

## Project Pitch: Temporal Multimodal RAG for Lifelong Contextual Understanding

### Problem

Modern Retrieval-Augmented Generation (RAG) systems excel at processing static text, but they **lack context over time**, **lose multimodal integration**, and **struggle to link events across different types of data (text, image, audio, video)**. In critical domains like **medicine**, **law**, or **security**, this limits their ability to reason causally, temporally, or contextually across cases.

### Solution

We're developing a **Temporal-Aware Multimodal RAG system** — an AI memory architecture that models **real-time evolving knowledge** much like a human brain:

“Every memory is a node; every node connects causally and temporally to others — across all senses.”

Our system supports:

- **Text, Image, Audio, and Video embeddings**
- **Scene-level segmentation for video and audio**
- **Causal-temporal linkage** between memory nodes
- **OLAP-enhanced SQL** backend for flexible and explainable reasoning
- **Unlearning capability**, allowing dynamic memory rewriting without full retraining