■ Documentation: Meeting Notes & Action Item Extractor

1. Overview

This application is a Streamlit-based tool that converts meeting audio recordings into structured outputs. It provides: a full transcript of the conversation, a concise summary of meeting notes, and a list of action items and tasks identified from the discussion. The app is designed to help teams save time by quickly capturing the essence of meetings and identifying responsibilities.

2. Features

- Upload meeting audio files in formats such as MP3, WAV, and M4A
- Automatic speech-to-text transcription using Whisper (runs locally, no external API required)
- Summarization of the transcript into clear meeting notes using a Hugging Face transformer model
- Simple action item detection by identifying sentences with task-related language
- User-friendly interface built with Streamlit

3. Requirements

- Python 3.8 or higher
- Streamlit
- Whisper (speech-to-text)
- Hugging Face Transformers (summarization)
- PyTorch (required for Whisper)

4. How It Works

- Upload Audio: Users upload a meeting recording through the Streamlit interface.
- Transcription: The audio is transcribed into text using the Whisper model.
- Summarization: The transcript is condensed into concise meeting notes using a summarization model.
- Action Item Extraction: Sentences indicating tasks or responsibilities are extracted and displayed.
- Output Display: The transcript, summarized notes, and action items are shown clearly in the app.

5. Usage Instructions

- Launch the app by running it through Streamlit.
- Upload an audio recording of your meeting.
- Wait for the transcription and summarization process to complete.
- View the full transcript, summarized meeting notes, and extracted action items.

6. Example Workflow

A user uploads a 30-minute meeting recording. The app generates a full transcript of the discussion, summarizes it into key points, and extracts action items such as 'John will prepare the financial report' or 'We need to schedule the next review'.

7. Limitations

- Action item detection is based on simple keyword rules and may miss tasks expressed in indirect language.
- Long audio files may take time to process.
- Whisper requires sufficient compute power to transcribe large recordings efficiently.

8. Future Improvements

- Use advanced NLP models (such as GPT) for more intelligent task extraction.
- Add speaker identification (who said what).
- Enable export of transcripts and notes into Word, PDF, or Notion.
- Support for multiple languages.