

## Vector database

### INFORMATION EXTRACTION

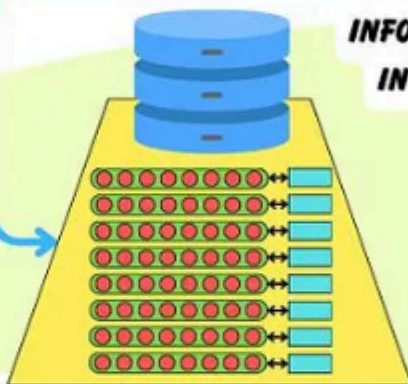


The data is partitioned in chunks

1

### INFORMATION INDEXING

The chunks are encoded into vectors with an LLM and indexed into a vector database by their vector representation



Vector database

2

### INFORMATION RETRIEVAL

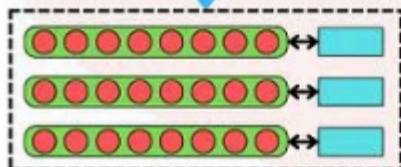
Question?



Embedding representation of the question



Vector database



Nearest neighbors to the question

Information retrieval is achieved by providing the nearest neighbors to the question as context to the LLM

3