Agentic RAG System

Objective

Build an Agentic RAG System that uses AI agents to intelligently retrieve and answer questions from documents.

Requirements

Your AI agent-powered system must:

1. Ingest data from the specified data sources
2. Store information in a vector database
3. Use agentic workflows to answer user queries

---

1. Agentic Workflow

- Demonstrate agentic behaviour in query processing

---

2. RAG Implementation

- Use any Vector Database
- Data Engineering is critical - show your thinking on data processing

---

3. LLM Integration

- Use any LLM of your choice
- Free APIs are acceptable (e.g., Google Gemini)

---

4. Data Source

System should support the following:

- Documents (PDF, DOCX, PPT)
- Excel
- Text File

Bonus Points

- **Milvus** as vector database
- Advanced retrieval techniques
- Clean UI
- Comprehensive error handling
- **Self-Hosted LLM** with any framework like Ollama or vLLM
- **MCP (Model Context Protocol) Servers** - Build or use MCP servers

---

Submissions

1. GitHub Repository (Public)

Must include:

- Complete working code
- README with setup instructions
- Sample data (if applicable)
- Requirements file

---

2. System Design Document (PDF)

Must explain:

- System architecture diagram
- Agentic workflow design
- Context construction strategy
- Technology choices and rationale
- Key design decisions
- Limitations

---

3. Video Recording (~3 minutes)

Must demonstrate:

- System flow and functionality

Submission Instructions

Submit the following through email:

- GitHub repository URL
- Video recording link
- System Design Document (PDF)

Good luck! Show us how you think about building intelligent AI systems.