

**Department of Computer Science
Amrita School of Computing, Amritapuri Campus
B Tech CSE (2022 Admission)**

19CSE495 PROJECT PHASE-1

PROJECT DESIGN SUBMISSION

Group Number: D7

Team Members:

Roll No	Student Name
AM.EN.U4CSE22042	Pirajesh M R
AM.EN.U4CSE22346	Rahul Rajesh Kumar
AM.EN.U4CSE22367	Nikhil Kumar Mishra

Project Guide : Dr. Swaminathan J.

Project Title - RoME: Role-aware Multimodal Meeting Summarizer with Temporal Reasoning

Novelty: This project pioneers personalized, role-aware meeting summarization by fusing audio, video, and transcript data to deliver highlights tailored to each stakeholder's needs. Unlike standard summarizers, RoME tracks topics, decisions, and action items across multiple meetings and uses user feedback for continuous adaptation, introducing an enterprise-relevant and context-rich thread of meeting intelligence.

Objectives:

Phase I (S7) Objectives:

- Develop multimodal data ingestion and synchronization pipeline (audio, video, transcript, metadata).
- Extract role information and assign embeddings to speakers.
- Implement highlight candidate segmentation and scoring using AI models across modalities.
- Build a prototype that generates role-specific summary highlights for a single meeting.
-

Phase II (S8) Objectives:

- Extend summarizer to link topics, decisions, and action items across multiple meetings using temporal graph memory.
- Integrate user feedback (thumbs up/down) for online personalization.
- Expand export and visualization features (thread view, timeline, downloadable reports).
- Test scalability with diverse meeting types and larger datasets.

Novelty - Objective mapping :

The system's ability to generate role-aware summaries (Objective S7-3, S7-4) directly implements the novel idea of tailored meeting intelligence, while the use of multimodal fusion (S7-1, S7-3) sets it apart from vanilla text-only approaches. The extension to track context across meetings (S8-1) and adapt based on feedback (S8-2) addresses the gap of continuity and personalization, ensuring RoME's solution strongly embodies its core innovative claims.

Scope of the project (S7):

Inclusions:

- Multimodal data processing (audio, video, transcript)
- Role-aware segment scoring and highlight generation for meetings
- One-meeting prototype with multiple user roles
- Simple web-based UI for results review

Exclusions:

- Real-time/live streaming summarization (deferred to future work)
- Non-English meeting processing
- Advanced UI features (mobile app, external integrations)

Deliverables:

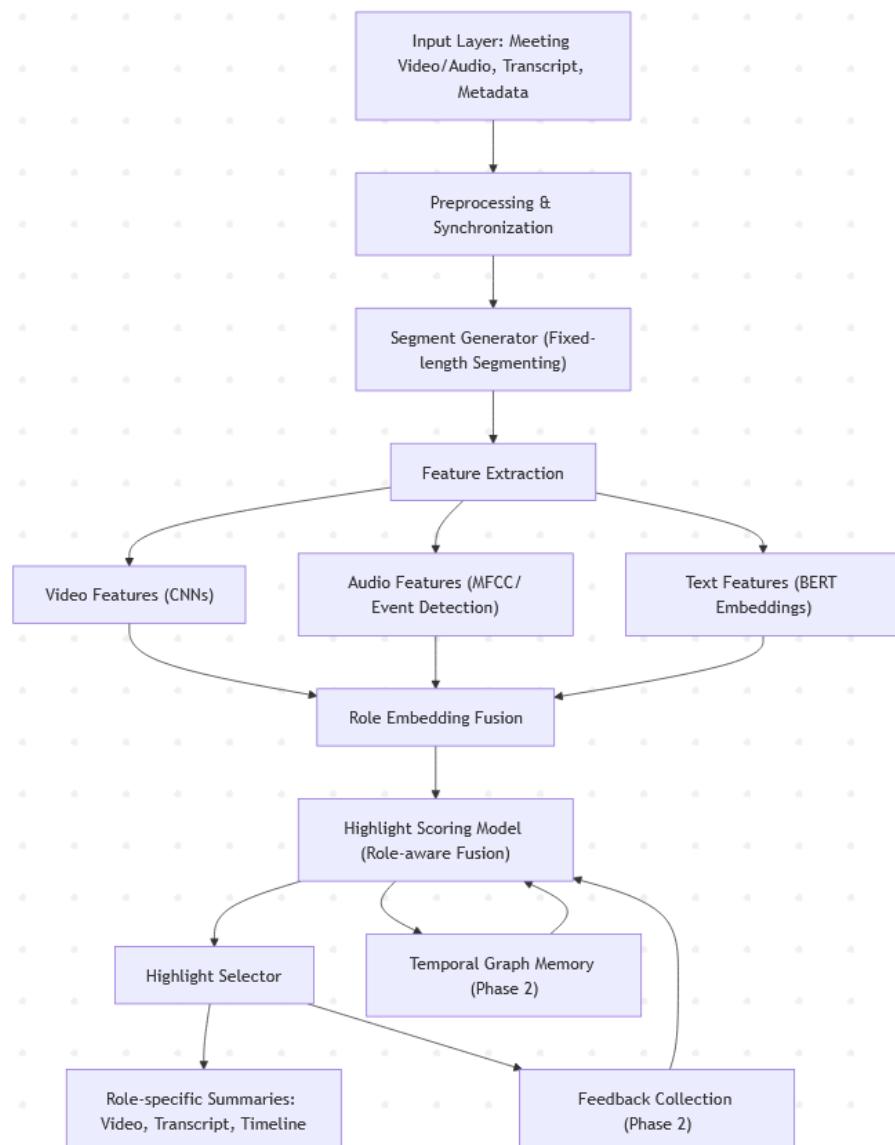
- Working prototype with highlight export
- Documentation and codebase
- Demo video showing system on sample meeting(s)

Boundaries & Limitations:

- Accuracy and personalization limited by available datasets
- Technical constraints on processing speed and scalability
- Limited to offline/batch summarization for project scope

Expected Outcomes:

- Demonstration of system producing role-aware highlights from real/simulated meetings
- Validation of approach using user studies / project presentations
- Foundational pipeline that can be extended in the next phase

Project Design :

Design- Objective mapping :

- Data Ingestion/Preprocessing: Meets S7-1 by preparing synchronized multimodal data and metadata
- Role Embedding and Multi-Modal Fusion: S7-2, S7-3 (role-aware scoring and segment creation)
- Highlight Scoring/Selection and Output: S7-4 (prototype of summary export)
- Temporal Reasoning/User Feedback Modules: Tie in with S8 objectives for continuity and personalization extensions.

Signature of the guide: