Azure Developer Community

Tamil Nadu

What we do?

Meetups

Network of Beginners,
Professionals &
Experts

Azure Cert Study Jam

Corporate / Startup
Talk

Vector DB's Growth

Agenda

Welcome Note

Why Vector DB

Types

Demo

Target Audience:

Level: Intermediate

The session is intended for developers and beginners who are new to Vector DBs.

Introduction

- Experience in Analytics, Reporting and data engineering

- Azure Dev Lead and Microsoft Certified - Azure and AI Engineer

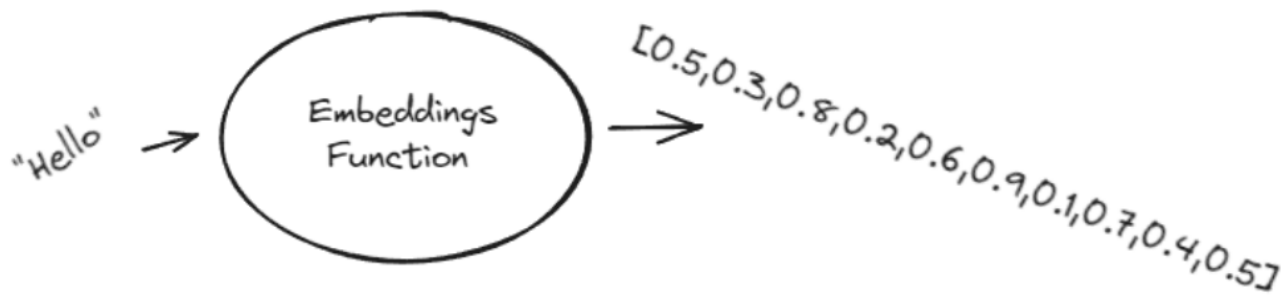- Enthusiastic learner and tech community contributor

Navaneethan Gopal
Data Science Manager @ Ingram Micro

LinkedIn – https://www.linkedin.com/in/navaneethan-gopal-9b101612/
Twitter - https://twitter.com/NavaneethhGopal

Vector Example

"Hello" → Embeddings Function → [0.5,0.3,0.8,0.2,0.6,0.9,01,0.7,0.4,0.5]

10 dimensional Vector

Vector Example

Ship sails in the ocean
boat floats in the sea
A vessel cruises through the vast ocean.
"A boat peacefully glides on the tranquil lake

all-MiniLM-L6-v2

0.1,-0.2,0.5,...,0.3]

(384 values for each word)
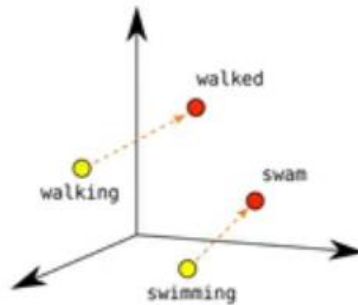
lake Ocean
Sea

Ship vessel
boat

floats
sails
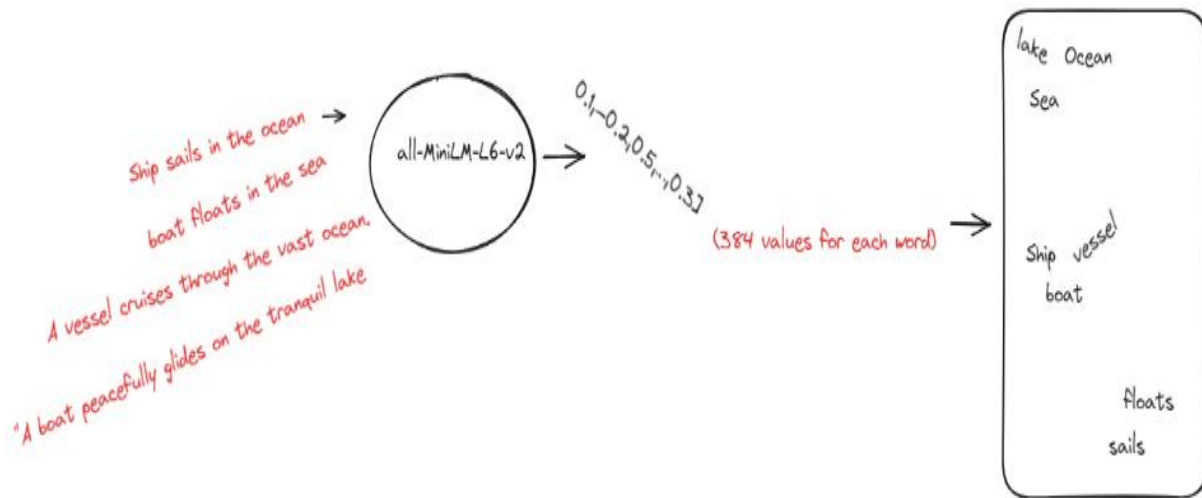
Embeddings Example
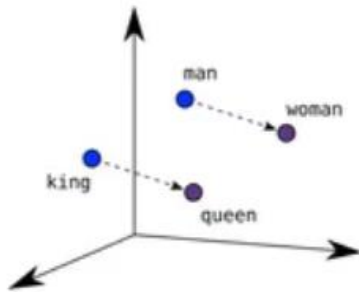
Male-Female   Verb Tense   Country-Capital
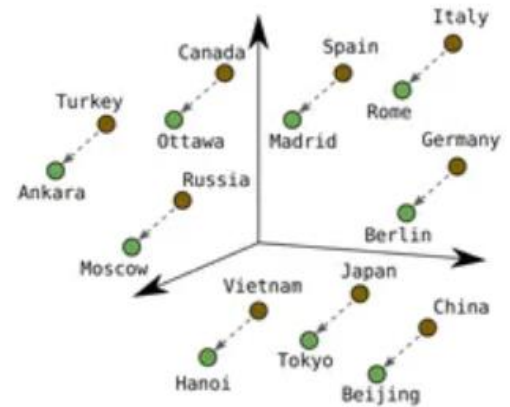
Vector DBs are evolving

🧠 Vector databases are a critical component in the AI stack, solving the stateless problem of LLMs + raising $175.5M in April
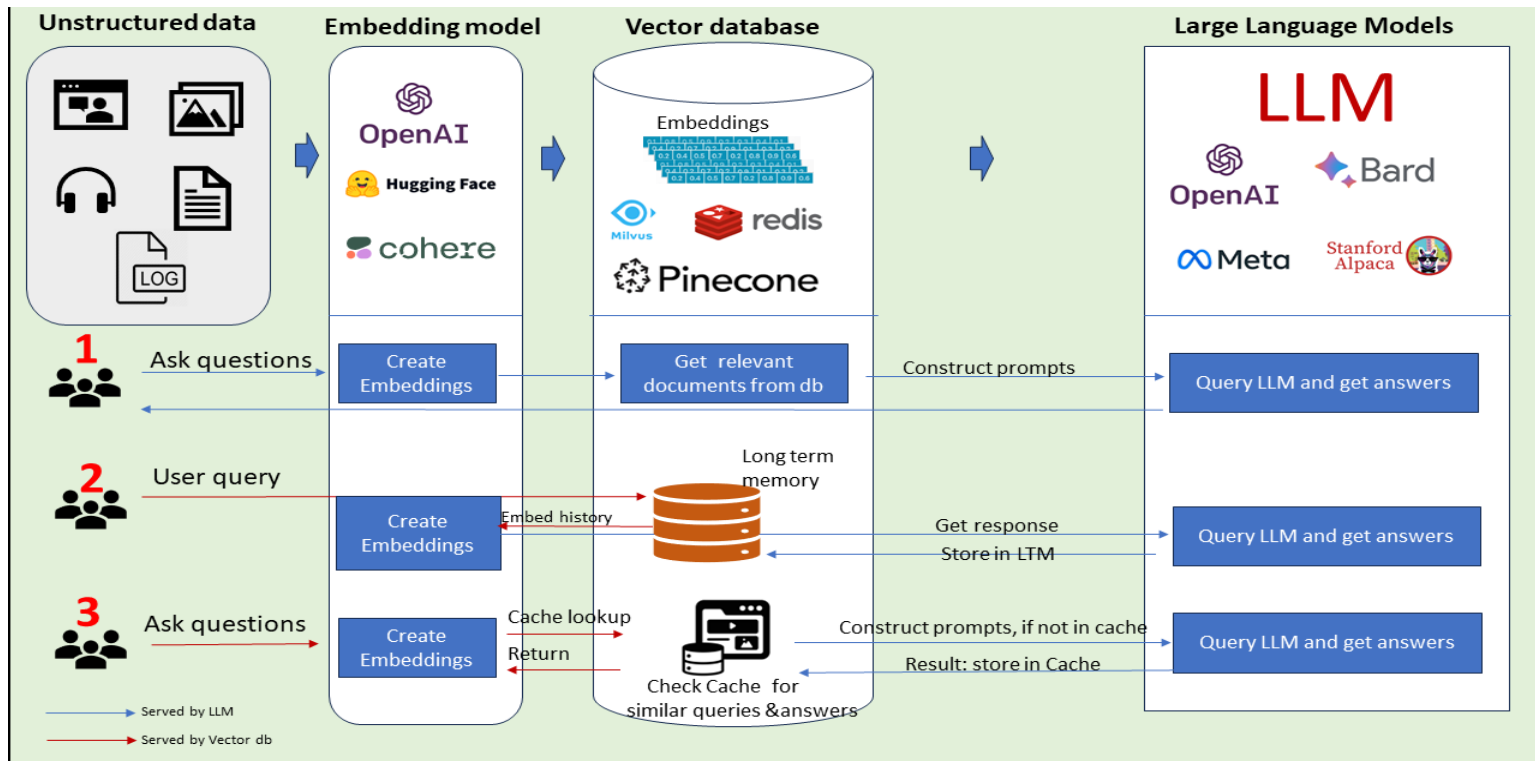
The top tools 👇

– Pinecone ($138M funding)
– Zillis ($113M funding)
– Weaviate ($67.7M funding)
– Chrome ($20.3M funding)
– Qdrant ($9.8 M funding)

Emerging LLM Stack

Emerging LLM Tech Stack

Raw Data → Data Connectors / Preprocessing Tooling → Store Embeddings / Document Index in Vector Database

LLM Endpoint
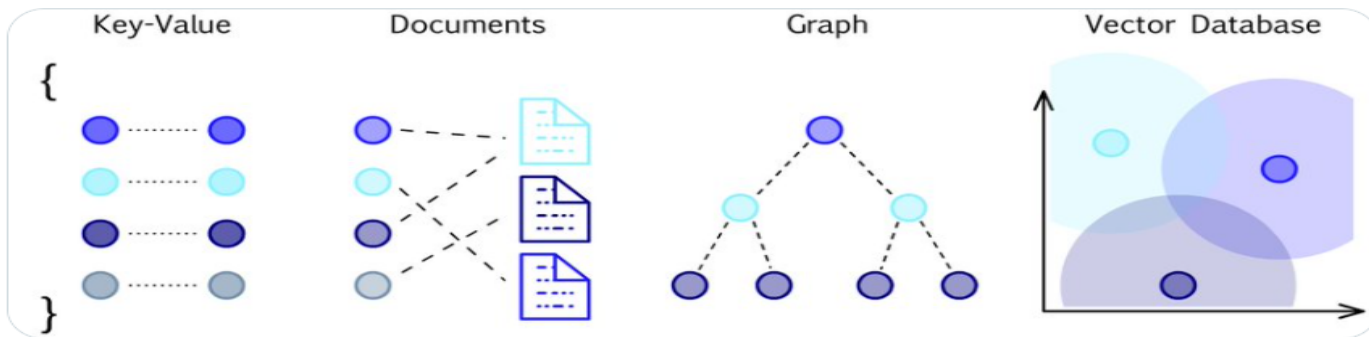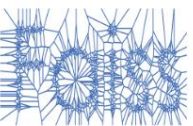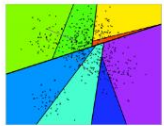
LLM Programming Framework
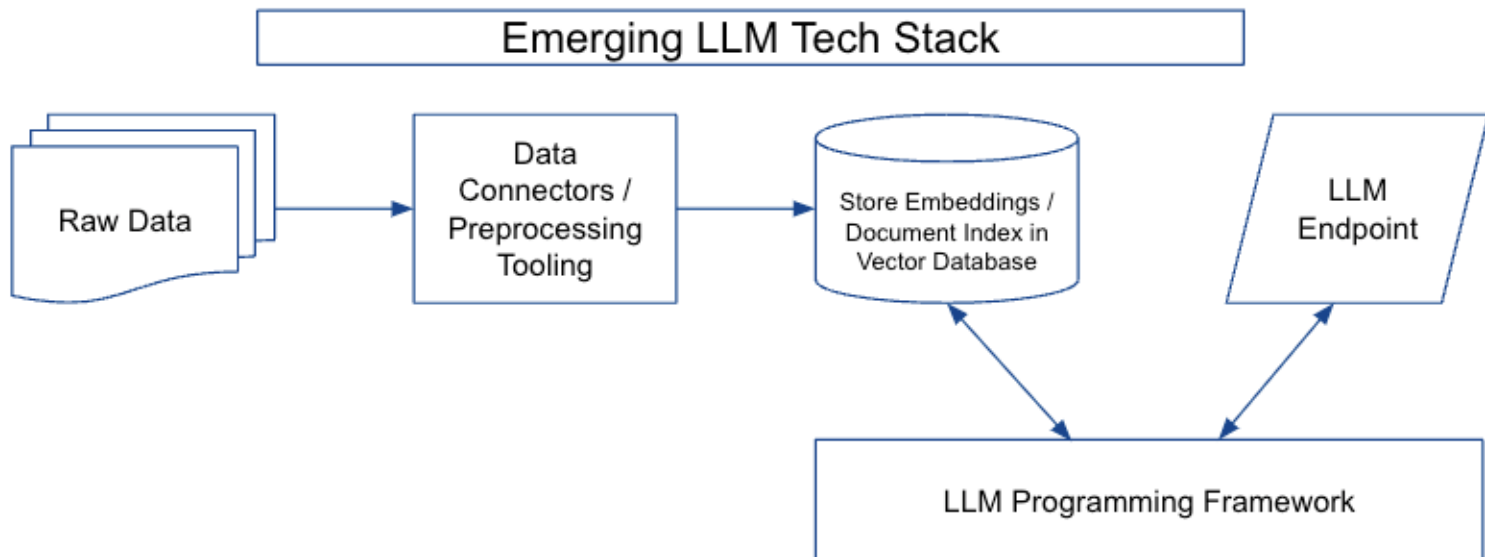
Vector DBs Architecture in Gen AI

Source:

Available Vector DBs

Pure vector databases

Milvus · zilliz · Pinecone · Weaviate · vespa · marqo · qdrant · Chroma · Vald · LanceDB

Text search databases

elasticsearch · APACHE LUCENE · OpenSearch · Apache Solr

Vector-capable NoSQL databases

cassandra · neo4j · MongoDB · DataStax Astra · Azure Cosmos DB · redis

Vector libraries

FAISS · Annoy · Hnswlib

Vector-capable SQL databases

SingleStore · pgvector for Postgres · ClickHouse · kinetica · ROCKSET

Source:

Vector database in Azure

- Azure Cosmos DB Vector Database Extension

- Azure PostgreSQL Server pgvector Extension

- Azure AI Search

Azure Vector Search

Data sources

Images

Audio

Video

Text

Transform into embeddings

Vector representation

$[-2, -1, 0, 1]$

$[2, 3, 4, 5]$

$[6, 7, 8, 9]$

Approximate nearest neighbor

Azure Cognitive Search

Results

Vector representation

$[2, 3, 4, 5]$

Transform into embedding

App/UX
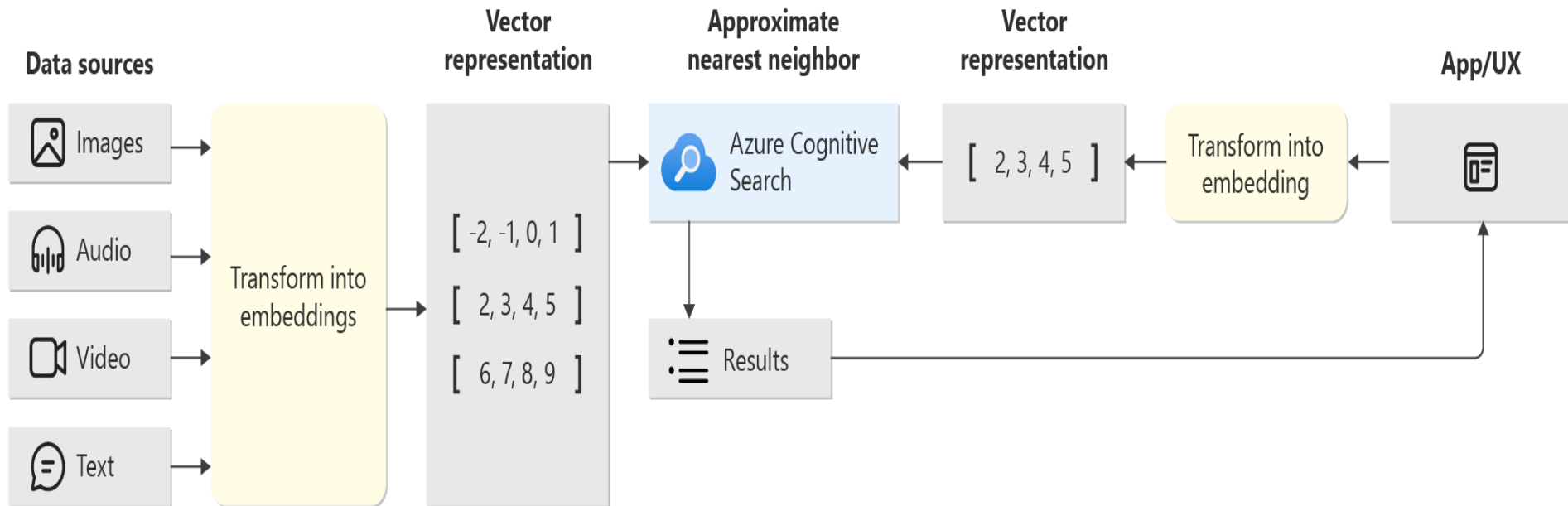
Azure AI Search - RAG patterns

App UX

App Server, Orchestrator

Azure AI Search

Query → Knowledge

Prompt + Knowledge → Response

Azure OpenAI (GPT/ChatGPT)

Data Sources (files, databases, etc.)

SQL
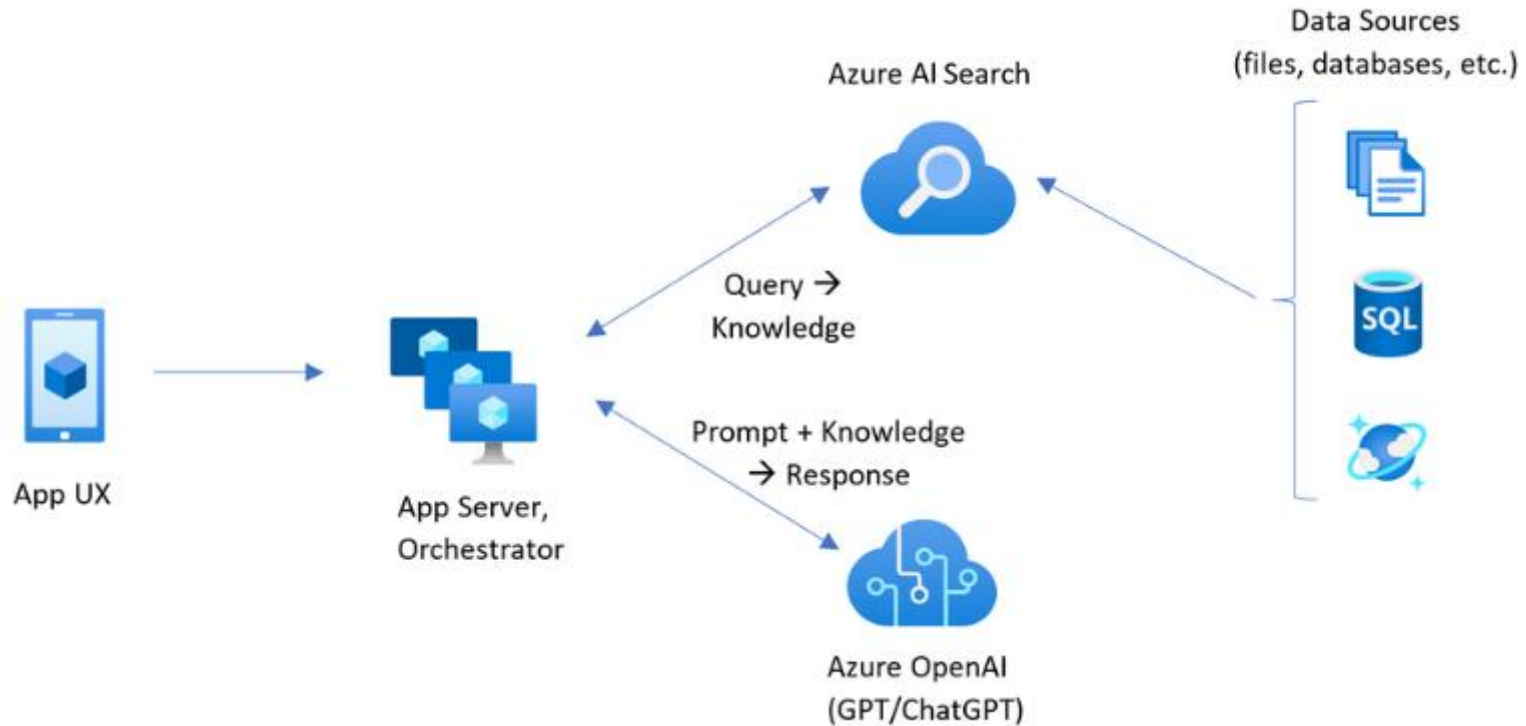
Use case

Why Vector DBs for Gen AI?

High-Dimensional Data Management

Enhanced for LLM Applications

Optimized for Similarity Search

Vector search Benefits

Handling Unstructured Data

Advanced Indexing Mechanisms

Scalability, Speed and Accuracy
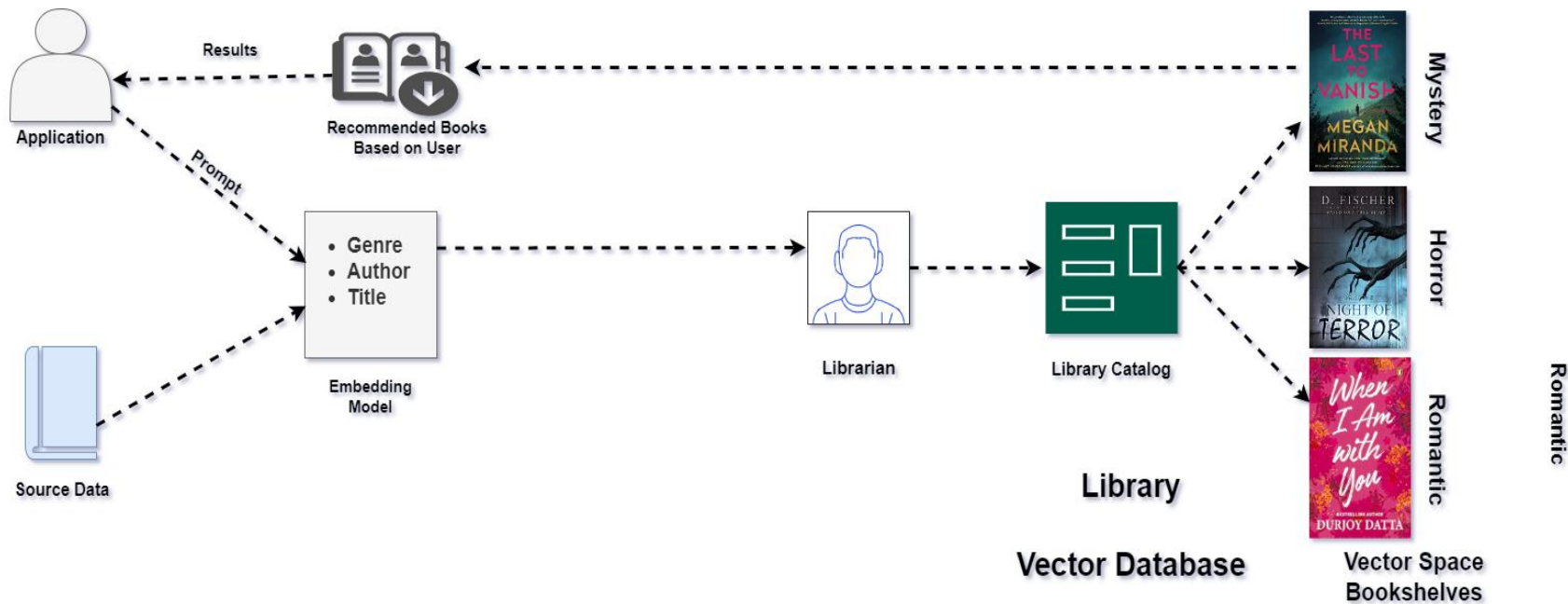
Source:

Learning Resources

- Azure free account: https://azure.microsoft.com/en-in/free/

- QuickStart : https://learn.microsoft.com/en-us/semantic-kernel/memories/vector-db#use-cases-for-vector-databases

- https://github.com/anyscale/ray-summit-2023-training/tree/main/Ray-LlamaIndex/notebooks

- https://learn.microsoft.com/en-gb/azure/search/retrieval-augmented-generation-overview

Source:

Open for Q&A

INGRAM MICRO ®

Thank you!!

Source: