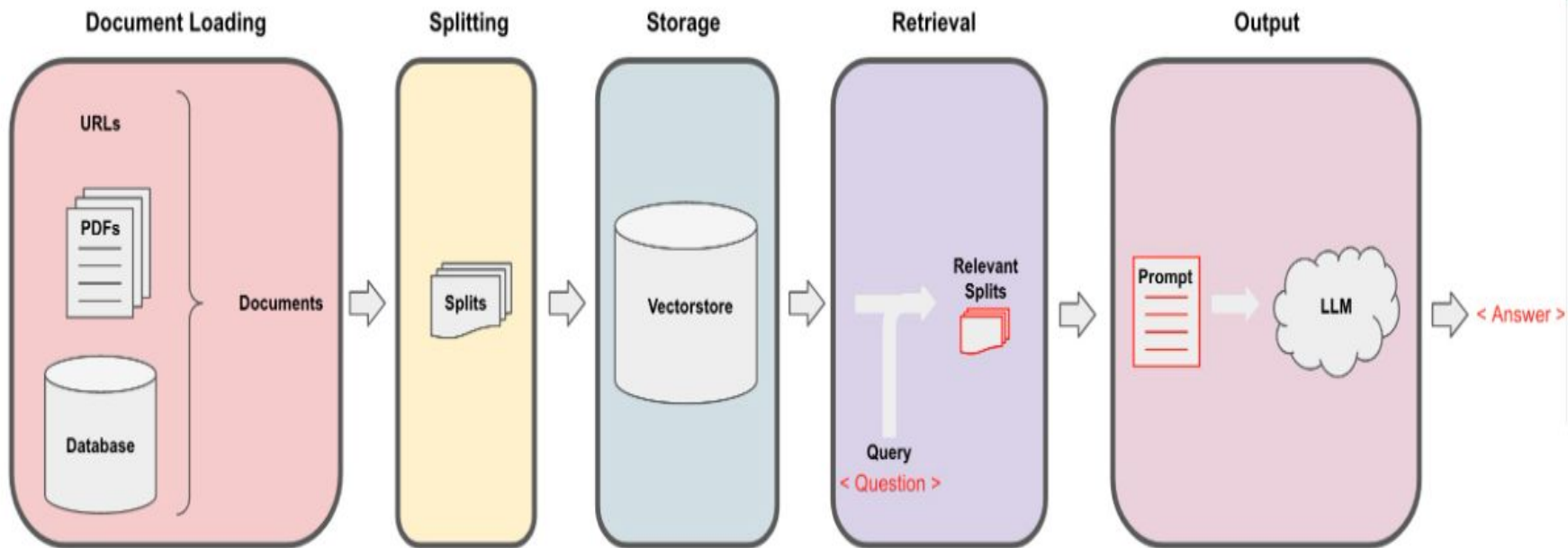


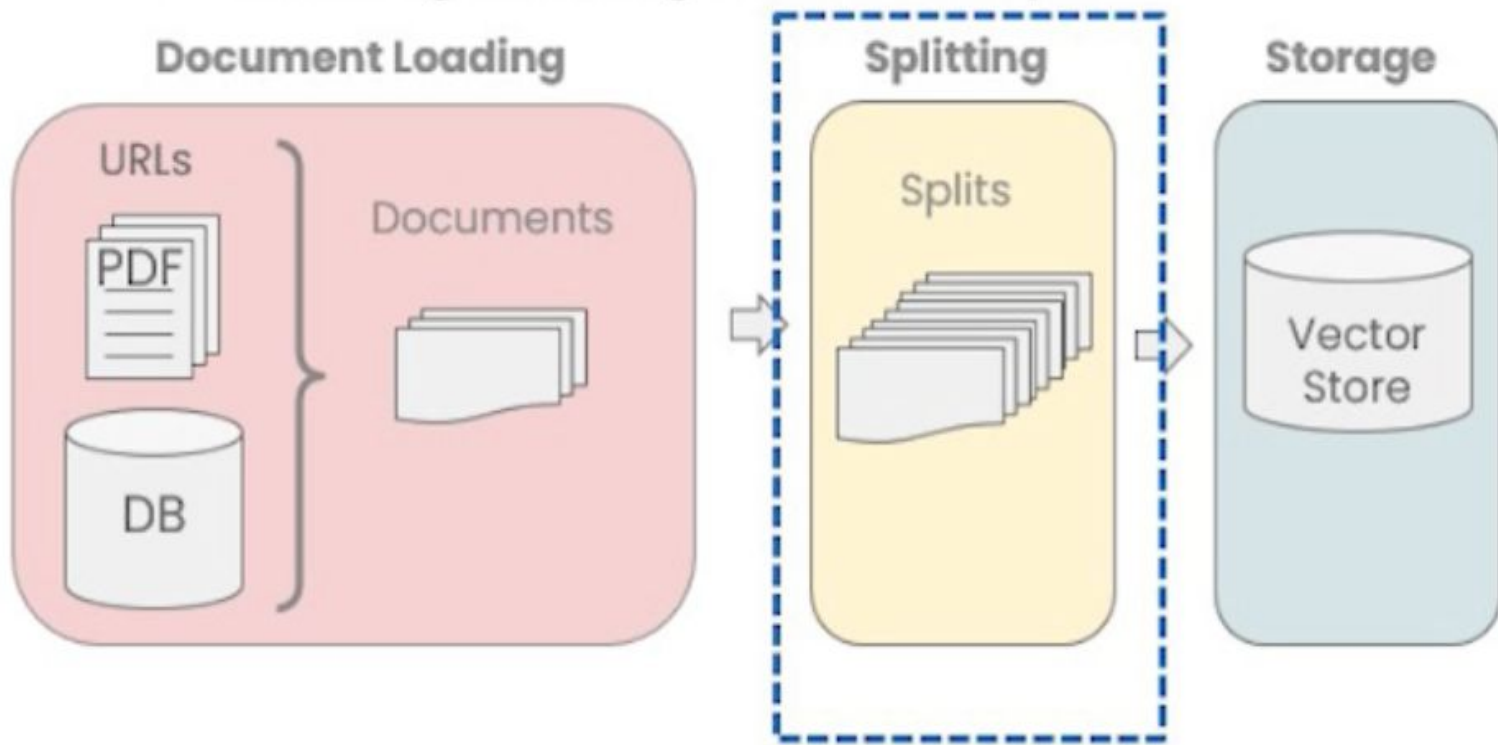
CWB-Module-2

GenAI

Process

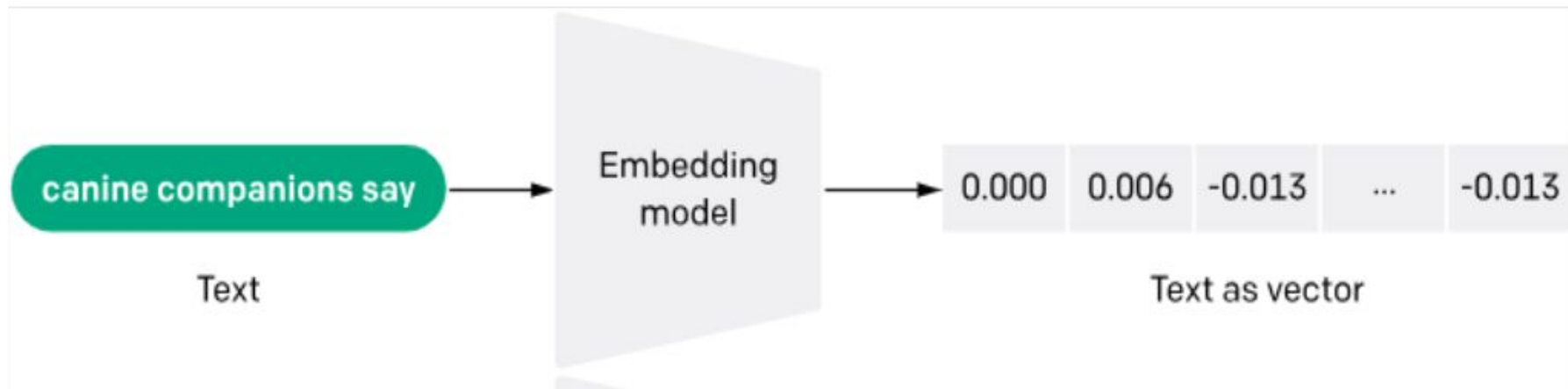


- Splitting Documents into smaller chunks
 - Retaining meaningful relationships!



Embedding

Text embeddings are a way to convert words or phrases from text into numerical data that a machine can understand. Think of it as turning text into a list of numbers, where each number captures a part of the text's meaning.



Vector Stores

1. Vector stores are **databases explicitly designed for storing and retrieving vector embeddings efficiently.**
2. They are needed because **traditional databases like SQL are not optimized for storing and querying large vector data.**
3. Vector stores **can index and quickly search for similar vectors** using similarity algorithms. It allows applications to find related vectors given a target vector query.

Chroma DB

Chroma DB is an open-source vector store used for storing and retrieving vector embeddings. Its main use is to save embeddings along with metadata to be used later by large language models.

