Sr. LLM / Backend Engineer

3 Hour Assignment

Scenario: Healthcare Knowledge Assistant

You're building a backend for a RAG-powered assistant that helps clinicians retrieve medical guidelines and research summaries in both English and Japanese. The system must ingest documents in either language and respond to queries like "What are the latest recommendations for Type 2 diabetes management?" in either language.

Tasks

Section	Task Description
Setup	Initialize a FastAPI project with FAISS and sentence-transformers.
Ingestion Endpoint	/ingest: Accept .txt documents in English or Japanese. Detect language, generate embeddings, and store in FAISS.
Retrieval Endpoint	/retrieve: Accept a query in English or Japanese. Return top-3 relevant documents with similarity scores.
Generation Endpoint	/generate: Combine retrieved docs + query into a mock LLM response. Support bilingual output (same as query language).
Translation Feature	Add optional translation toggle (e.g., output_language: "en" or "ja"). Use a mock or open-source translation tool (e.g., transformers or googletrans).
Security	Implement API key authentication (X-API-Key header).
Deployment	Implement an automatic CI/CD pipeline with Github actions and Docker
Design Notes	Write 1-2 paragraphs on scalability, modularity, and future improvements.

Deliverables

- Source code (modular, clean)
- README with setup + design notes
- Github Actions & Dockerfile

Scoring Criteria

Area	Weight
Code structure & clarity	15%
API functionality	15%
FAISS & embedding integration	15%
Bilingual support & translation	15%
Deployment	10%
Design explanation	10%
Security Implementation	10%
Documentation & packaging	10%

AI Usage Disclaimer

- Use of AI tools (e.g., Copilot, ChatGPT, Gemini, etc.) is strictly prohibited unless explicitly stated.
- If you use any AI-generated content (e.g., for mock responses or code snippets), you must cite it clearly in your submission.
- We use AI detection tools to verify originality. Submissions found to be AI-generated without disclosure may be disqualified.

Submission guideline: All tasks, code files, and written responses must be compiled into a single PDF document. Please include any external references as hyperlinks within the document.
File name format: Name your submission file as: SLLM-AAI-[Your Full Name]
Submission link: Once completed, upload the PDF to the Google Form linked here.
Deadline: 07 November 2025, by 3.00 PM