



AURA

SAY YOUR  
WORDS  
AND THE  
WORLD  
TRANSFORMS



LET US DO OUR  
WORK

## OVERVIEW

Using the chatgpt API, we are creating a speech assistant. The idea is to record our speech and feed it into chatgpt, which will find the answer and convert it back to audio for the user.



A U R A

# THIS IS WHAT IT LOOKS LIKE...

### AI Voice Assistant with ChatGPT AI

user\_voice

Record from microphone

Clear

ChatGPT Text

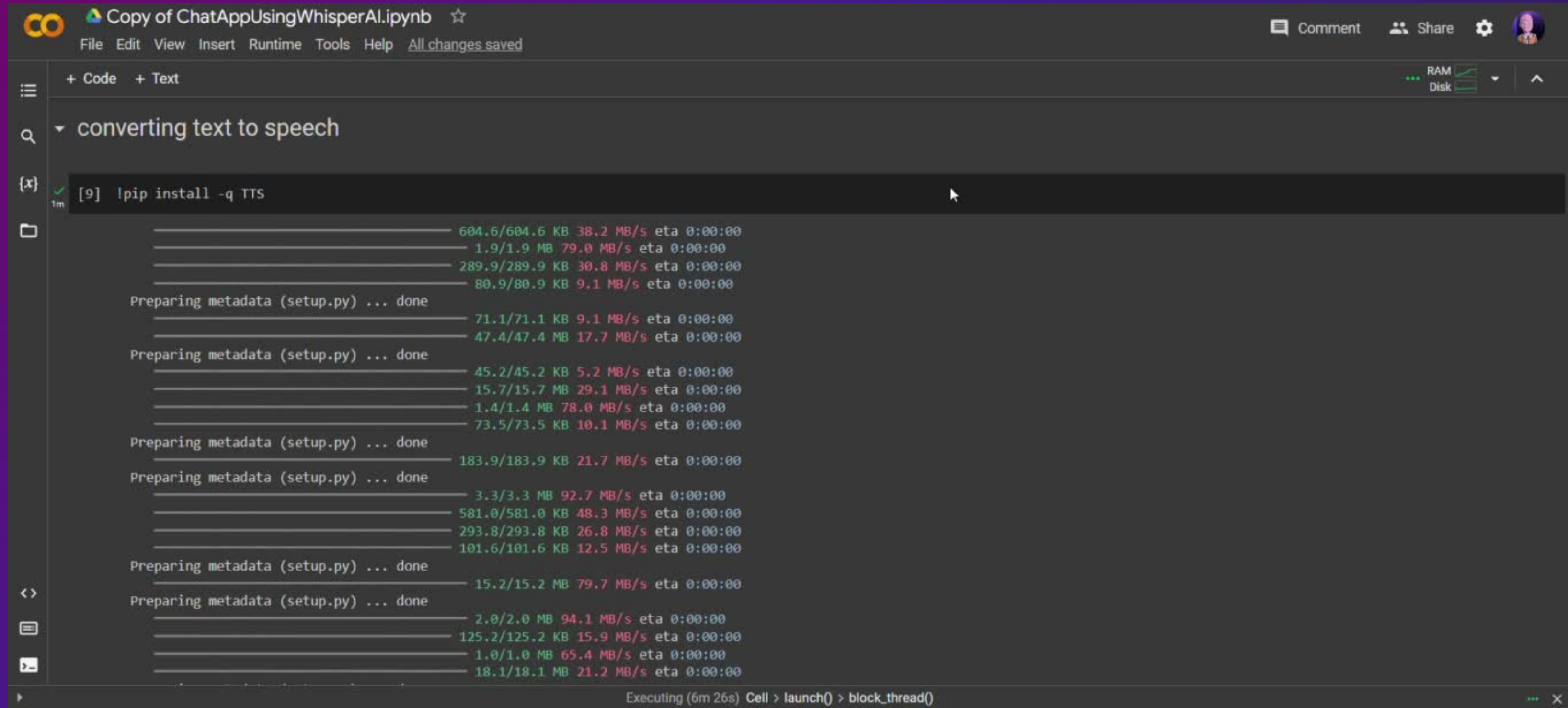
output 1

0:00 / 0:01

Flag

The image shows a dark-themed user interface for an AI voice assistant. At the top, the title "AI Voice Assistant with ChatGPT AI" is centered. Below the title, the interface is split into two main sections. The left section, labeled "user\_voice", contains a "Record from microphone" button with a red indicator light and a "Clear" button below it. The right section, labeled "output 1", features a "ChatGPT Text" input field, a media player with a play button, a progress bar showing "0:00 / 0:01", a volume icon, and a "Flag" button at the bottom.

# THIS IS WHAT IT LOOKS LIKE...



The screenshot shows a Jupyter Notebook interface with a terminal output for a pip install command. The notebook title is "Copy of ChatAppUsingWhisperAI.ipynb". The terminal output shows the command `!pip install -q TTS` and its execution progress, including package names, sizes, download speeds, and estimated times.

```
[9] !pip install -q TTS
----- 604.6/604.6 KB 38.2 MB/s eta 0:00:00
----- 1.9/1.9 MB 79.0 MB/s eta 0:00:00
----- 289.9/289.9 KB 30.8 MB/s eta 0:00:00
----- 80.9/80.9 KB 9.1 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 71.1/71.1 KB 9.1 MB/s eta 0:00:00
----- 47.4/47.4 MB 17.7 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 45.2/45.2 KB 5.2 MB/s eta 0:00:00
----- 15.7/15.7 MB 29.1 MB/s eta 0:00:00
----- 1.4/1.4 MB 78.0 MB/s eta 0:00:00
----- 73.5/73.5 KB 10.1 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 183.9/183.9 KB 21.7 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 3.3/3.3 MB 92.7 MB/s eta 0:00:00
----- 581.0/581.0 KB 48.3 MB/s eta 0:00:00
----- 293.8/293.8 KB 26.8 MB/s eta 0:00:00
----- 101.6/101.6 KB 12.5 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 15.2/15.2 MB 79.7 MB/s eta 0:00:00
Preparing metadata (setup.py) ... done
----- 2.0/2.0 MB 94.1 MB/s eta 0:00:00
----- 125.2/125.2 KB 15.9 MB/s eta 0:00:00
----- 1.0/1.0 MB 65.4 MB/s eta 0:00:00
----- 18.1/18.1 MB 21.2 MB/s eta 0:00:00
```

Executing (6m 26s) Cell > launch() > block\_thread()



# TOOLS USED :



**Whisper:** Whisper is an AI technology that can accurately recognize speech in a multitude of languages, accents, and environments.

**ChatGPT API (Turbo GPT 3,5):** ChatGPT is a specialized variant of the GPT-3 language model designed to generate human-like responses in conversational contexts.





HOW THIS COULD  
TRANSFORM THE  
PERSPECTIVE



- Hands-free and intuitive way of interacting with technology.
- More logical thinking with modern GPT
- Can be integrated with smart home devices for seamless control of appliances and systems.
- Increases accessibility for people with disabilities or mobility impairments.



# CURRENT LIMITATIONS

- **Can't handle os related operations**
- **Dont have personal opinions**
- **Accuracy Problems of language model**







# EXTENDING OUR PROJECT

- **including mood prediction by passing the text through some pre-trained models**
- **Creating live Lyrics for songs**
- **Language translation**
- **Live Video Captions**





THANK YOU!