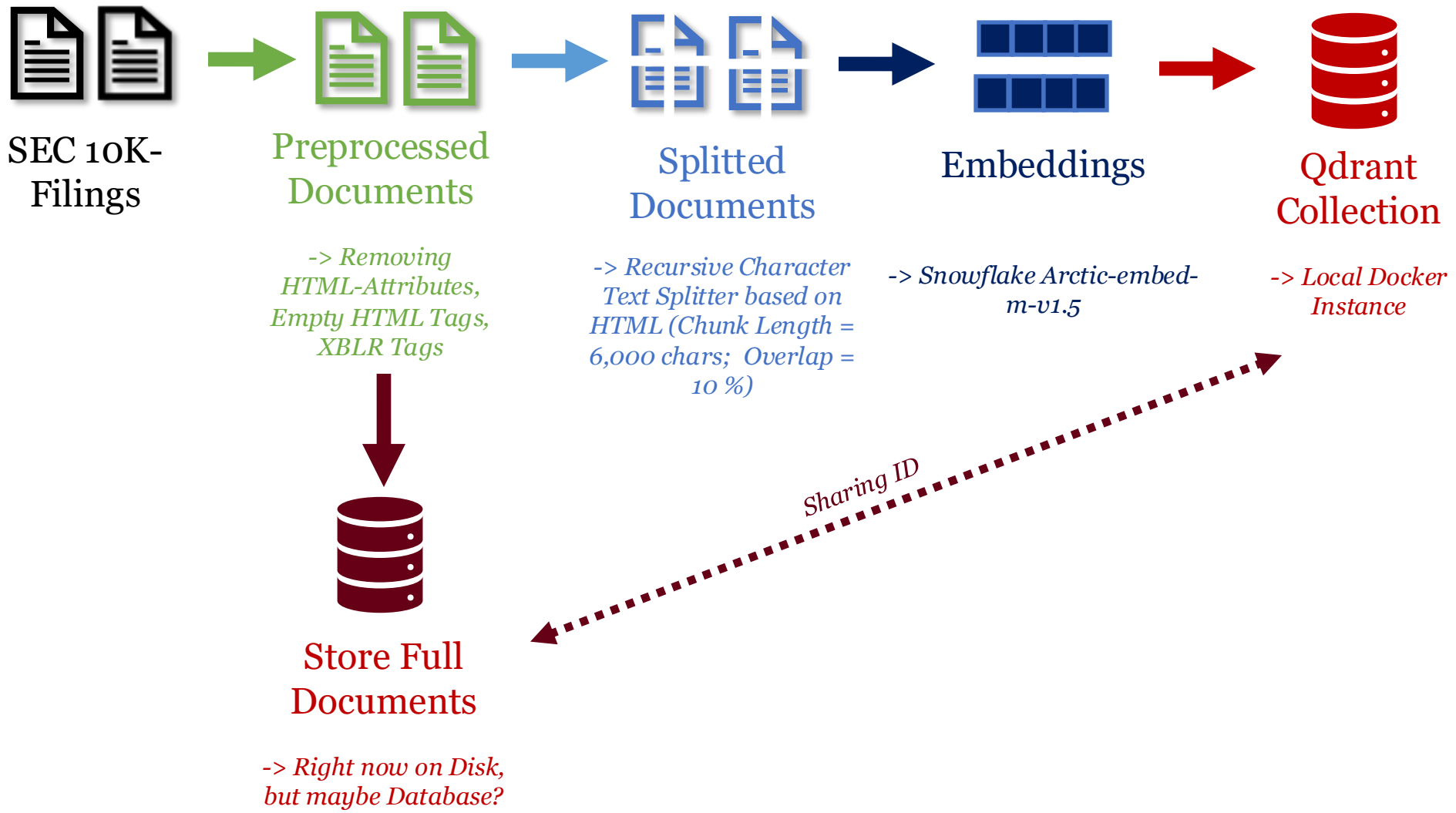
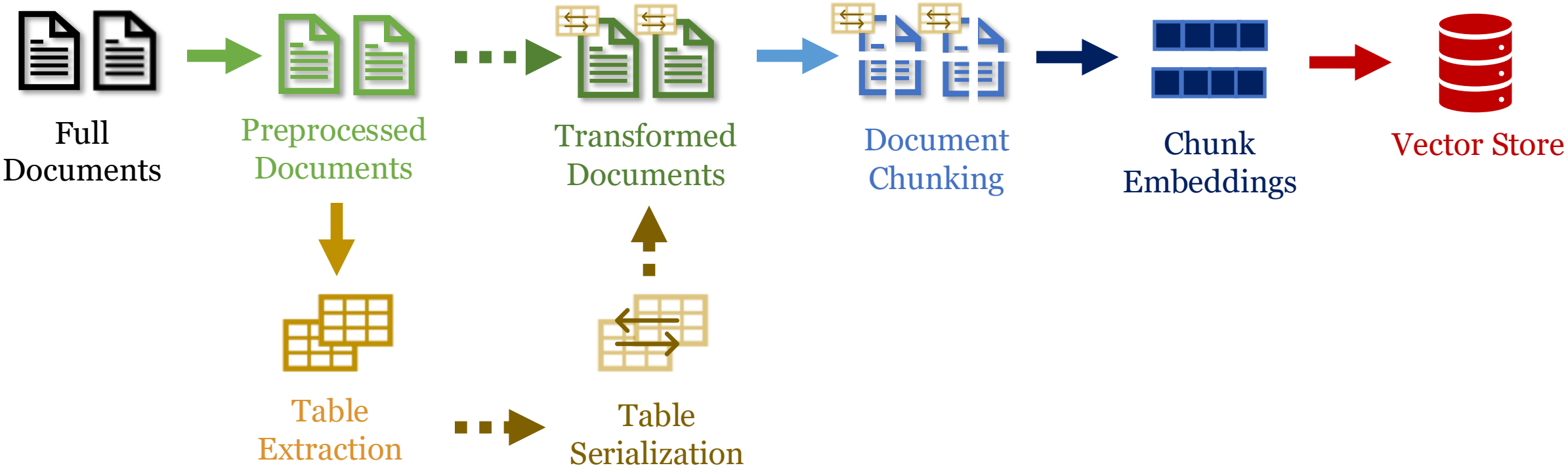
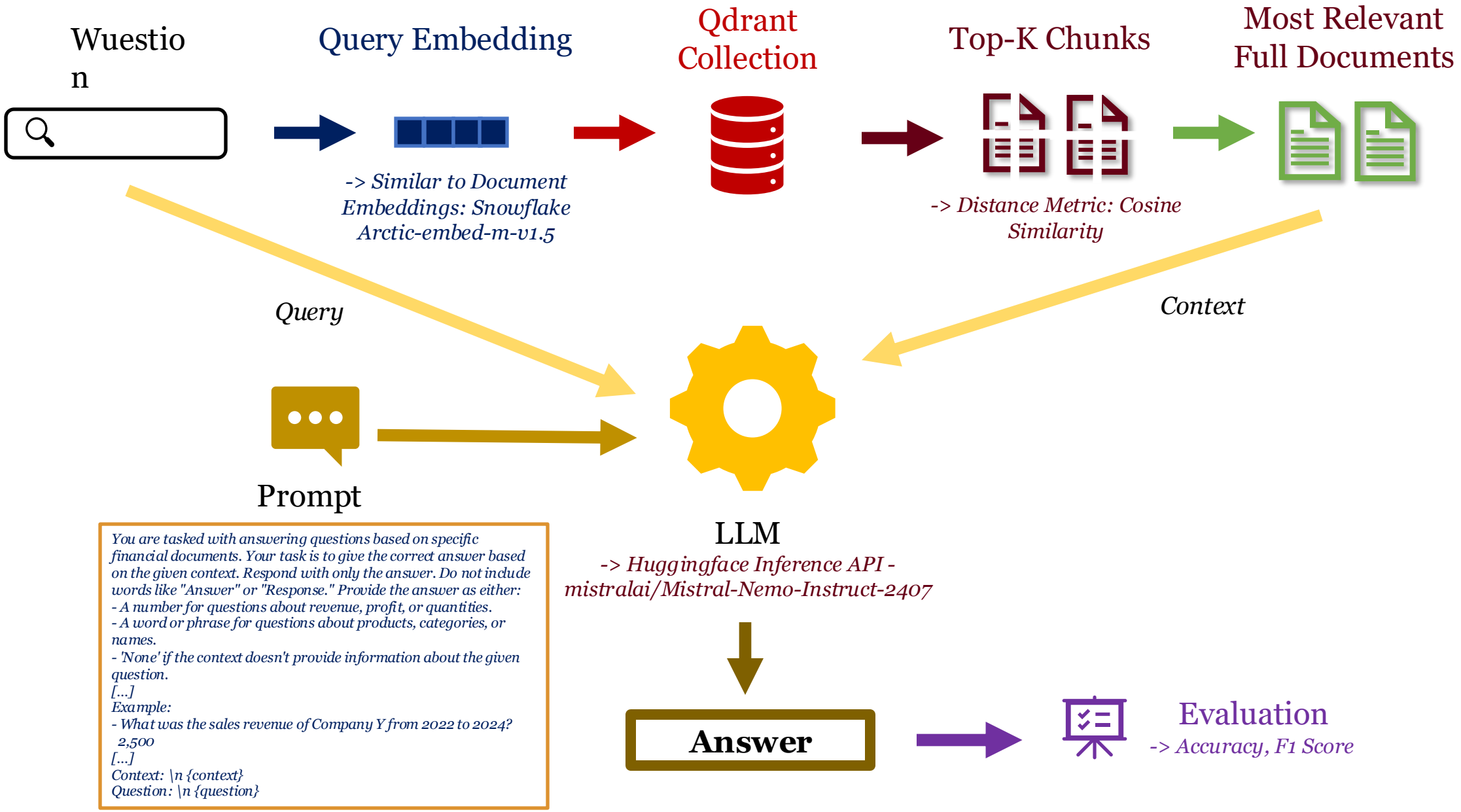
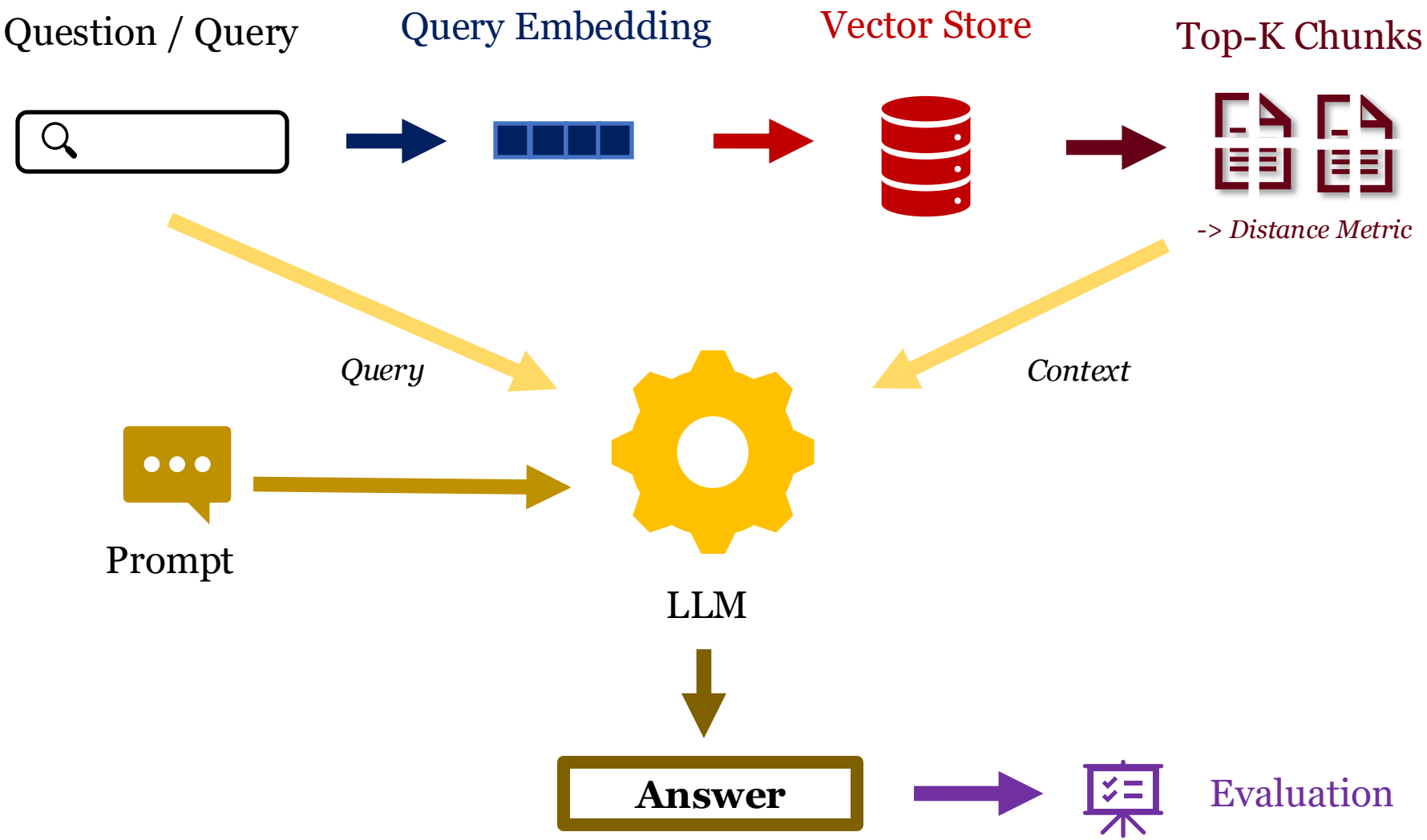

SEC 10K-Filings
Preprocessed Documents
-> Removing HTML-Attributes, Empty HTML Tags, XBLR Tags
Splitted Documents
-> Recursive Character Text Splitter based on HTML (Chunk Length = 6,000 chars; Overlap = 10 %)
Embeddings
-> Snowflake Arctic-embed-m-v1.5
Qdrant Collection
-> Local Docker Instance
Sharing ID
Store Full Documents
-> Right now on Disk, but maybe Database?


Full Documents
Preprocessed Documents
Transformed Documents
Document Chunking
Chunk Embeddings
Vector Store
Table Extraction
Table Serialization


Wuestio
n
Query Embedding
-> Similar to Document Embeddings: Snowflake Arctic-embed-m-v1.5
Qdrant Collection
Top-K Chunks
-> Distance Metric: Cosine Similarity
Most Relevant Full Documents
Query
Context
Prompt
You are tasked with answering questions based on specific financial documents. Your task is to give the correct answer based on the given context. Respond with only the answer. Do not include words like "Answer" or "Response." Provide the answer as either:
- A number for questions about revenue, profit, or quantities.
- A word or phrase for questions about products, categories, or names.
- 'None' if the context doesn't provide information about the given question.
[...]
Example:
- What was the sales revenue of Company Y from 2022 to 2024?
2,500
[...]
Context: \n {context}
Question: \n {question}
LLM
-> Huggingface Inference API - mistralai/Mistral-Nemo-Instruct-2407
Answer
Evaluation
-> Accuracy, F1 Score


Question / Query
Query Embedding
Vector Store
Top-K Chunks
-> Distance Metric
Query
Context
Prompt
LLM
Answer
Evaluation