**Problem Statement**

**Scenario**:  
A major banking institution manages **millions of documents** spread across multiple teams, departments, and repositories (e.g., local drives, SharePoint sites, third-party cloud storage). The documents include **policies, client agreements, compliance reports, operational manuals**, and **training materials**. This scattered storage leads to challenges in document retrieval, version control, and inter-departmental collaboration.

**Business Need**:  
The bank’s leadership wants to **consolidate** all documents into a **centralized, secure document management platform**. Additionally, they require the system to:

* Allow users to **search** for information across all documents quickly.
* Provide a **chatbot interface** where users can **ask questions** in natural language and retrieve relevant document excerpts.
* Allow users to **download documents or excerpts** along with **citations** (i.e., references to the document source, title, and date).
* Ensure **data security**, **compliance auditing**, and **role-based access control**.

**Requirements**:

* Propose an **architecture** or **solution framework** that enables this centralized document management.
* Explain how **document ingestion**, **indexing**, and **retrieval** will work at scale (millions of documents).
* Suggest technologies that can enable **chatbot interactions** over the documents.
* Discuss how **citations** can be automatically generated and included with chatbot responses.
* Outline security measures to protect sensitive banking information.