

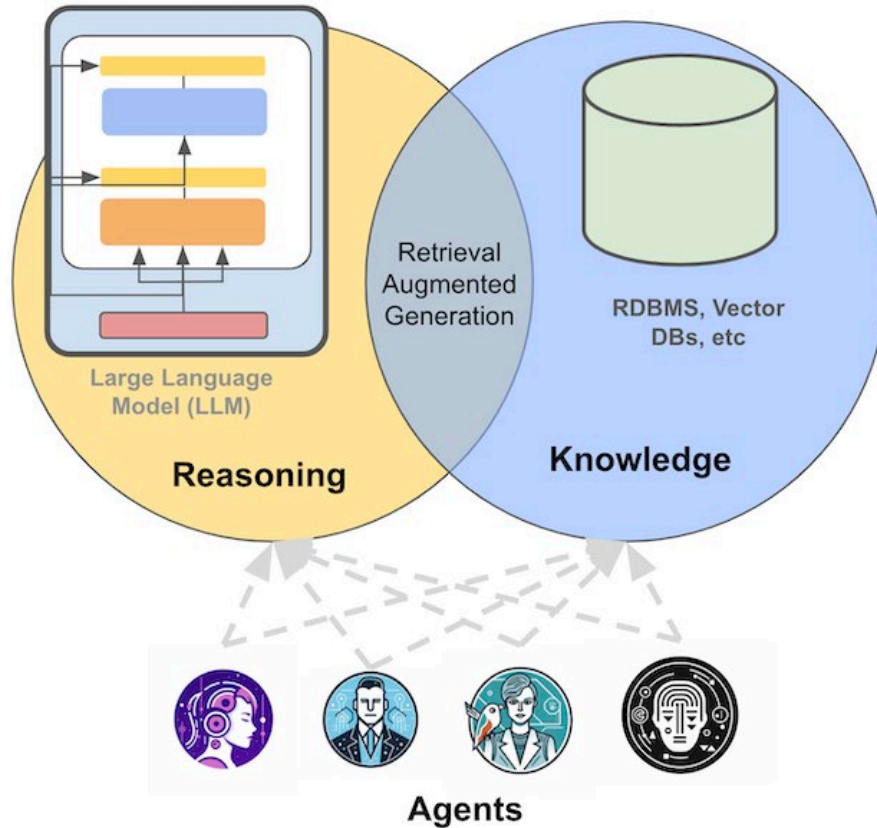
# Architecting Intelligence: Next-Generation Databases for RAG- Powered AI

## **Arun Nandi**

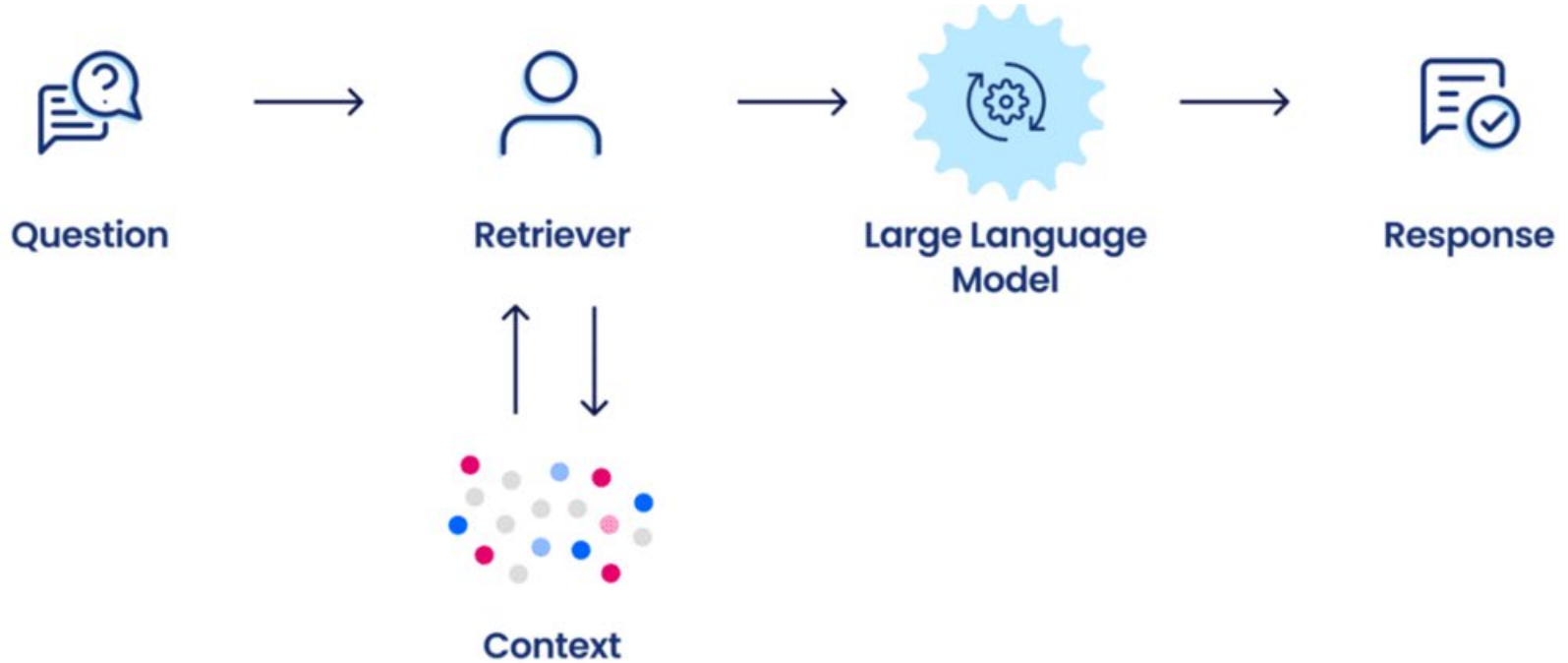
Head of Data and Analytics at Unilever  
Board Advisor ♦ Top 100 Leaders in AI ♦ Keynote Speaker

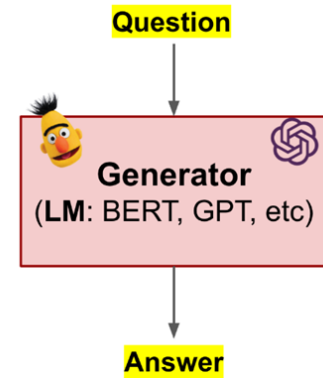


# What does RAG do?



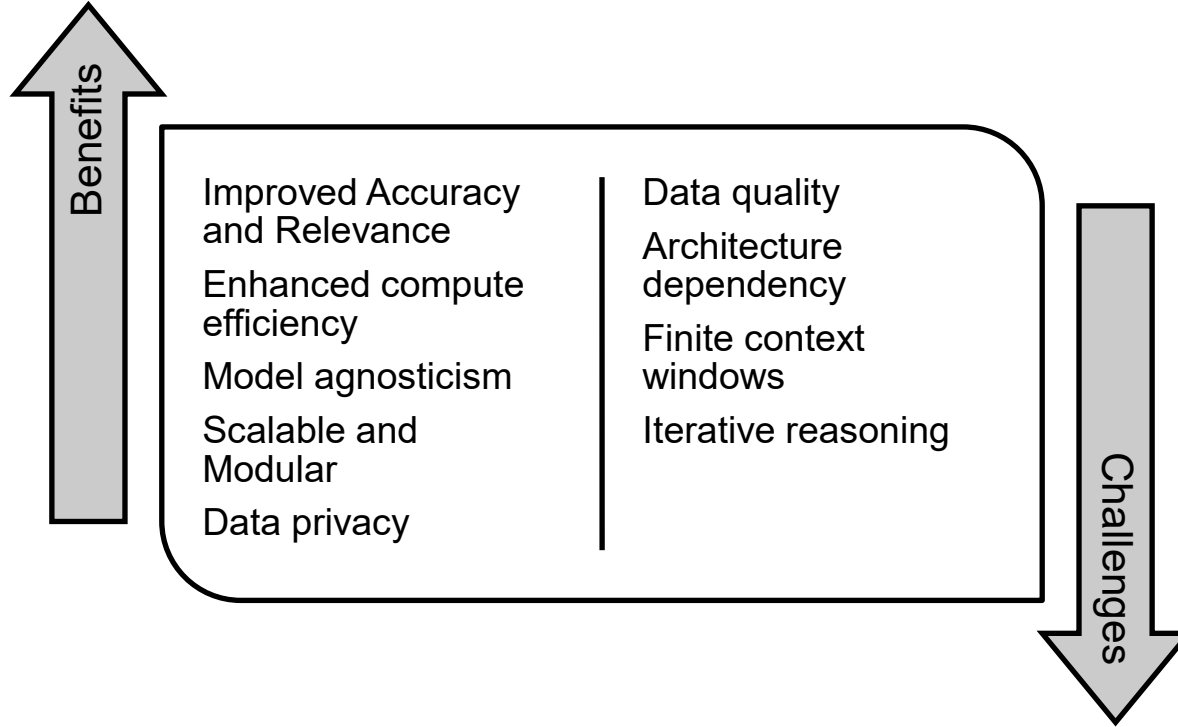
# Retrieval Augmented Generation Schema



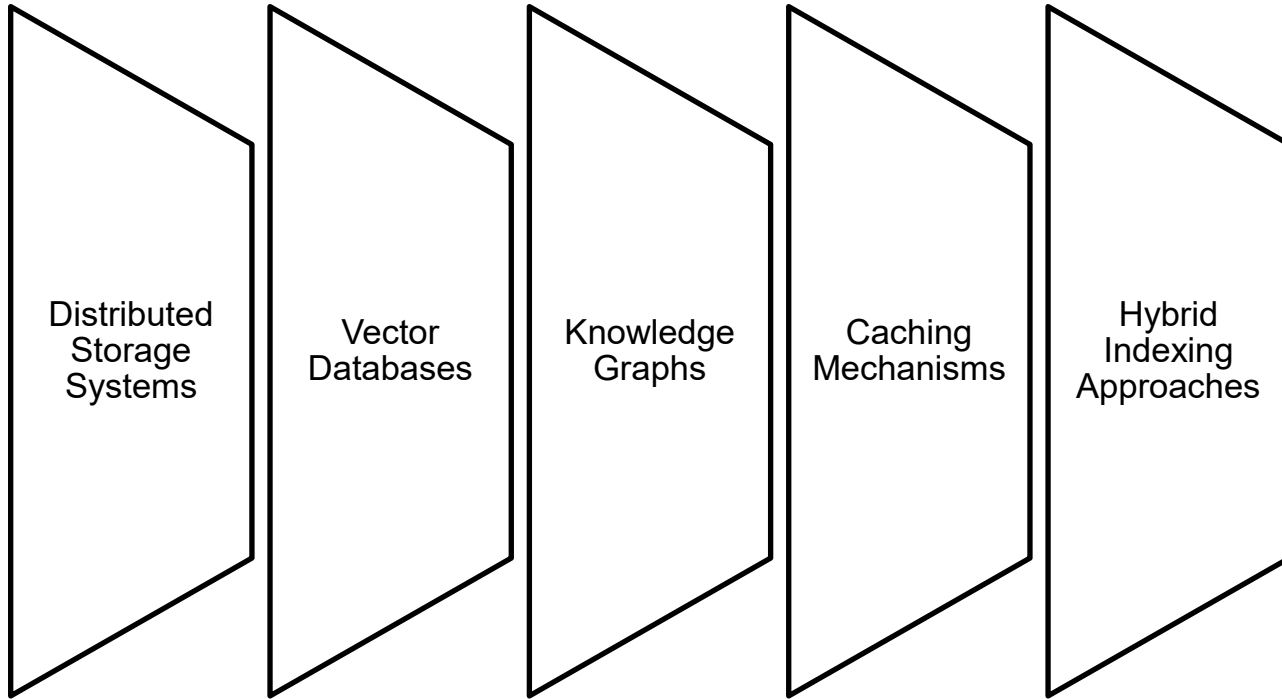


Generator

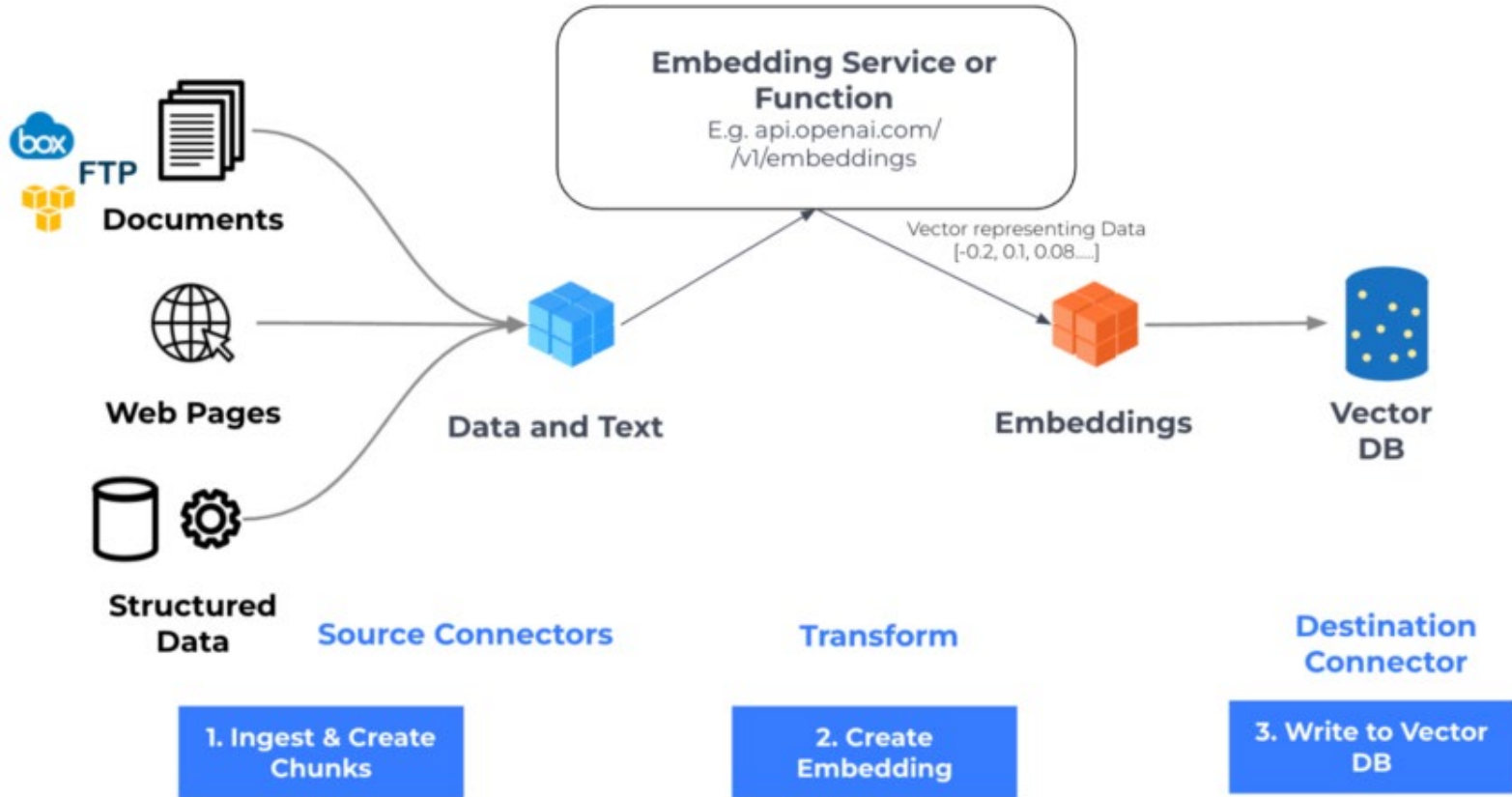
# Query Routing and RAG



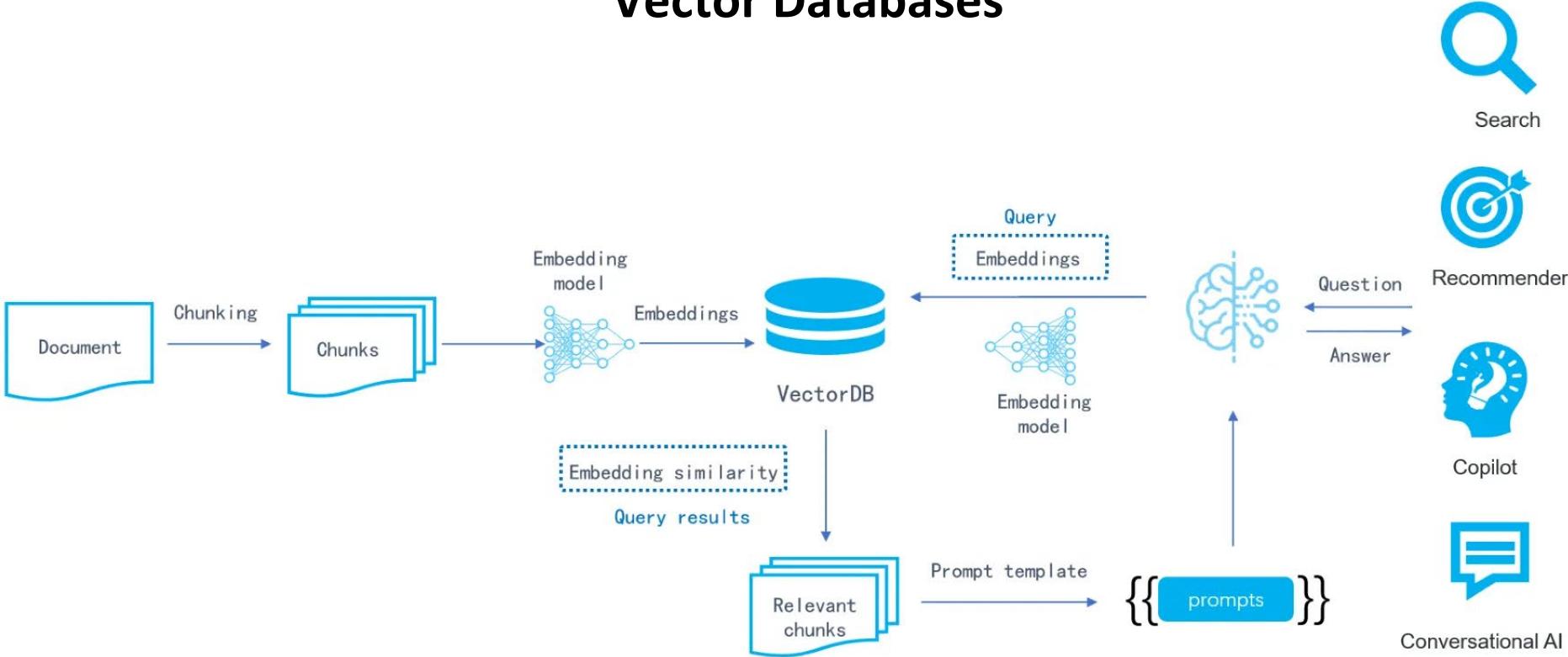
# Database Architectures for Efficient Applied LLM use-cases



# The case for Vector Databases



# Current Generation Architecture using Vector Databases



Search



Recommender



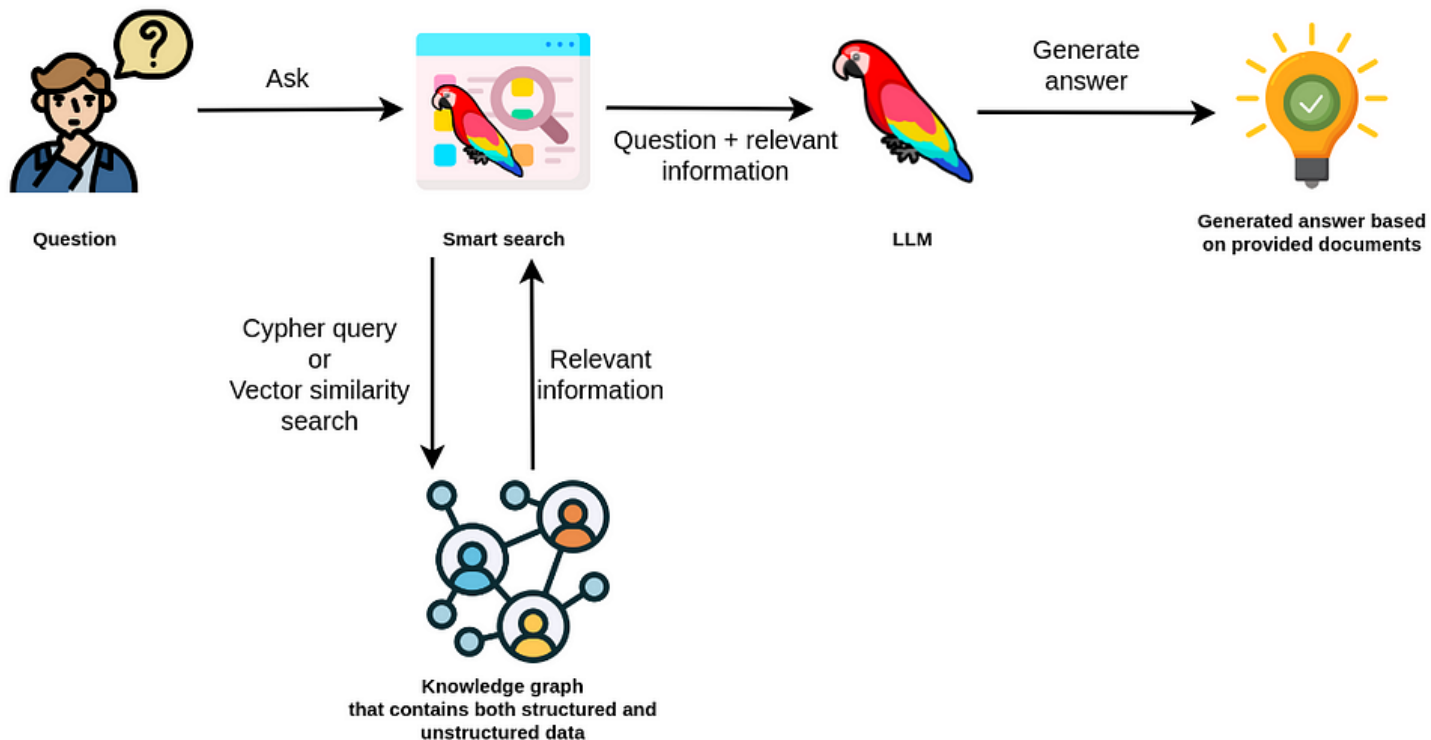
Copilot



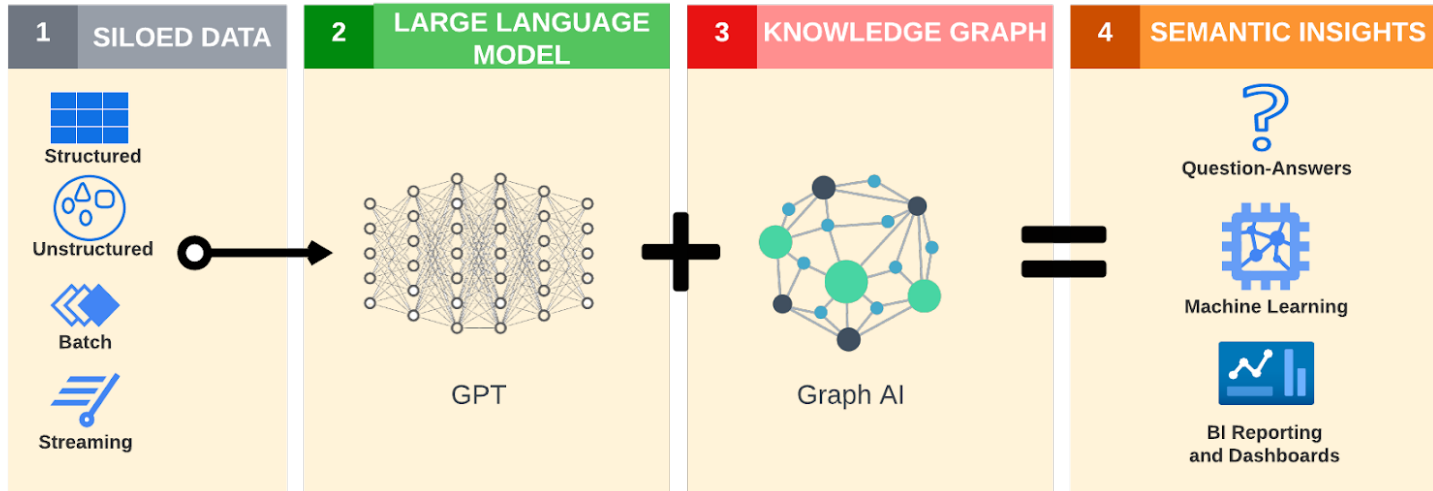
Conversational AI



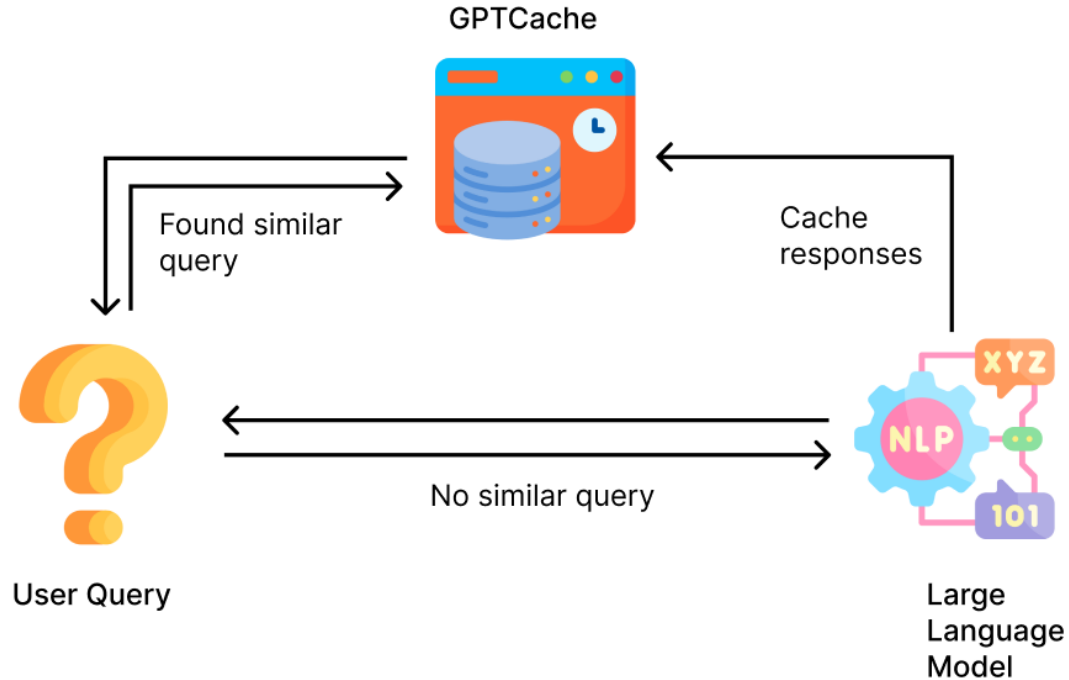
# Knowledge Graphs to feed RAG



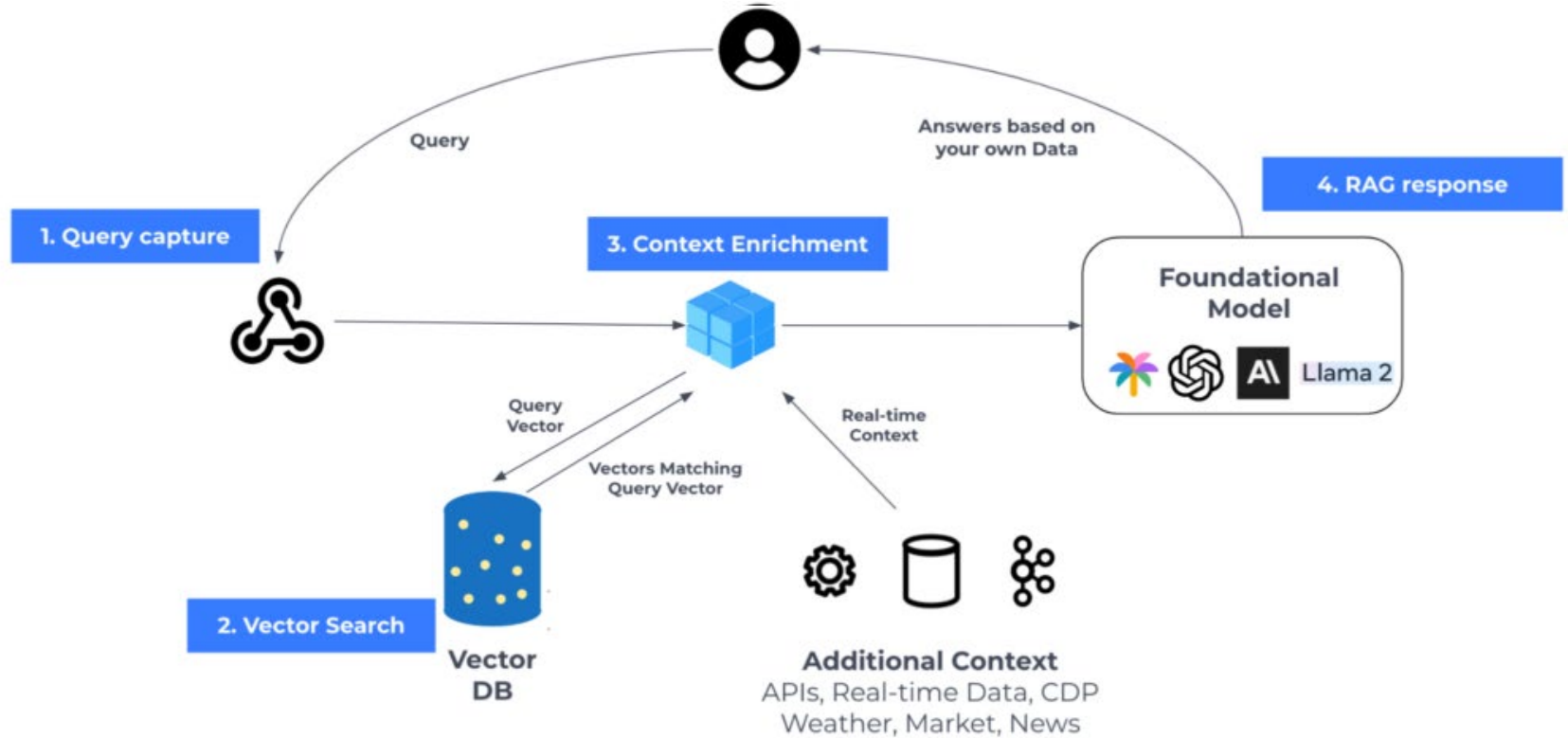
# Knowledge Graphs to augment LLMs



# Caching Mechanisms



# Hybrid Approaches



# References

- [1] Snorkel AI: Fine Tuning vs. RAG: <https://snorkel.ai/which-is-better-retrieval-augmentation-rag-or-fine-tuning-both/>
- [2] The Evolution of Retrieval Augmented Generation - [https://pattersonconsultingtn.com/blog/evolution\\_of\\_retrieval\\_augmented\\_generation\\_jan\\_2024.html](https://pattersonconsultingtn.com/blog/evolution_of_retrieval_augmented_generation_jan_2024.html)
- [3] Lil'Log: How to Build an Open-Domain Question Answering System? - <https://lilianweng.github.io/posts/2020-10-29-odqa/>
- [4] Looking to the future of vector databases - <https://medium.com/@infiniflowai/looking-to-the-future-of-vector-databases-be86f218dc94>
- [5] Embracing the future of data with Vector Databases - <https://jillanisofttech.medium.com/embracing-the-future-of-data-with-vector-databases-af4055a1f57a>
- [5] RAG Best Practices - <https://nexla.com/ai-infrastructure/retrieval-augmented-generation/>
- [6] Knowledge Graphs & LLMs: Multi-Hop Question Answering - <https://neo4j.com/developer-blog/knowledge-graphs-llms-multi-hop-question-answering/>
- [7] ChatGPT is not all you need - <https://www.wisecube.ai/blog/chatgpt-is-not-all-you-need/>
- [8] Cookbook for solving common problems in building GPT/LLM apps - <https://bootcamp.uxdesign.cc/cookbook-for-solving-common-problems-in-building-gpt-llm-apps-93fcdbe3f44a>

Thank You

AMA

**Arun Nandi**

Head of Data and Analytics at Unilever

Board Advisor ♦ Top 100 Leaders in AI ♦ Keynote Speaker

