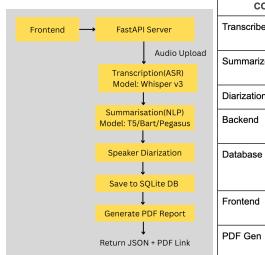
Multilingual Meeting Note-Taking Agent – Design Document

Overview

This system automates the process of converting multilingual meeting audio into structured summaries, action items, and decisions, using advanced speech-to-text and NLP techniques. It is built as a full-stack web application with a Python backend and React frontend.

System Architecture



COMPONENT	TECHNOLOGY	DESCRIPTION
Transcriber	faster-whisper	Efficient transcription using Whisper Large V3
Summarizer	transformers	Extractive & abstractive summarization
Diarization	pyannote-audio	Identifies speaker segments
Backend	FastAPI + Flask	API endpoints, PDF generation, DB interaction
Database	SQLite	Stores transcripts, summaries, meeting metadata
Frontend	React + Axios	UI for upload, summary, search, PDF download
PDF Gen	fpdf	Structured downloadable meeting reports

Future Enhancements

- Advanced NLP-based action & decision extraction
- Role assignment via email or dashboard
- Real-time transcription
- Language detection and dynamic model selection