

Línea de comandos con docker

Speech to text

<https://github.com/alphacep/vosk-api>

<https://alphacephei.com/vosk/install>

```
apt-get install python-is-python3 python-is-python3 python3-pip ffmpeg  
pip3 install vosk
```

```
vosk-transcriber -l es -i fichero.wav -o texto.txt
```

Dockerfile

```
FROM debian  
  
RUN apt-get update && apt-get install -y python-is-python3 python-is-python3  
python3-pip ffmpeg  
RUN pip3 install vosk  
  
WORKDIR /audio  
  
ENTRYPOINT ["bash", "-c", "vosk-transcriber -l es -i $1 -o texto.txt  
>/dev/null 2>&1 && cat texto.txt", "--"]  
  
docker run --rm -v "$PWD":/audio local/speechtotext fichero.wav
```

Como servicio web

Creamos el fichero vosk.py que levantará un servidor web que al hacer un post de un fichero de audio, devuelve el texto. Lo levantamos con el comando:

```
python3 vosk.py
```

```
from flask import Flask, request, jsonify  
import subprocess  
  
app = Flask(__name__)  
  
@app.route('/transcribir', methods=['POST'])  
def transcribir_audio():  
    archivo_audio = request.files['audio']  
    archivo_audio.save('audio.wav')  
    subprocess.run(['vosk-transcriber', '-l', 'es', '-i', 'audio.wav', '-o',  
                    'texto.txt'])
```

```
with open('texto.txt', 'r') as f:
    texto = f.read()
    return jsonify({'texto': texto})

if __name__ == '__main__':
    app.run(debug=True, host='0.0.0.0', port=3000)
```

Se envía un audio con este comando y te devuelve el texto:

```
curl -X POST -F 'audio=@a.ogg' http://10.103.0.1:3000/transcribir
```

Servicio de Google

```
import speech_recognition as sr
import time

# Creamos un objeto Recognizer
r = sr.Recognizer()

# Abrimos el archivo de audio y lo pasamos al objeto AudioFile
with sr.AudioFile('audio.wav') as source:
    # Leemos el audio del archivo
    audio = r.record(source)

# Convertimos el audio a texto
text = r.recognize_google(audio, language='es-ES')

# Generamos un nombre de archivo único para el texto
filename = 'texto_' + str(int(time.time())) + '.txt'

# Guardamos el texto en un archivo de texto
with open(filename, 'w') as f:
    f.write(text)
```

API TELEGRAM

/dades/web/htdocs/apitelegram.lob99.info/speechtotext/bot.php

Graba el archivo en el file system como AwAxxxxxxxxx.ogg

Lo pasa a <http://vosk.lob99.com/transcribir>

From:

<http://wiki.legido.com/> - **Legido Wiki**

Permanent link:

http://wiki.legido.com/doku.php?id=informatica:speech_to_text



Last update: **2023/04/14 08:31**