Global Architecture Document (CMI IS & Document Reader)

The flow of the architecture of the CMI IS & Document Reader consists of multiple stages. The reader can be invoked programatically via APIs, and will also provide a User Interface (UI) enabling end users to upload a document and launch a classification, summarization, topic modelling, NER or Q&A feature.

1. Entry Points - Documents can enter the system either through the User Interface (UI), where end users upload files, or via an API Gateway, where external systems send documents programmatically.

2. Document Ingestion and Normalization - This layer standardizes all the inputs, format conversion (PDF → text), OCR for scanned images, and attach metadata is done in this layer.

3. Routing & Orchestration - The Orchestrator decides how the document will be processed i.e. synchronously or asynchronously. The synchronous mode is for quick and immediate processing of small/medium size documents, for large documents the processing is done asynchronously and is directed to a message queue for batch processing. The routing also invokes different engines based on the requested task i.e. classification, summarization, topic modelling, NER or Q&A.

4. Results & Secure Storage - Results of every engine are saved in a Results & Secure Storage Layer, along with raw and processed documents. A Security Layer enforces role-based access control, encryption, and audit logging.

5. **Results Delivery -** Processed results are delivered back through multiple channels, User Interface (UI) → for end users to view/download, Notifications/Callbacks → alerts or job completion messages for async requests, API Responses → direct delivery of results to external systems.

This is a high level architecture for the proposed CMI & Document Reader, we can improve it further, also the architecture is scalable as each engine is an independent entity and we can always switch with the best models or retrain models for the specific tasks.

Global Architecture Diagram -