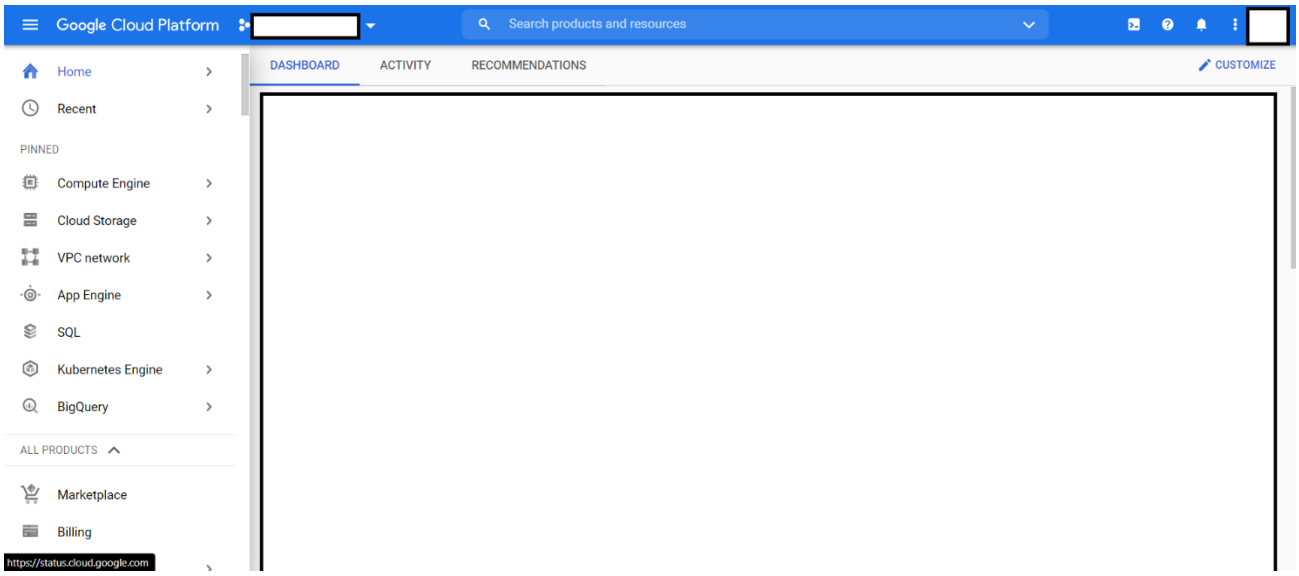
Multimodal Human Computer Interaction

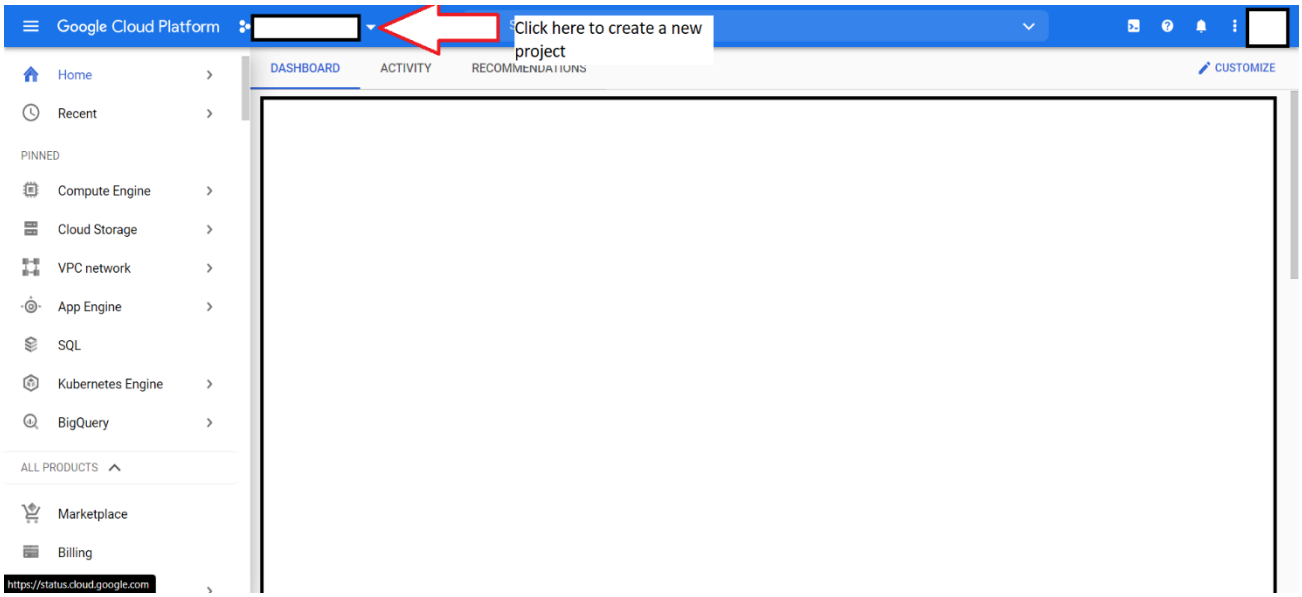
To set up the speech recognition, follow the steps below:

Step 1: Create a Google Cloud Platform Account (https://console.cloud.google.com/). A free trial account will be sufficient.

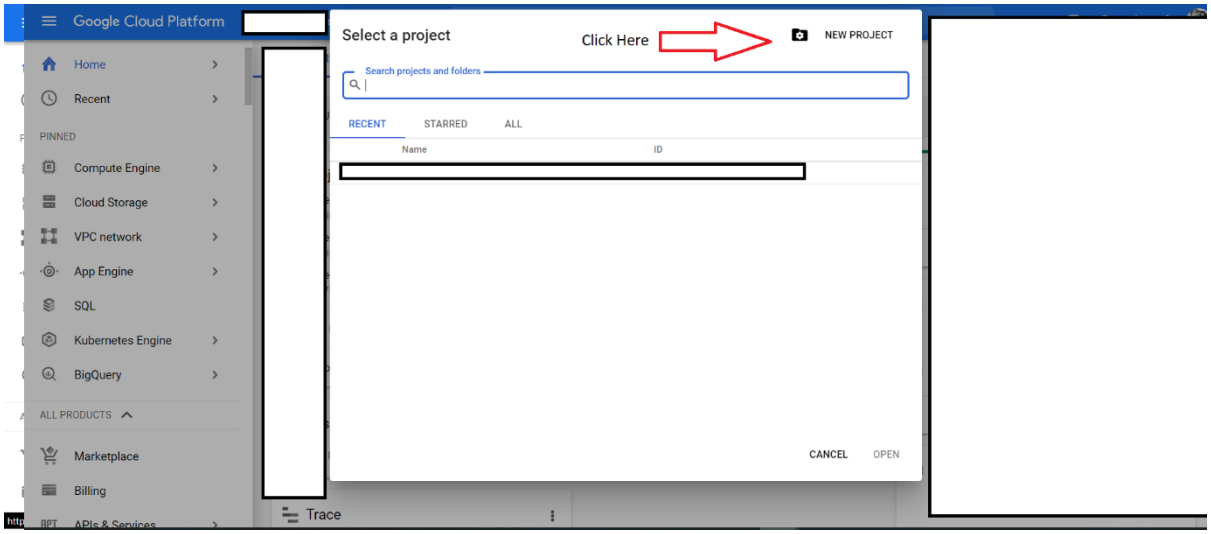
Step 2: After you're done creating the account, you'll land on the following page



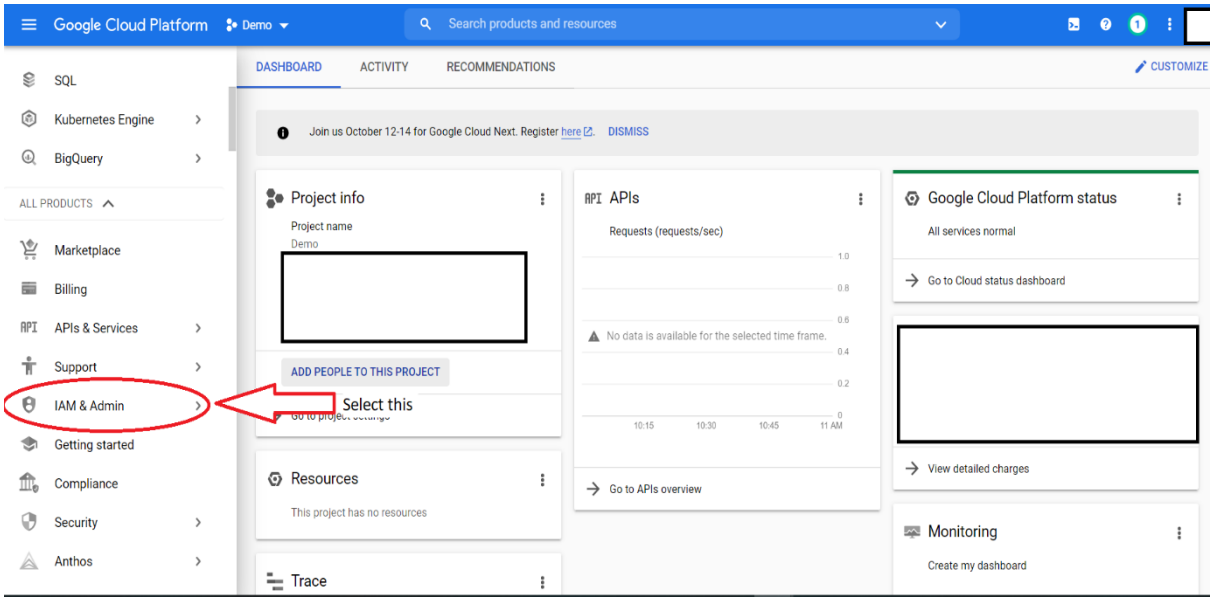
Next create a new project



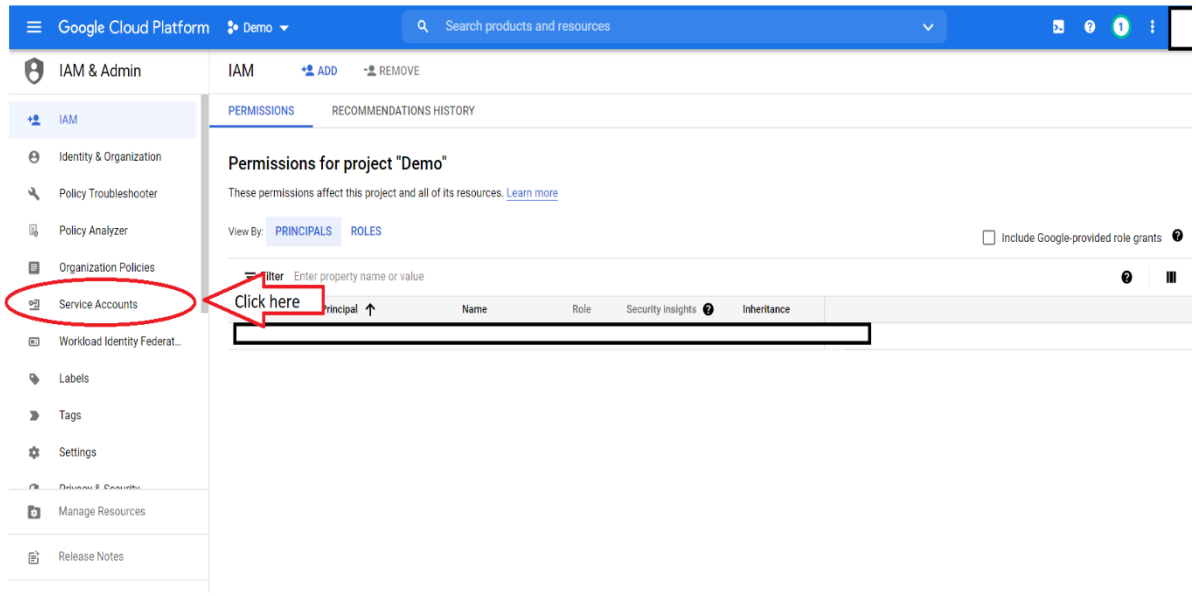
Select the following in the popup. Then continue to creating the project.



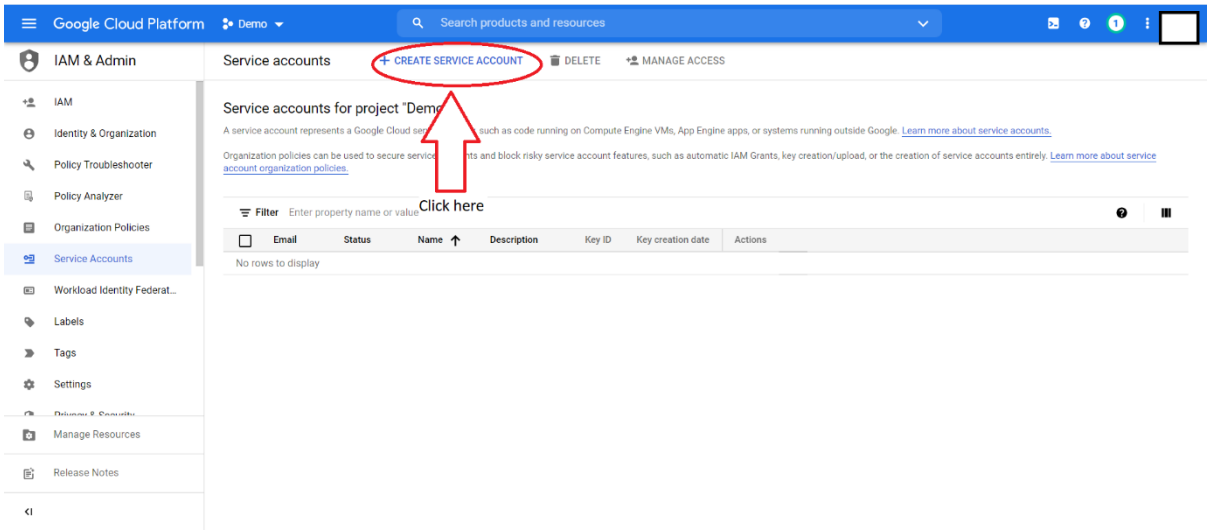
Step 3: Scroll down and select **IAM & Admin** from the side menu



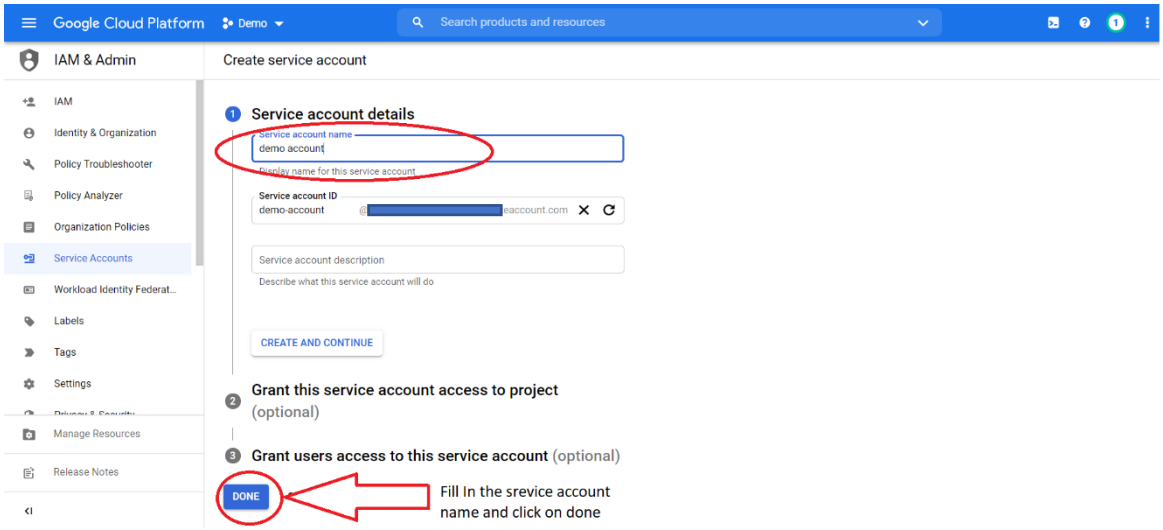
Then the following will appear. Here, click on **Service Accounts** from the side menu.



Then the following will appear. Here, click on **Create Service Account** from the side menu.

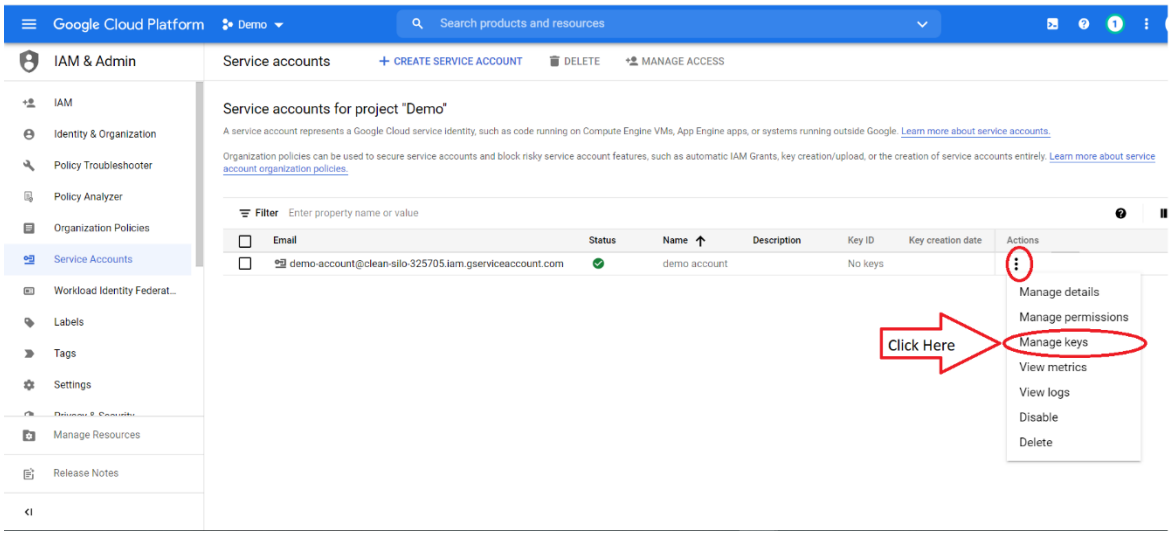


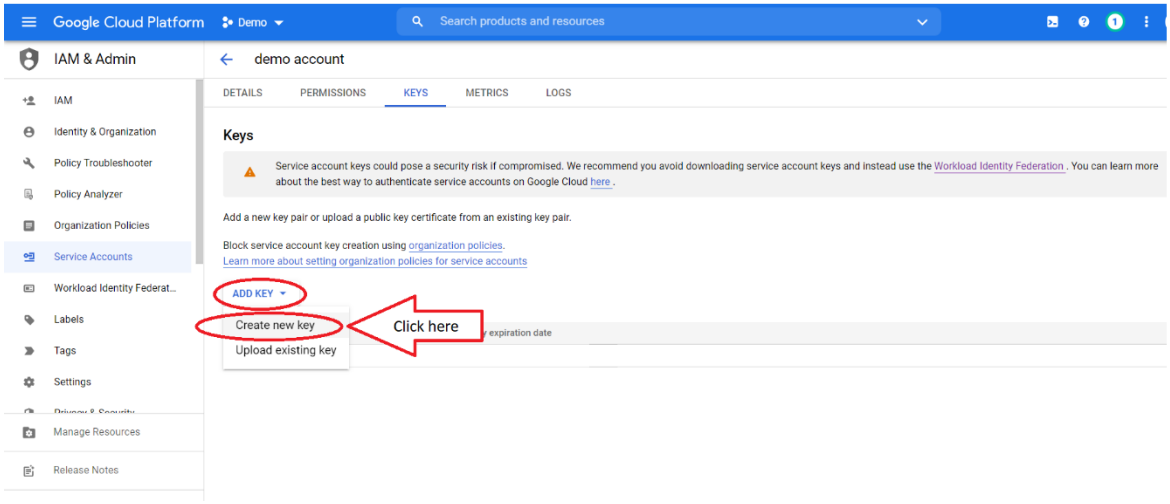
Then the following will appear. Here, fill in the name and click on done.

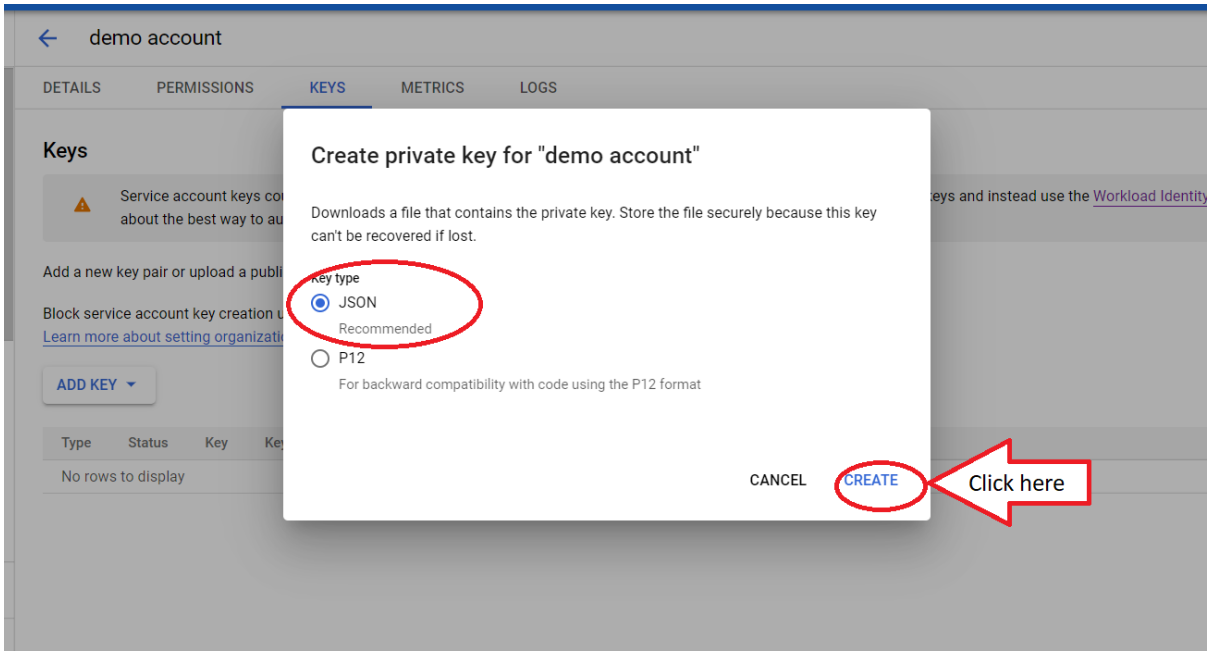


You'll be redirected to the following page, where it shows you your new service account. Next, we'll have to create an API Key for your account.

Note: The API key generated in the following process SHOULD NOT be shared.





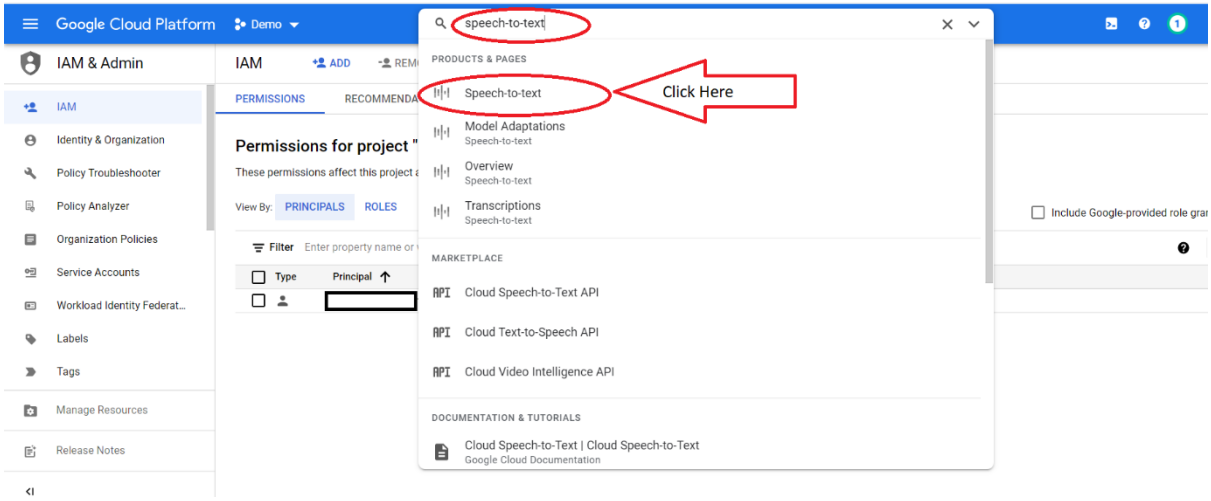


When you click create, a .json file containing the API key is downloaded. Rename this file to " SERVICE_KEY.json ", and place it in the same folder as the main program file.

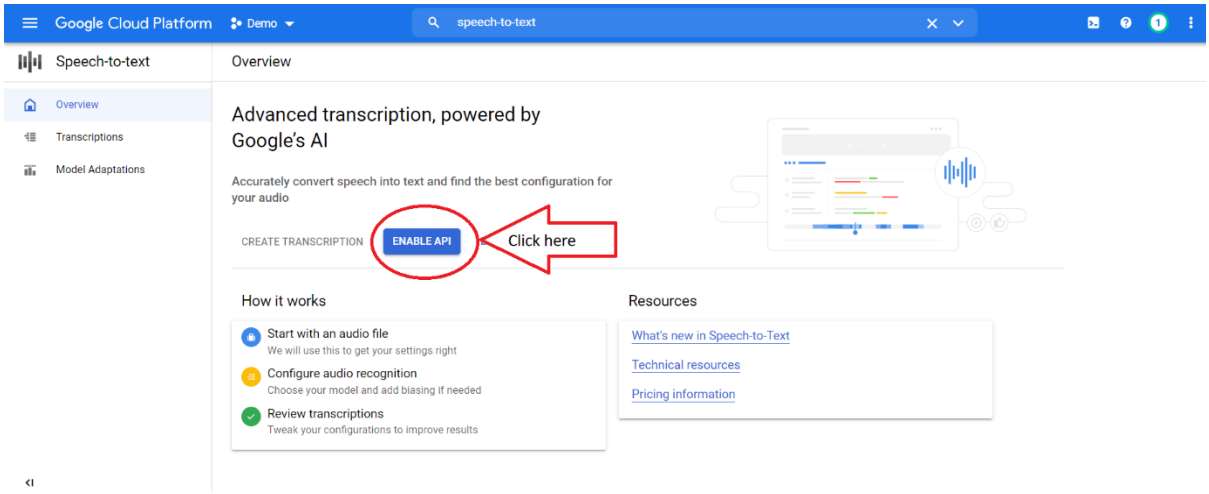
Do not share the .json file anywhere public.

Step 4: Activating the speech-to-text API.

In the above search bar, search for “speech-to-text”.



Then the following will appear. Here, click on “Enable API”.



Your Cloud speech-to-text API will be activated.

Program:

Before you run the program make sure the following Python packages are installed.

cv2

numpy

mediapipe

time

autopy

threading

re

sys

google.cloud

pyaudio

pywin32

keyboard

json

os

Speech Commands Available:

Open, double click

Click, tap, left click

Right click

Copy

Cut, move

Paste

Scroll up

Scroll down

Zoom in

Zoom out

Exit, quit, close, done (these commands are used to stop the program execution)