DS565 - Generative AI-Driven Intelligent Apps Development

Project: Speech to text-text to speech(OpenAI)

Course: DS565
Professor: Dr. Henry Chang

Presented by: Karan Shrestha (20087)

Introduction to OpenAI API

- **What is OpenAI API?**
  - Provides access to powerful language models (GPT series)
  - Applications: Text generation, chatbots, translation, etc.
- **Why GPT-4o Mini?**
  - Efficient, lightweight version of GPT-4, optimized for quick interactions

Flask-based web application for real-time voice-driven AI interaction

Features:

- **Audio transcription using Whisper**
- **Response generation with GPT-3.5 Turbo**
- **Text-to-Speech feedback using openaiTTS-1**

Core Technologies

- **Flask**: Backend framework
- **Whisper**: Audio transcription model
- **OpenAI GPT-3.5 Turbo**: AI response generation
- **Open AI tts-1**: Text-to-Speech conversion
- **Werkzeug**: Secure file handling
- **Dotenv**: API key management

System Architecture

1. **Audio Upload**: Users upload their audio files via the web interface.
2. **Transcription**: Audio is transcribed using Whisper.
3. **Response Generation**: GPT-3.5 generates an intelligent response.
4. **TTS Feedback**: The response is converted into speech and returned to the user.

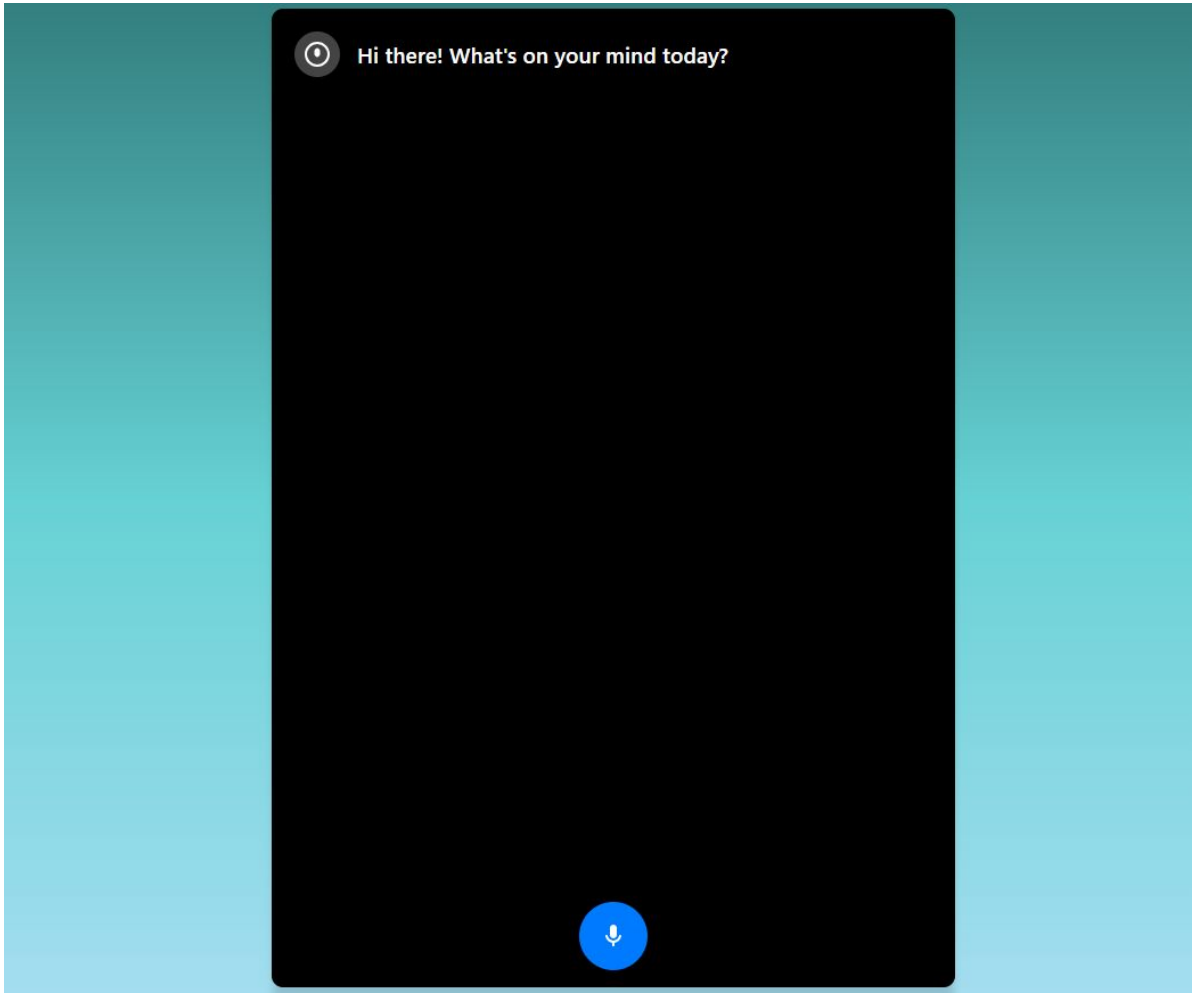
Highlights

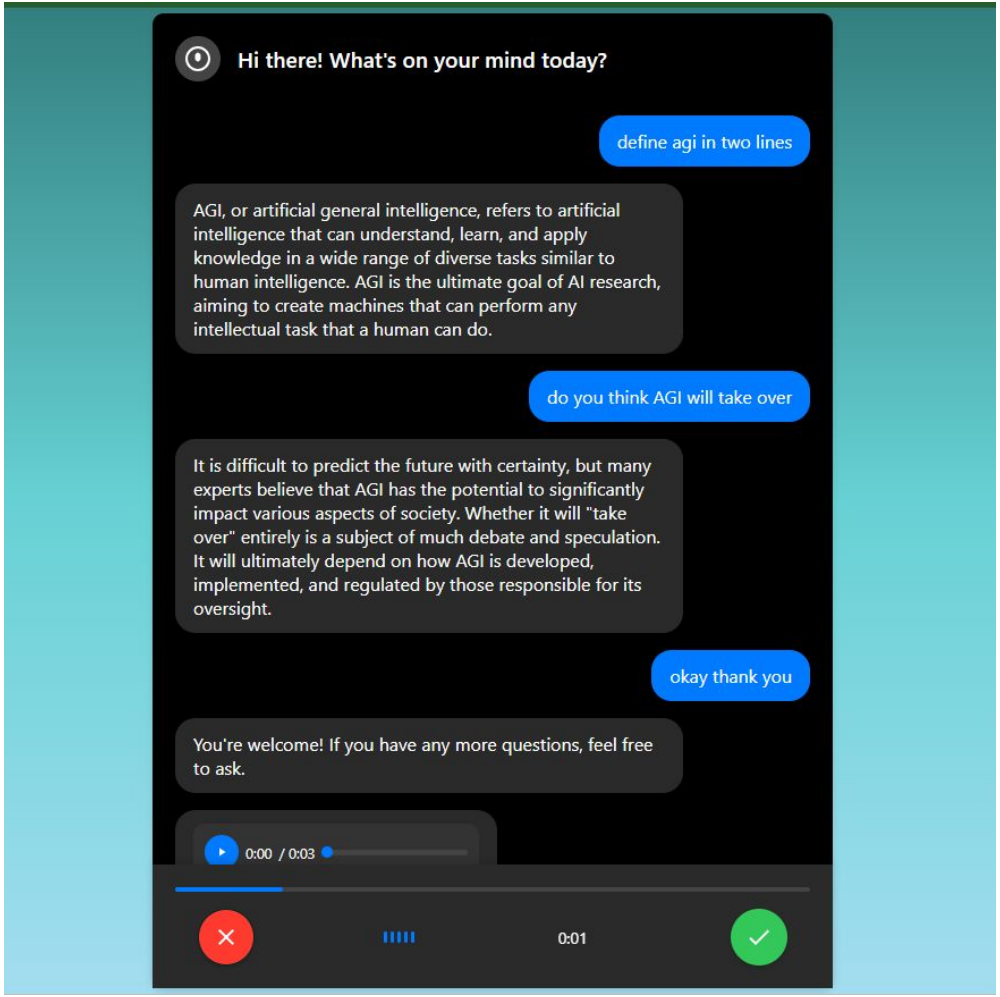
- **Real-Time Interaction**: Immediate transcription and response
- **Seamless Audio Integration**: Easy upload and playback
- **Dynamic TTS Feedback**: Unique response generation for every interaction
- **Scalable Design**: Modular architecture for future improvements

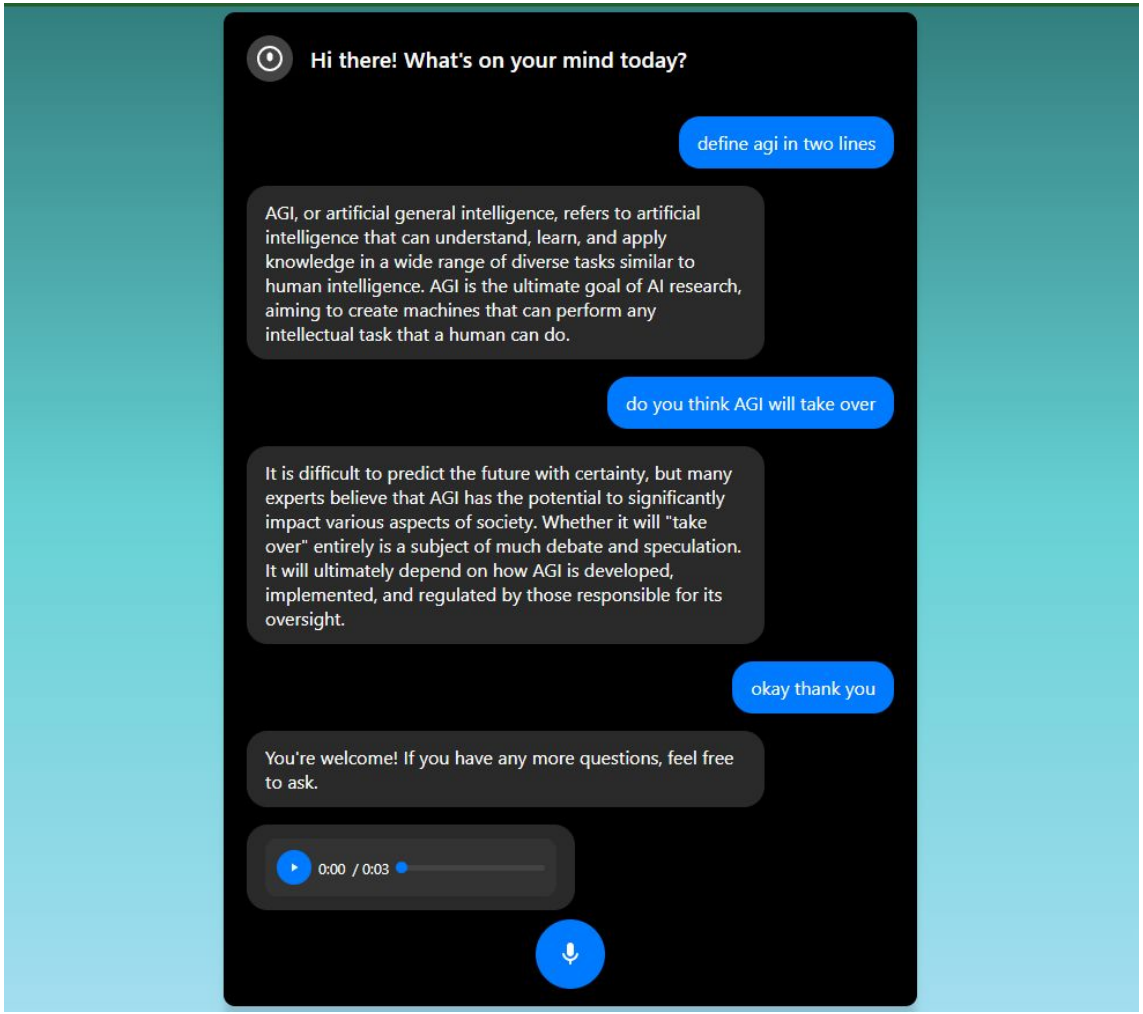
```python
def generate_tts(text, output_dir="static/audio"):
    global previous_audio_file
    os.makedirs(output_dir, exist_ok=True)

    # Generate a unique filename for the audio file
    unique_filename = f"response_{uuid.uuid4().hex}.mp3"
    # unique_filename = f"response.mp3"
    filepath = os.path.join(output_dir, unique_filename)


    with client.audio.speech.with_streaming_response.create(
        model="tts-1",
        voice="onyx",
        input=text,
    ) as response:
        response.stream_to_file(filepath)

    return filepath  # Return the full path of the audio file
```

```python
def transcribe_audio(audio_file):
    try:
        # Resolve the absolute path to the `temp` folder
        base_dir = os.path.dirname(os.path.abspath(__file__))  # Get directory of `utility.py`
        temp_dir = os.path.join(base_dir, "temp")  # Resolve the temp directory path

        # Ensure the temp directory exists
        os.makedirs(temp_dir, exist_ok=True)

        # Save the audio file to the `temp` directory
        filename = secure_filename(audio_file.filename)
        temp_path = os.path.join(temp_dir, filename)
        audio_file.save(temp_path)

        audio_file= open(temp_path, "rb")
        transcription = client.audio.transcriptions.create(
        model="whisper-1",
        file=audio_file
        )
        print(transcription.text)
        return transcription.text

    except Exception as e:
        return f"Error during transcription: {str(e)}"
```

Amazon Q Tip 2/3: Invoke suggestions


Hi there! What's on your mind today?


Hi there! What's on your mind today?
define agi in two lines
AGI, or artificial general intelligence, refers to artificial intelligence that can understand, learn, and apply knowledge in a wide range of diverse tasks similar to human intelligence. AGI is the ultimate goal of AI research, aiming to create machines that can perform any intellectual task that a human can do.
do you think AGI will take over
It is difficult to predict the future with certainty, but many experts believe that AGI has the potential to significantly impact various aspects of society. Whether it will "take over" entirely is a subject of much debate and speculation. It will ultimately depend on how AGI is developed, implemented, and regulated by those responsible for its oversight.
okay thank you
You're welcome! If you have any more questions, feel free to ask.
0:00 / 0:03
0:01


Hi there! What's on your mind today?
define agi in two lines
AGI, or artificial general intelligence, refers to artificial intelligence that can understand, learn, and apply knowledge in a wide range of diverse tasks similar to human intelligence. AGI is the ultimate goal of AI research, aiming to create machines that can perform any intellectual task that a human can do.
do you think AGI will take over
It is difficult to predict the future with certainty, but many experts believe that AGI has the potential to significantly impact various aspects of society. Whether it will "take over" entirely is a subject of much debate and speculation. It will ultimately depend on how AGI is developed, implemented, and regulated by those responsible for its oversight.
okay thank you
You're welcome! If you have any more questions, feel free to ask.
0:00 / 0:03