Company Design Doc — Knowledge Discovery Assistant Proposal

To: IT Architecture Council
From: Sr. Data Architect
Subject: Proposed Integration Architecture - Knowledge Assistant

Architecture Notes:
- Document ingestion pipeline will normalize PDFs and emails.
- Sentence embedding models (MiniLM) selected for initial vector search.
- Retrieval logic should support hybrid: SQL, vector, and entity graph.
- Model output must include traceable citations per document chunk.

Dependency:
- Qdrant instance setup in internal Docker cluster
- Neo4j sandbox for entity graph tests

Expected Outcome:
AllyIn Compass should support natural queries like:
> "Show emissions violations since Q2 by location"
> "Summarize audit issues in compliance emails"

— Sr. Data Architect