Problem Statement

In modern organizations, meetings often contain critical discussions, decisions, and action items. However, manually transcribing, summarizing, and extracting tasks from these meetings is:

- **Time-consuming** (listening and writing notes takes hours)

- **Error-prone** (important points may be missed)

- **Inefficient** (team members may not review full recordings)

Need:
A system that automatically processes meeting audio and produces:

1. A **transcript** (raw conversation text)

2. A **concise summary** (key points & decisions)

3. **Action items** (structured checklist for follow-up)

Solution Architecture

The solution is a **Streamlit-based AI-powered Meeting Notes Extractor** with the following flow:

1. **User Upload**

   - User uploads audio files (`.wav`, `.mp3`, `.m4a`, `.mp4`, `.mpeg`).

2. **Speech-to-Text (ASR)**

   - Model: **openai/whisper-tiny**

   - Converts audio into text transcript.

3. **Summarization**

   - Model: **sshleifer/distilbart-cnn-12-6**

- Creates a short, readable summary of the transcript.

4. **Action Item Extraction**

   - Model: **google/flan-t5-base**

   - Generates action items in a checklist format.

5. **User Interface (Streamlit)**

   - Displays transcript, summary, and action items.

   - Options to enable/disable AI processing and show detailed outputs.

📊 Architectural Diagram

Here's a simplified architecture flow:

```
┌─────────────────┐
│   User Uploads  │
│   Meeting Audio │
│ (wav, mp3, m4a...) │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Speech-to-Text  │
│ (Whisper Tiny ASR) │
└─────────────────┘
         │ Transcript
         ▼
┌─────────────────┐
│  Summarization  │
│ DistilBART CNN  │
└─────────────────┘
         │ Summary
         ▼
┌─────────────────┐
│ Action Extraction │
│   FLAN-T5 Base  │
└─────────────────┘
         │ Action Items
         ▼
┌─────────────────┐
│   Streamlit UI  │
│  - Transcript   │
│  - Summary      │
│  - Action Items │
└─────────────────┘
```