Meeting Notes and Action Item Extractor - Documentation

# 1. Project Overview

The Meeting Notes and Action Item Extractor is a tool that:  
- Converts meeting audio recordings into text using local Whisper (speech-to-text model).  
- Generates a summary of the transcript.  
- Extracts action items (tasks, follow-ups, responsibilities).  
- Provides a web frontend for uploading audio and viewing results.  
  
Unlike many AI apps, this project does not rely on OpenAI APIs. Instead, it uses local open-source models.

# 2. Tech Stack

## Backend

- FastAPI → REST API for handling uploads & processing.  
- Whisper (open-source) → For transcription.  
- Python modules → Summarizer, extractor, file handling.  
- FFmpeg → Required for audio preprocessing.

## Frontend

- Streamlit → Simple UI to upload audio and view structured notes.

# 3. Setup & Installation

Clone the Project:  
git clone <your-repo-url>  
cd Meeting-Notes-and-Action-Item-Extractor  
  
Install Dependencies:  
pip install -r requirements.txt  
  
Key dependencies:  
- fastapi  
- uvicorn  
- transformers  
- torch  
- streamlit  
- ffmpeg-python  
  
Install FFmpeg (Windows):  
1. Download from: https://www.gyan.dev/ffmpeg/builds/   
2. Extract (e.g., C:\ffmpeg).  
3. Ensure ffmpeg.exe, ffplay.exe, ffprobe.exe are inside C:\ffmpeg\bin\.  
4. Add C:\ffmpeg\bin to your system PATH.  
  
Test with: ffmpeg -version

# 4. How the App Works

1. Upload Audio File (frontend → backend).  
2. Backend saves file in /uploads.  
3. Whisper transcribes speech → text.  
4. Summarizer generates short structured notes.  
5. Extractor identifies action items.  
6. Frontend displays results (Transcript, Summary, Action Items).

Example Flow:  
Input Audio (meeting speech) →  
Transcript: "Schedule a meeting with the design team on Friday." →  
Summary: "Meeting planned with design team." →  
Action Items:  
- Schedule a meeting with the design team on Friday.

# 5. Issues Encountered & Fixes

* Issue: Error: Could not import module 'main'  
  Root Cause: Running uvicorn main:app outside correct folder  
  Fix: Run from project root → uvicorn main:app --reload
* Issue: Error processing audio  
  Root Cause: FFmpeg not installed / not in PATH  
  Fix: Installed FFmpeg and added to PATH
* Issue: Didn’t see ffmpeg.exe  
  Root Cause: Extracted wrong folder  
  Fix: Checked inside bin/ → confirmed ffmpeg.exe exists
* Issue: Action items not showing  
  Root Cause: Extractor was too strict, no task-like language detected  
  Fix: Updated extractor to look for keywords (schedule, send, prepare…) or use spaCy
* Issue: Summary looks like transcript  
  Root Cause: Summarizer too weak on short inputs  
  Fix: Tuned model / fallback to rule-based sentence shortening
* Issue: Frontend didn’t load results  
  Root Cause: Wrong API endpoint or CORS issue  
  Fix: Confirmed backend running at 127.0.0.1:8000, frontend calls /process-audio/

# 6. Future Improvements

- Better Action Item Extraction → Use LLMs locally (LLaMA, Mistral) instead of keyword rules.  
- Speaker Diarization → Identify who said what.  
- Database Storage → Save transcripts & tasks for later use.  
- Integration with Jira/Trello → Auto-create tasks.  
- Multi-language Support → Whisper handles multiple languages, extend extractor too.