Documentations for running the two mini projects:

We are going to see two mini projects where we will be using Google Cloud Vision API for extracting the text from the image and audio.

To start with we have to get an API-key of Google-Cloud-Vision in order to use their services.

Step 1:
Setting up Your Google Platform Account

Google Cloud Platform Free An account is required to get an api-key.json file.

- Sign-in to Google Cloud Console
- Click “API Manager”
- Click “Credentials”
- Click “Create Credentials”
- Select “Service Account Key”
- Under “Service Account” select “New service account”
- Name service (whatever you’d like)
- Select Role: “Project” -> “Owner”
- Leave “JSON” option selected
- Click “Create”
- Save generated API key file
- Rename file to api-key.json

Step 2:
Convert the Audio file to .WAV format. We can use any online tools to do it.

Step 3(For Audio to text):
Break up the audio file into smaller parts. Google Cloud Speech API only accepts files no longer than 60 seconds. To be on safe side, either break your files in 30-seconds chunks or select audio file less than 60 seconds.

To Break:
We can either use any online tools or we can use an open source command line library called ffmpeg. It can be downloaded from its site and install it in your machine.

Here is the command to break up the file.

- ➔ #Clean out the old parts if needed via **rm -rf parts/***

Here is the command I used to break up my file:

```
# Clean out old parts if needed via rm -rf parts/*
ffmpeg -i source/filename.wav -f segment -segment_time 30 -c copy parts/out%09d.wav
```

Where, source/ filename.wav is the name of the input file, and parts/out%09d.wav is the format for output files. %09d indicated that the file number will be padded with 9 zeros (i.e. out000000001.wav), allowing files to be sorted alphabetically. This way ls command returns files sorted in the right order.

Step 3(For Image to text):
For Image we don’t need to do much pre-work. We have to select the image and keep them in the local directories or we have to mentioned the proper address if the location of the image is different.

Step 4:

Install the requirements.txt file using pip command which contains the required libraries.

```
pip install -r requirements.txt
```

Step 5:
Running the Code

Step 6:

These two mini projects should gives an amazing result and it does recognize the words properly even from a song which is amazing. Same goes with the image-to-text project, it reads the words properly, but it is not able to format then properly which is something we have to take care.