<https://chatgpt.com/share/68810b26-049c-8005-949f-50332f3644cb>

**Step 5: Advanced: Build tools combining LLM + structured search**

You can:

* Create a **RAG (retrieval-augmented generation)** pipeline:
  + Store metadata (titles, abstracts, sample info) in a vector database (FAISS, Chroma).
  + Index with sentence embeddings (e.g., using sentence-transformers).
  + Ask natural language questions → retrieve relevant datasets → LLM summarizes.
* Build dashboards: search for scRNA-seq datasets by keyword and get summaries.