OpenAI Platform   Docs API   Log in   Sign up

Speech to text

Learn how to turn audio into text

Introduction

The Audio API provides two speech to text endpoints, `transcriptions` and `translations`, based on our state-of-the-art open source large-v2 Whisper model. They can be used to:

Transcribe audio into whatever language the audio is in.

Translate and transcribe the audio into english.

File uploads are currently limited to 25 MB and the following input file types are supported: `mp3`, `mp4`, `mpeg`, `mpga`, `m4a`, `wav`, and `webm`.

Quickstart

Transcriptions

The transcriptions API takes as input the audio file you want to transcribe and the desired output file format for the transcription of the audio. We currently support multiple input and output file formats.

Transcribe audio   python ⌄   ⧉

```python
1  from openai import OpenAI
2  client = OpenAI()
3
4  audio_file= open("/path/to/file/audio.mp3", "rb")
5  transcription = client.audio.transcriptions.create(
6      model="whisper-1",
7      file=audio_file
8  )
9  print(transcription.text)
```