34-56_UTC.jpg [Take your chance, baby! ☺  Pr…] #target_directory\2023-04-30_15-34-56_UTC.mp4 json
Downloaded successfully.
```

In [56]:

```
from pydub import AudioSegment
import moviepy.editor as mp

def convert_to_wav(video_path, output_path):
    video = mp.VideoFileClip(video_path)
    audio = video.audio
    audio = audio.to_audiofile(output_path, codec='pcm_s16le', bitrate='16k')

    print("Conversion to WAV completed.")

input_video_path = r"C:\Users\lenovo\ds project\#target_directory\2023-04-30_15-34-56_UTC.mp4"
output_wav_path = r"C:\Users\lenovo\ds project\#target_directory\insta1_output_audio.wav"

convert_to_wav(input_video_path, output_wav_path)
```

```
MoviePy - Writing audio in C:\Users\lenovo\ds project\#target_directory\insta1_output_audio.wav
```

```
MoviePy - Done.
Conversion to WAV completed.
```

In [58]:

```
filename=r"C:\Users\lenovo\ds project\#target_directory\insta1_output_audio.wav"
language = 'hi-IN'
transcription_paragraph = process_audio(filename, language)

print("Transcription:")
print(transcription_paragraph)
```

```
I am exporting ./chunked/insta1_output_audio.wav_0.wav
I am exporting ./chunked/insta1_output_audio.wav_1.wav
I am exporting ./chunked/insta1_output_audio.wav_2.wav
I am exporting ./chunked/insta1_output_audio.wav_3.wav
Transcription:
तुम तो बड़े दोस्तों से गाड़ी खरीदी 5 साल पहले अब तक राज पकड़े गए मुझे जिस दिन हो गए उस दिन में गाड़ी बाहर तो कैसे पता चलेगा कि बाहर निकालो फ्यूचर मेकर
```

In [60]:

```
model_file_path = r"C:\Users\lenovo\Downloads\deepspeech-0.9.3-models.pbmm"
lm_file_path = r"C:\Users\lenovo\Downloads\deepspeech-0.9.3-models.scorer"
```

In [62]:

```
!pip install deepspeech
```

```
[notice] A new release of pip is available: 23.0.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
Collecting deepspeech
  Using cached deepspeech-0.9.3-cp39-cp39-win_amd64.whl (8.0 MB)
```

```
Requirement already satisfied: numpy>=1.19.4 in c:\users\lenovo\anaconda3\lib\site-packag
es (from deepspeech) (1.21.5)
Installing collected packages: deepspeech
Successfully installed deepspeech-0.9.3
```

In [65]:

```
!pip install pydub
```

```
Requirement already satisfied: pydub in c:\users\lenovo\anaconda3\lib\site-packages (0.25
.1)

[notice] A new release of pip is available: 23.0.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [64]:

```
!apt install libasound2-dev portaudio19-dev libportaudio2 libportaudiocpp0 ffmpeg
```

```
'apt' is not recognized as an internal or external command,
operable program or batch file.
```

In [66]:

```
video_url='https://youtu.be/26PrgjTboVQ'
output_path = r"D:\pytube\audio\english_6min.mp3"
download_audio(video_url, output_path)
```

In [68]:

```
input_file = r"D:\pytube\audio\english_6min.mp3"
output_file = r"D:\pytube\audio\english_6min1.wav"

command = ["C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe", '-i', input_
file, output_file]
subprocess.run(command, capture_output=True)
```

Out[68]:

```
CompletedProcess(args=['C:\\Users\\lenovo\\OneDrive\\Desktop\\ffmpeg\\bin\\ffmpeg.exe', '
-i', 'D:\\pytube\\audio\\english_6min.mp3', 'D:\\pytube\\audio\\english_6min1.wav'], retu
rncode=0, stdout=b'', stderr=b"ffmpeg version 6.0-full_build-www.gyan.dev Copyright (c) 2
000-2023 the FFmpeg developers\r\n  built with gcc 12.2.0 (Rev10, Built by MSYS2 project)
\r\n  configuration: --enable-gpl --enable-version3 --enable-static --disable-w32threads
--disable-autodetect --enable-fontconfig --enable-iconv --enable-gnutls --enable-libxml2
--enable-gmp --enable-bzlib --enable-lzma --enable-libsnappy --enable-zlib --enable-libr
ist --enable-libsrt --enable-libssh --enable-libzmq --enable-avisynth --enable-libbluray
--enable-libcaca --enable-sdl2 --enable-libaribb24 --enable-libdav1d --enable-libdavs2 -
-enable-libuavs3d --enable-libzvbi --enable-librav1e --enable-libsvtav1 --enable-libwebp
--enable-libx264 --enable-libx265 --enable-libxavs2 --enable-libxvid --enable-libaom --e
nable-libjxl --enable-libopenjpeg --enable-libvpx --enable-mediafoundation --enable-liba
ss --enable-frei0r --enable-libfreetype --enable-libfribidi --enable-liblensfun --enable
-libvidstab --enable-libvmaf --enable-libzimg --enable-amf --enable-cuda-llvm --enable-c
uvid --enable-ffnvcodec --enable-nvdec --enable-nvenc --enable-d3d11va --enable-dxva2 --
enable-libmfx --enable-libshaderc --enable-vulkan --enable-libplacebo --enable-opencl --
enable-libcdio --enable-libgme --enable-libmodplug --enable-libopenmpt --enable-libopenc
ore-amrwb --enable-libmp3lame --enable-libshine --enable-libtheora --enable-libtwolame --
enable-libvo-amrwbenc --enable-libilbc --enable-libgsm --enable-libopencore-amrnb --enab
le-libopus --enable-libspeex --enable-libvorbis --enable-ladspa --enable-libbs2b --enabl
e-libflite --enable-libmysofa --enable-librubberband --enable-libsoxr --enable-chromaprin
t\r\n  libavutil      58.  2.100 / 58.  2.100\r\n  libavcodec     60.  3.100 / 60.  3.100
\r\n  libavformat    60.  3.100 / 60.  3.100\r\n  libavdevice    60.  1.100 / 60.  1.100\
r\n  libavfilter     9.  3.100 /  9.  3.100\r\n  libswscale      7.  1.100 /  7.  1.100\r
\n  libswresample   4. 10.100 /  4. 10.100\r\n  libpostproc    57.  1.100 / 57.  1.100\r\
nInput #0, mp3, from 'D:\\pytube\\audio\\english_6min.mp3':\r\n  Metadata:\r\n    encoder
: Lavf60.3.100\r\n  Duration: 00:06:19.20, start: 0.023021, bitrate: 82 kb/s\r\n  Stream
#0:0: Audio: mp3, 48000 Hz, stereo, fltp, 82 kb/s\r\n    Metadata:\r\n      encoder
: Lavc60.3.\r\nStream mapping:\r\n  Stream #0:0 -> #0:0 (mp3 (mp3float) -> pcm_s16le (nat
ive))\r\nPress [q] to stop, [?] for help\r\nOutput #0, wav, to 'D:\\pytube\\audio\\englis
h_6min1.wav':\r\n  Metadata:\r\n    ISFT            : Lavf60.3.100\r\n  Stream #0:0: Aud
io: pcm_s16le ([1][0][0][0] / 0x0001), 48000 Hz, stereo, s16, 1536 kb/s\r\n    Metadata:\
r\n      encoder         : Lavc60.3.100 pcm_s16le\r\nsize=       0kB time=00:00:00.00 bit
```

```
rate=N/A speed=N/A    \rsize=   68352kB time=00:06:04.77 bitrate=1535.0kbits/s speed= 176
x    \rsize=   71094kB time=00:06:19.15 bitrate=1536.1kbits/s speed= 171x    \r\nvideo:0k
B audio:71094kB subtitle:0kB other streams:0kB global headers:0kB muxing overhead: 0.0001
07%\r\n")
```

Process of audio transcription project:

- **Download the video or audio from YouTube using the yt-dlp library.**
- **Convert the downloaded file to the WAV format using ffmpeg.**
- **Firstly i started whisper module,due to memory allocation,i could use base and small pretrained whisper model when i make use of medium,large module im getiing memory error so i started use speech recognition model**
- **further,Utilize SpeechRecognition library for transcription as an alternative to the Whisper module.**
- **The Whisper module only supports tiny and small pretrained models due to memory allocation limitations and errors in output.**
- **Memory allocation errors occur when attempting to use the large or medium models with the Whisper module.**
- **Deploy the application using a Flask API, allowing users to upload audio files and transcribe them.**
- **The Flask API provides a user interface for uploading audio files, selecting the language, and initiating the transcription process.**
- **Process the audio file by splitting it into chunks and transcribing each chunk using the selected speech recognition engine.**
- **Joined the resulting transcriptions together to form the complete transcription of the audio file.**
- **Display the transcription in the Flask app for users to view. Additionally, adjusting certain thresholds in the audio file can improve the transcription results. Google Speech-to-Text API is also used and performs better compared to the Whisper module and SpeechRecognition library.**

In [ ]: