O'REILLY®

Data Storing

Bert Gollnick

Learning Objectives

By the end of this module, you will:

- Know how data can be stored in vector databases
- Why vector databases are necessary
- How similarity search works
- Which db providers are available

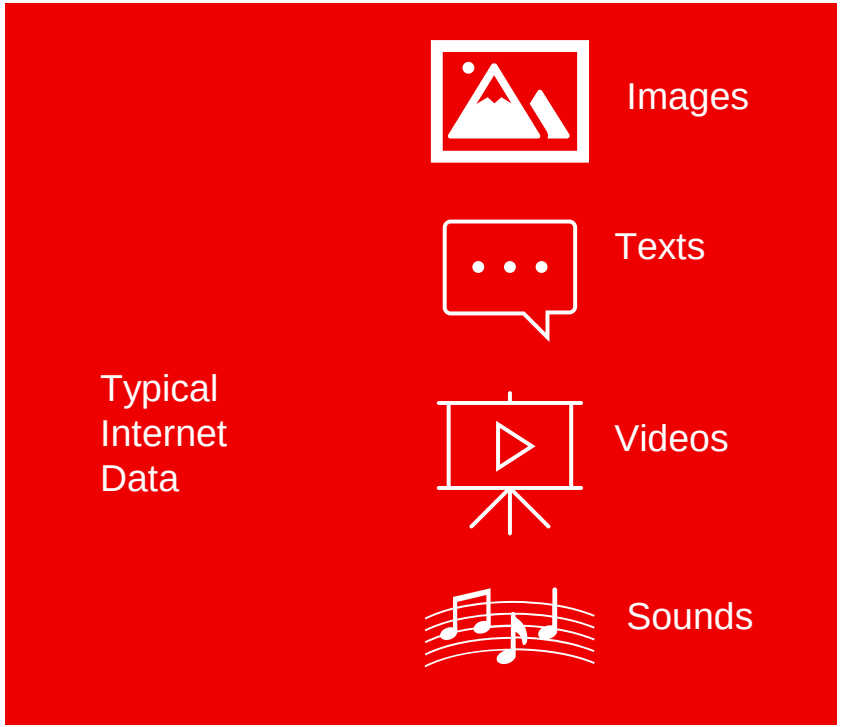
What is a vector database?

A vector database stores high-dimensional data (embeddings) for fast querying and similarity analysis.

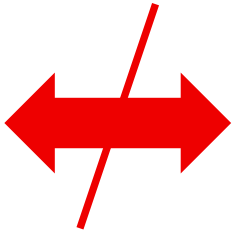
Features

- Special type of database
- Allows to store, manage, and query data which is represented in geometric formats
- Enables similarity search, clustering, real-time analytics

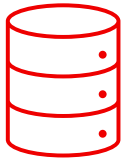
Why is a vector DB required?



Unstructured Data



Typical
Databases

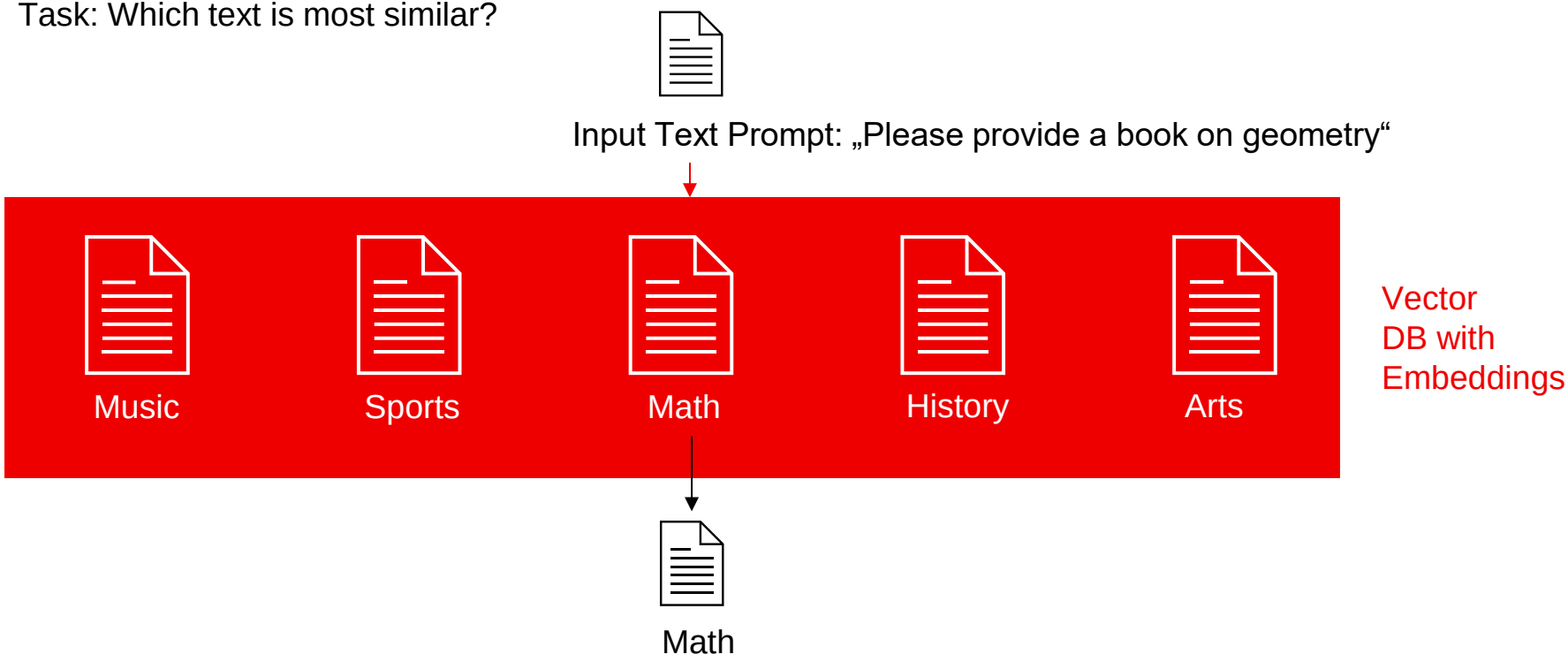


SQL

Structured Data

Use Case: Text Querying

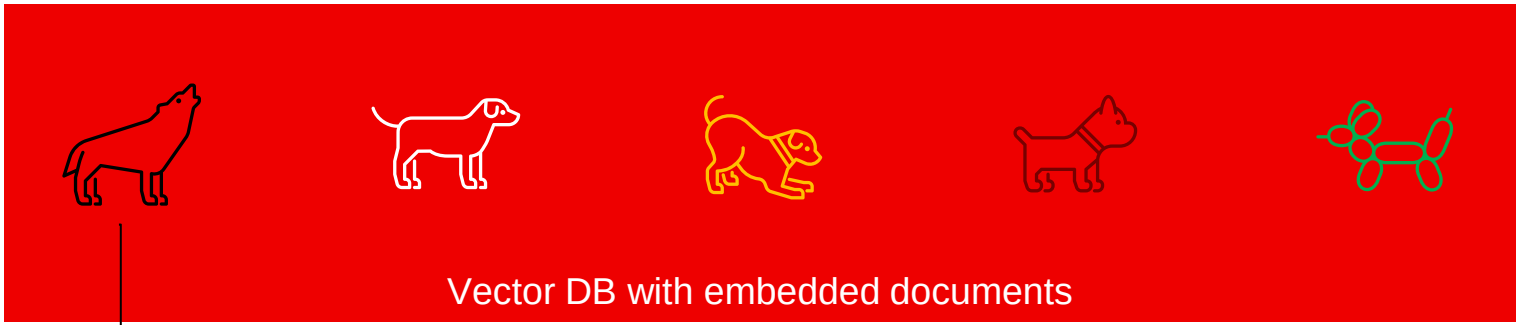
Task: Which text is most similar?

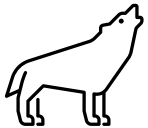


Use Case: Image Querying

Task: Which image is most similar to a text prompt?

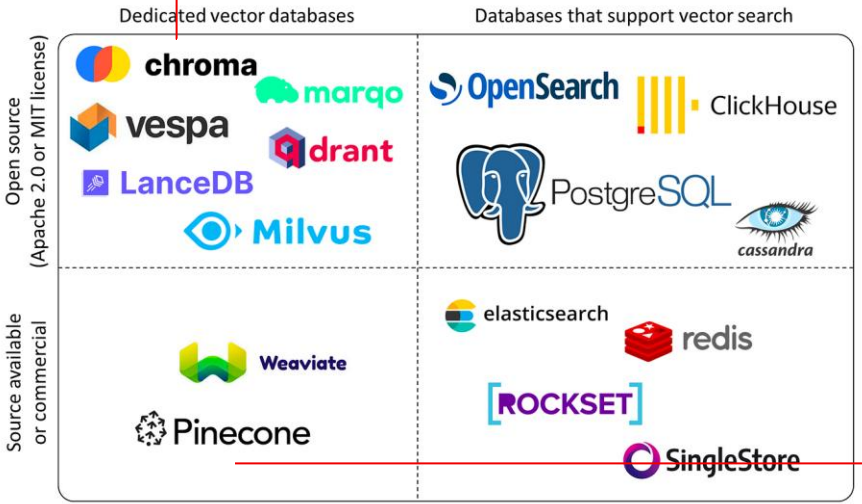
Prompt: Which picture shows a brown wolf?





Most similar document

Vector DB-Providers



Source: https://www.datacamp.com/blog/the-top-5-vector-databases

Pinecone

pinecone.io

web-based

freemium

trychroma.com

local

free

O'REILLY®