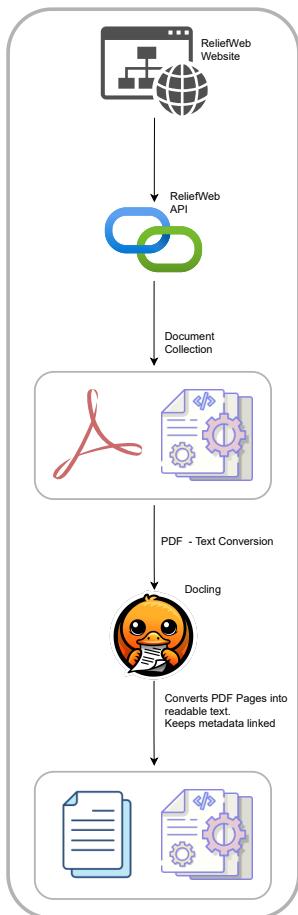
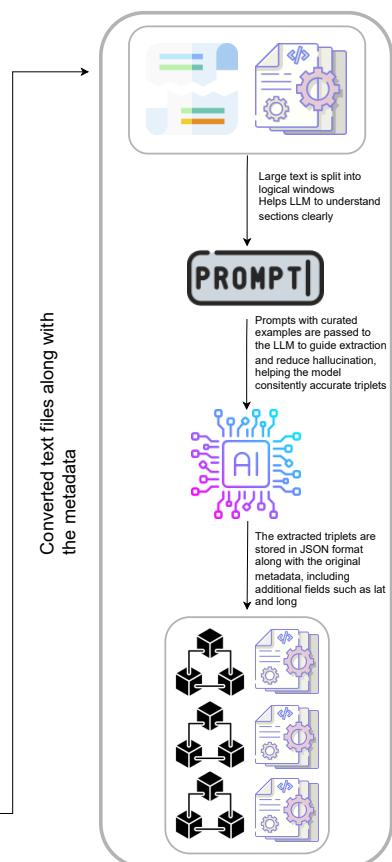


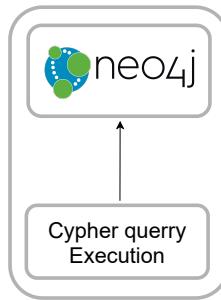
Data Preparation



Data Processing



Knowledge Graph Construction

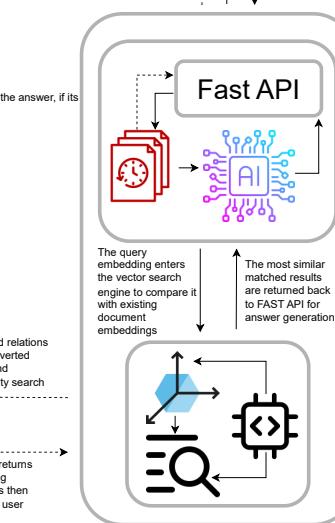
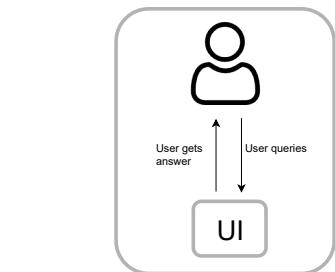


The generated JSON triplets are ingested through Cypher scripts into Neo4j, where subjects and objects become nodes, relations become edges, and all metadata is stored as properties inside the Knowledge Graph

- Selected required reports using ReliefWeb filters (country, category, date).
- Used ReliefWeb API to download PDFs + metadata automatically.
- Collected 5 key report types: Analysis, Appeals, New & Press Release, UN Documents, Situation Reports.
- Countries focused: South Sudan and Ukraine.
- Passed downloaded PDFs to Docing for text extraction.
- Extracted text and kept original metadata linked to each document.

- The extracted text + metadata are split into fixed window sizes.
- Each text window is combined with the relevant metadata.
- Windows are sent to the LLM through API calls with a focused prompt.
- LLM generates subject-relation-object triplets from each window.
- Triplets are saved together with original metadata.
- Final output is stored in structured JSON format.

Frontend: User Interaction



FastAPI Retrieval Service