

Presented to you by www.happyoverton.com

We need RAG - but

Same shelf, different job...

The confusion

PEOPLE NAME THE PATTERN – AND BUY THE STORE

Teams say "we need RAG" and shop for Pinecone

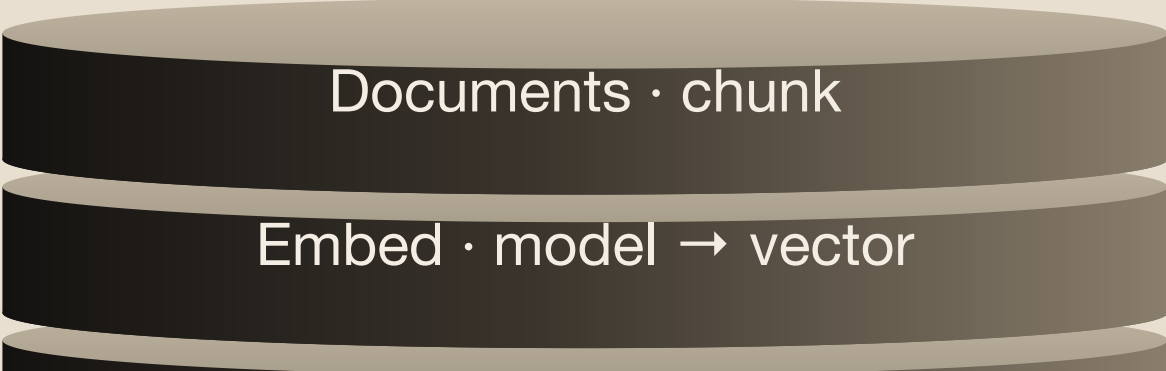
A vector database is infrastructure. RAG is a loop

Vector database

INFRASTRUCTURE

Embed, index, query by similarity.

Neighbours out — not an answer.



Documents · chunk

Embed · model → vector

Index · HNSW / IVF

Top-k · ids · scores

Does not write the answer.

Store + ANN index

out: cities · scores · stop · LLM? no

id	city
01	Lisbon
02	Reykjavik
03	Tokyo

YOU ASK

"spring city walks?"

embed

Vector DB

RESULT

id	city	score
01	Lisbon	0.89
03	Tokyo	0.71



STOP

Do not conf