Knowledge Management System with AI and Real-Time Collaboration

Overview

This project is a **Django-based knowledge management system** with AI-powered features, real-time collaboration, and multi-user support. It includes YouTube video summarization, AI-assisted notes, a markdown editor, multimedia storage, classrooms for discussions, and an interactive whiteboard.

Features

1. AI-Powered Summarization

- **YouTube Video Summarization**: Extracts transcripts, processes them with an LLM, and generates structured summaries.
- **AI Notes Summarization**: Summarizes user-generated notes for quick reviews.

2. Note-Taking & Multimedia Support

- **Markdown Editor** for structured note-taking.
- **Support for Images & Audio** within notes.
- **Collaborative Editing**: Anyone in a classroom can edit shared notes.

3. Classroom System

- **Public Classrooms**: Open for anyone to join.
- **Private Classrooms**: Invitation-only access.
- **Role-Based Moderation**: Moderators can manage discussions, delete messages, and pin notes.

4. Real-Time Communication & Collaboration

- **Django Channels** for chat and discussion within classrooms.

- **Interactive Whiteboard**:
    - Uses **Canvas API** for real-time drawing.
    - **WebRTC for Peer-to-Peer (P2P) communication** (direct updates without server load).
    - **Django Channels + WebSockets as a fallback** when P2P isn't possible.

5. Search & Organization

- **Full-Text Search** (PostgreSQL full-text search or Elasticsearch).
- **Tagging & Categorization** for easy note retrieval.

Tech Stack

- **Backend**: Django + DRF
- **Frontend**: HTMX for UI + JavaScript (for real-time whiteboard)
- **Database**: PostgreSQL (supports full-text search)
- **Authentication**: Django Allauth or JWT-based auth
- **Deployment**: Docker, Nginx, Gunicorn on a cloud platform (AWS/GCP)

Implementation Details

- **WebRTC for Whiteboard**:
    - Uses **Django Channels for signaling**.
    - **STUN/TURN servers for P2P connection**.
    - **Canvas API for drawing**, updating in real-time without hitting the server.
- **OCR for Board Text Detection**:
    - Uses **OpenCV to detect board changes**.
    - **Tesseract OCR extracts text** only when needed.
    - **FAISS for text-transcript alignment** instead of brute-force LLM matching.

Summary

This project is a **full-fledged knowledge management system** designed for seamless learning, discussion, and AI-powered assistance. It effectively utilizes Django's capabilities for CRUD, real-time communication, authentication, and AI integrations while maintaining performance and scalability. The combination of **HTMX for UI, WebRTC for**

P2P updates, Django Channels for chat, and AI-driven features makes it a production-ready, innovative solution.