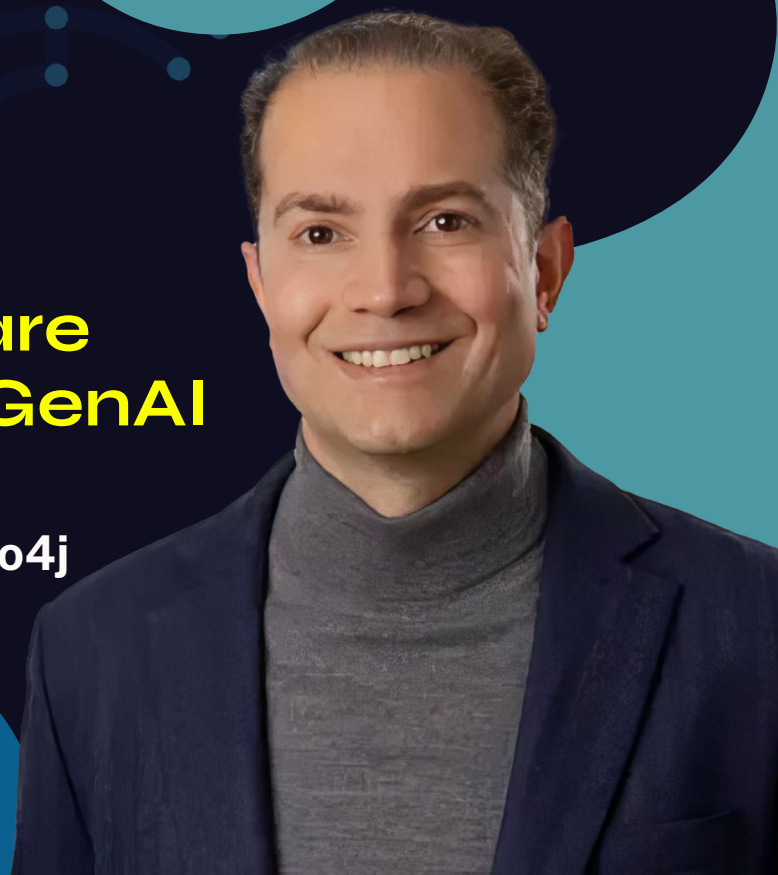
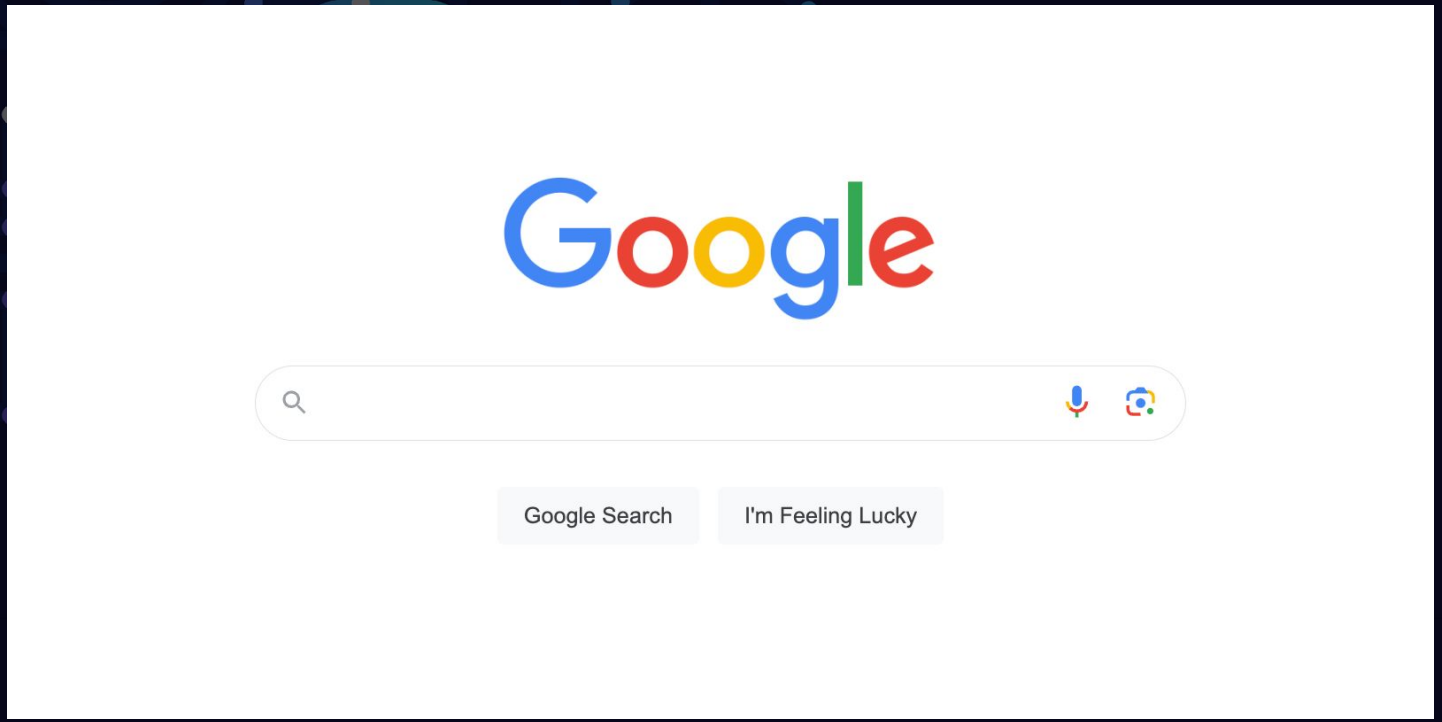
Why Knowledge Graphs are Necessary for Enterprise GenAI

Philip Rathle, Chief Technology Officer, Neo4j



The Evolution of... Web Search



 Neo4j Inc. All rights reserved 2024

How is GenAI like Web Search?

Web Search

1. Ask question
2. Technology operates on the q
3. Returns some text
4. Text = knowledge

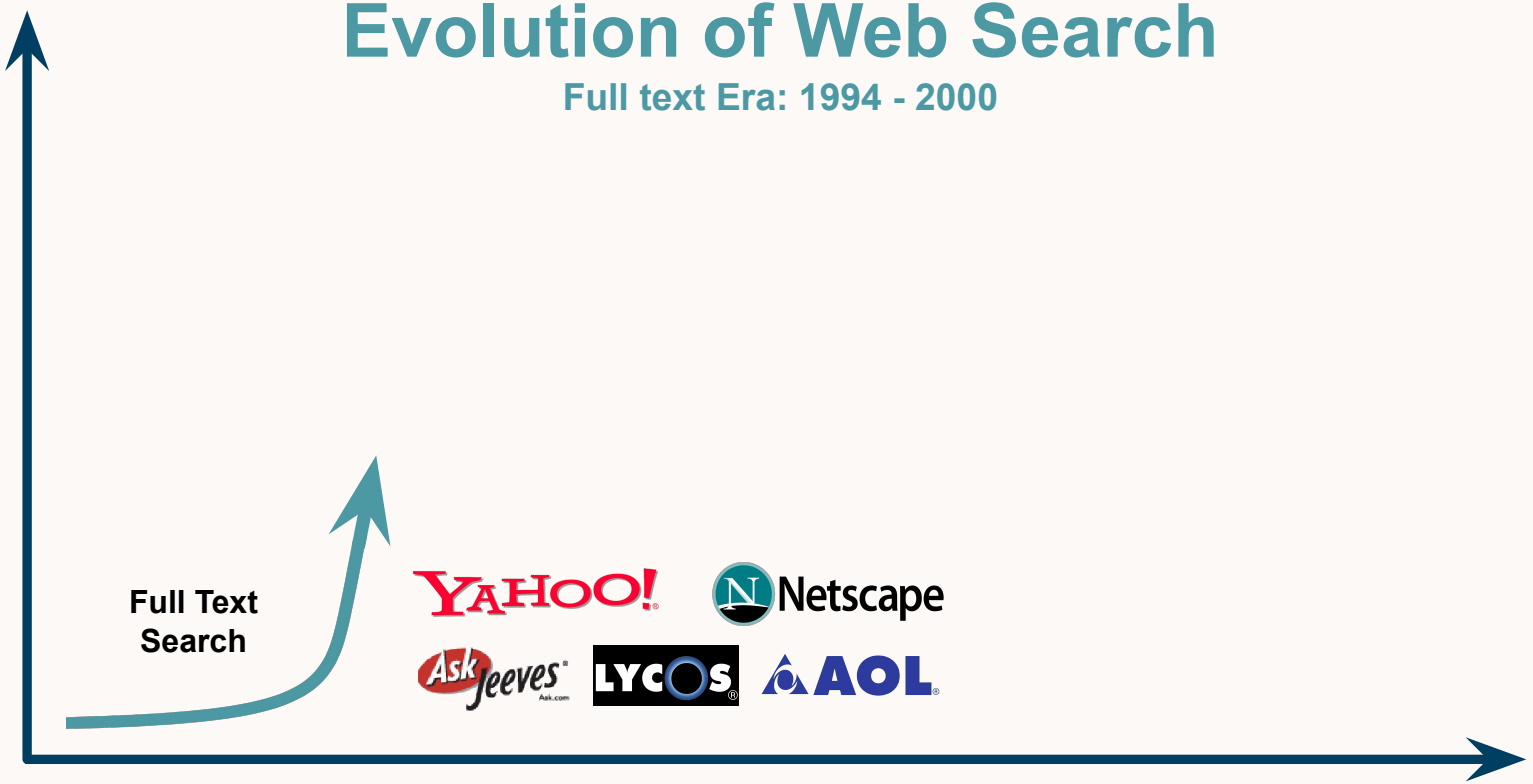
GenAI

1. Ask question
2. Technology operates on the q
3. Returns some text
4. Text = knowledge

Neo4j Inc. All rights reserved 2024

Evolution of Web Search

Full text Era: 1994 - 2000



 Neo4j Inc. All rights reserved 2024

Google Launches World's Largest Search Engine

Subscribe

Google Now Enables Internet Users to Search More Than 1 Billion URLs, Providing Quick and Easy Access to 560 Million Full-Text Indexed Web Pages and 500 Million Partially Indexed URLs

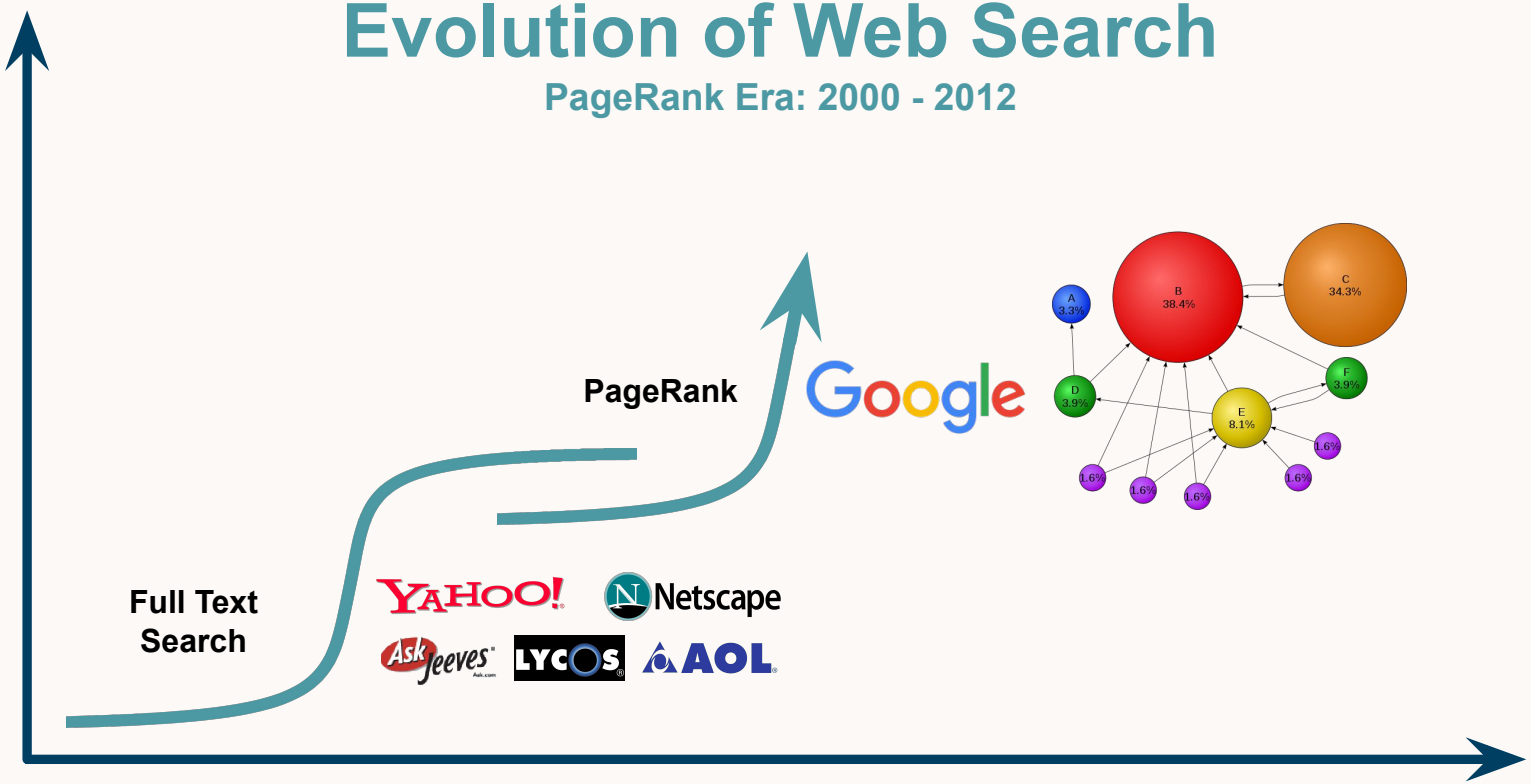
MOUNTAIN VIEW, Calif. – June 26, 2000 – Google Inc., one of the fastest growing search engines on the web, today announced it has released the largest search engine on the Internet. Google's new index, comprising more than 1 billion URLs, offers users the web's most comprehensive collection of websites, which can be easily searched with Google's fast and highly relevant search technology. Available now at www.google.com, Google's portal and destination site customers can also license this new index for integration with their own websites.

"Google is based on a variety of innovative technologies, including sophisticated text matching and its advanced, patent-pending technology called PageRank™, which ensures that the most important results always come up first."

Neo4j Inc. All rights reserved 2024

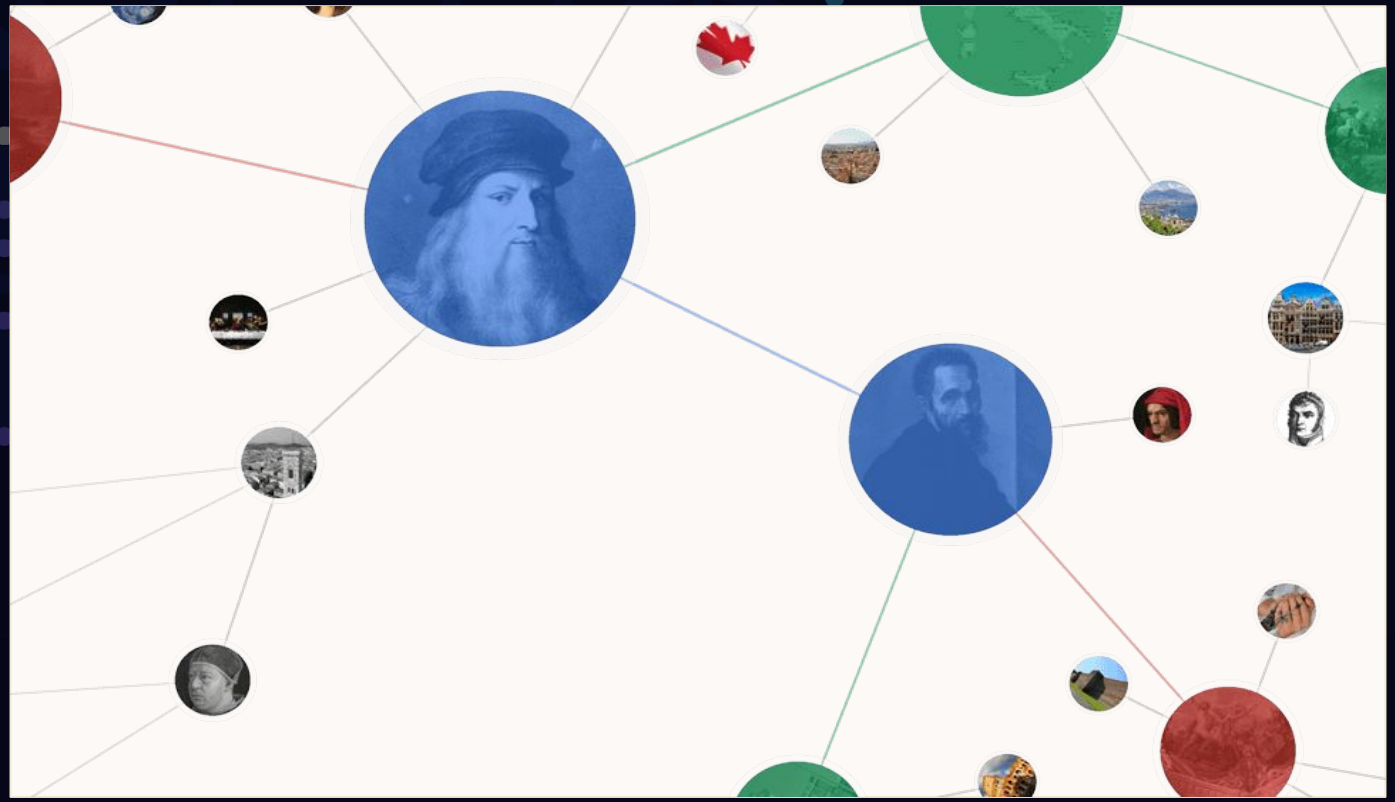
Evolution of Web Search

PageRank Era: 2000 - 2012

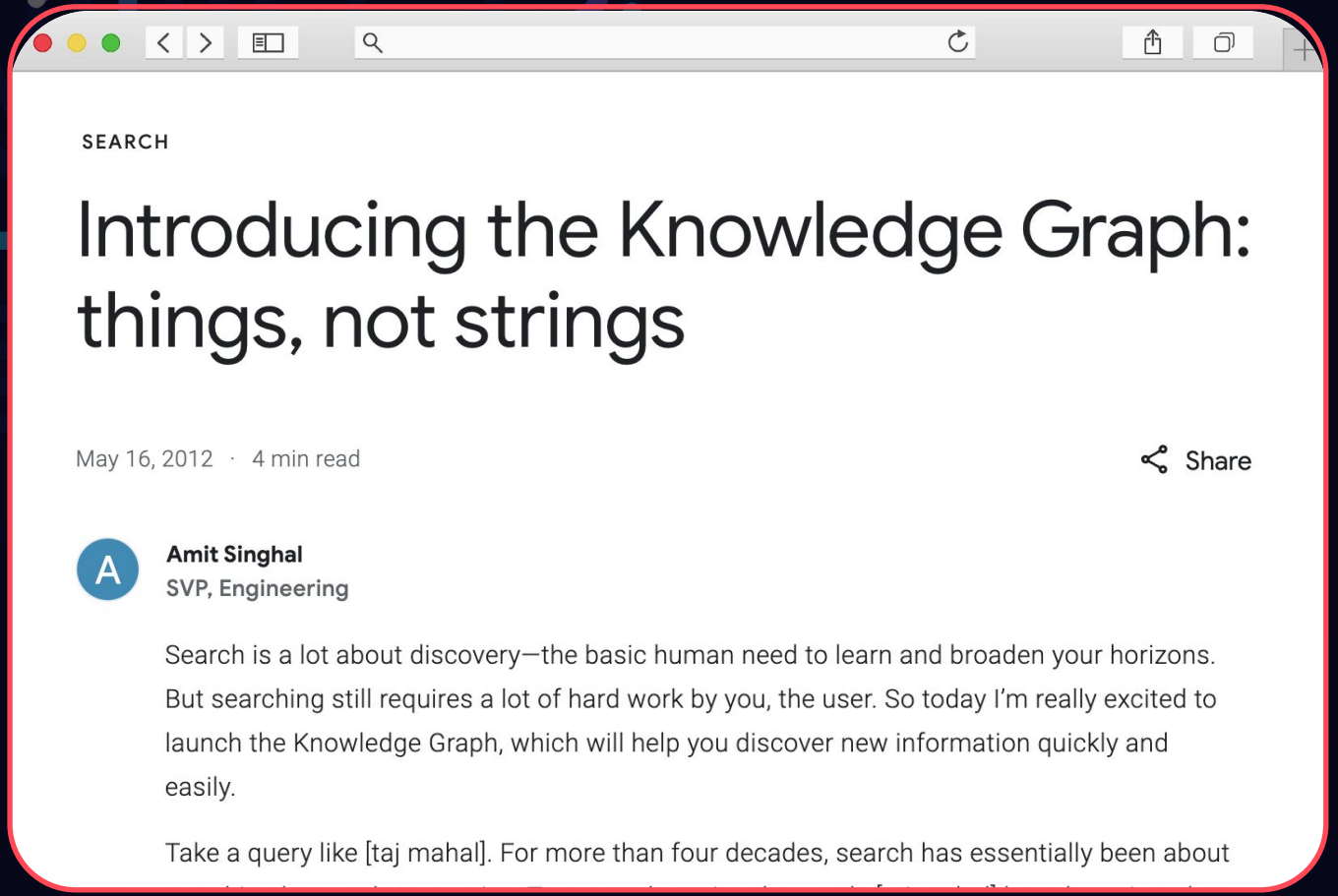


 Neo4j Inc. All rights reserved 2024

Google Knowledge Graph



 Neo4j Inc. All rights reserved 2024



SEARCH

Introducing the Knowledge Graph: things, not strings

May 16, 2012 · 4 min read

Share

Amit Singhal
SVP, Engineering

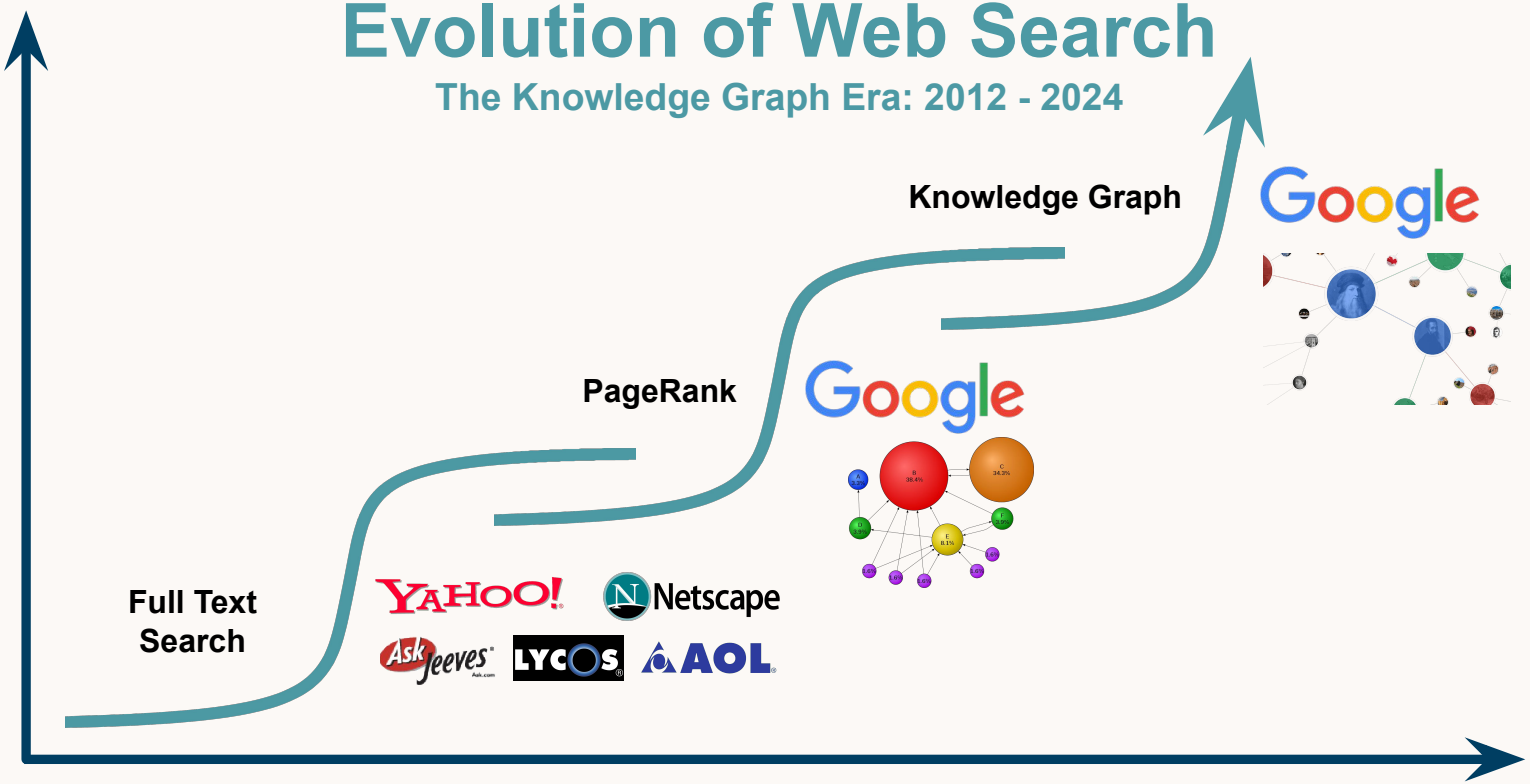
Search is a lot about discovery—the basic human need to learn and broaden your horizons. But searching still requires a lot of hard work by you, the user. So today I'm really excited to launch the Knowledge Graph, which will help you discover new information quickly and easily.

Take a query like [taj mahal]. For more than four decades, search has essentially been about

 Neo4j Inc. All rights reserved 2024

Evolution of Web Search

The Knowledge Graph Era: 2012 - 2024



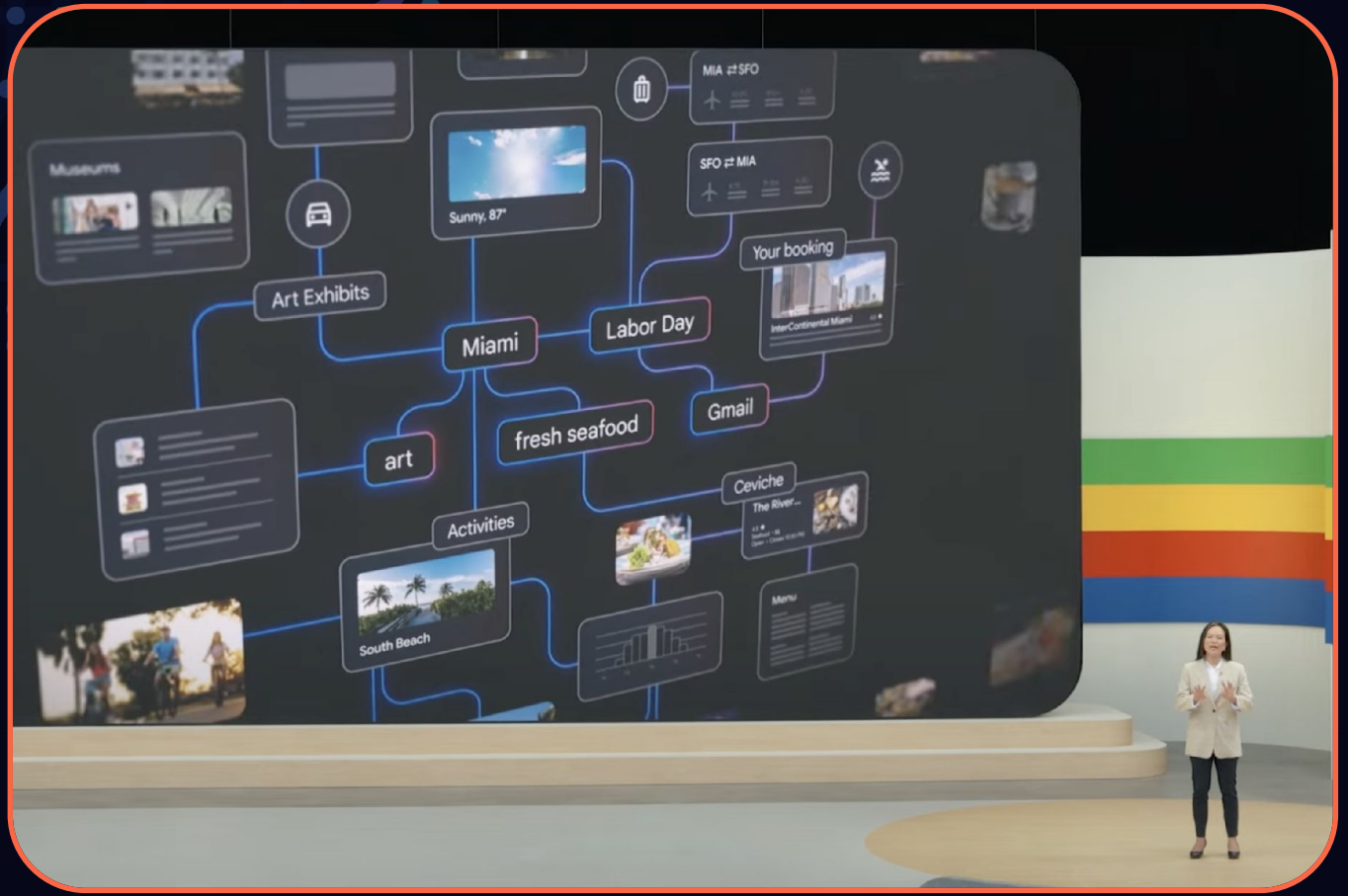
 Neo4j Inc. All rights reserved 2024

Google I/O ‘24 Keynote

 Neo4j Inc. All rights reserved 2024

GraphRAG @Google

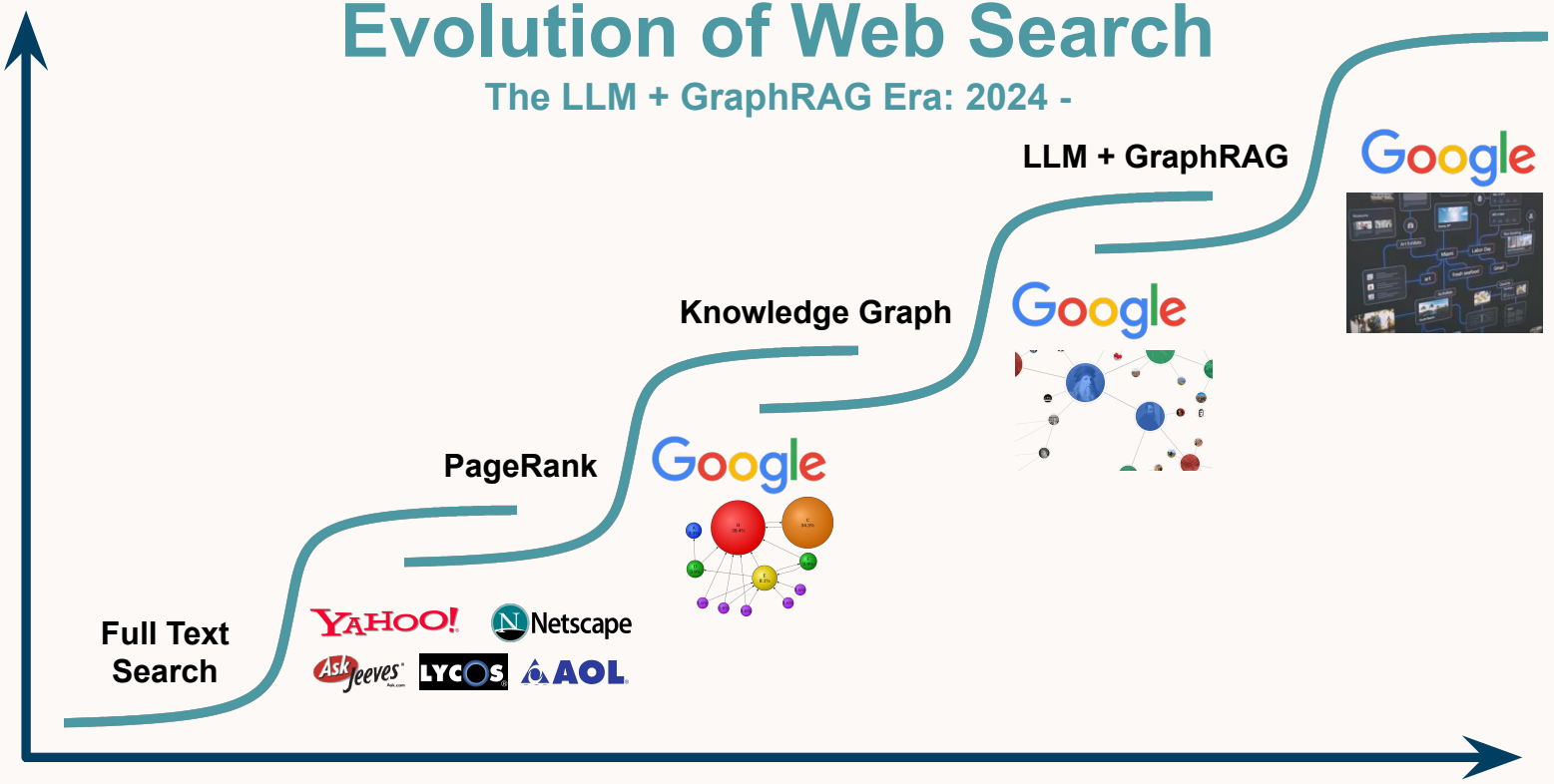
Google I/O ‘24 Keynote



Neo4j Inc. All rights reserved 2024

Evolution of Web Search

The LLM + GraphRAG Era: 2024 -

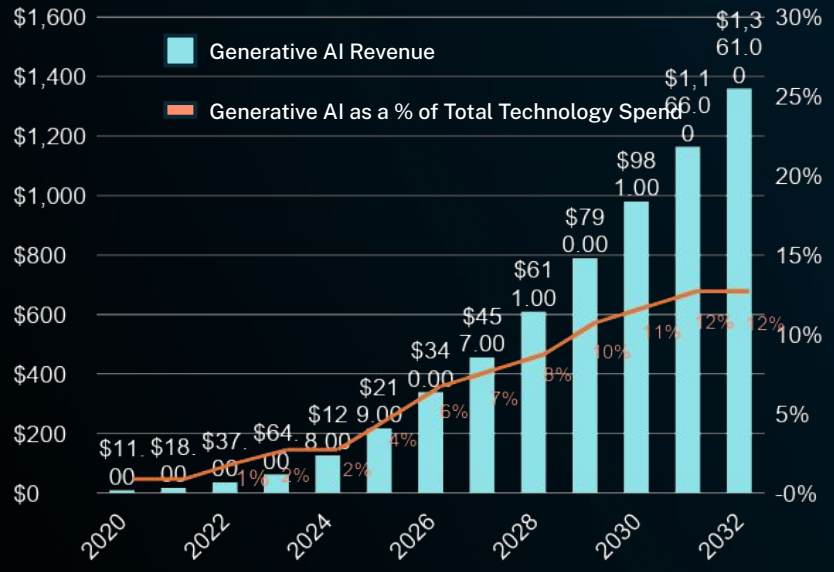


 Neo4j Inc. All rights reserved 2024

Back to GenAI...

Neo4j Inc. All rights reserved 2024

Generative AI races toward $1.3 trillion in revenue by 2032

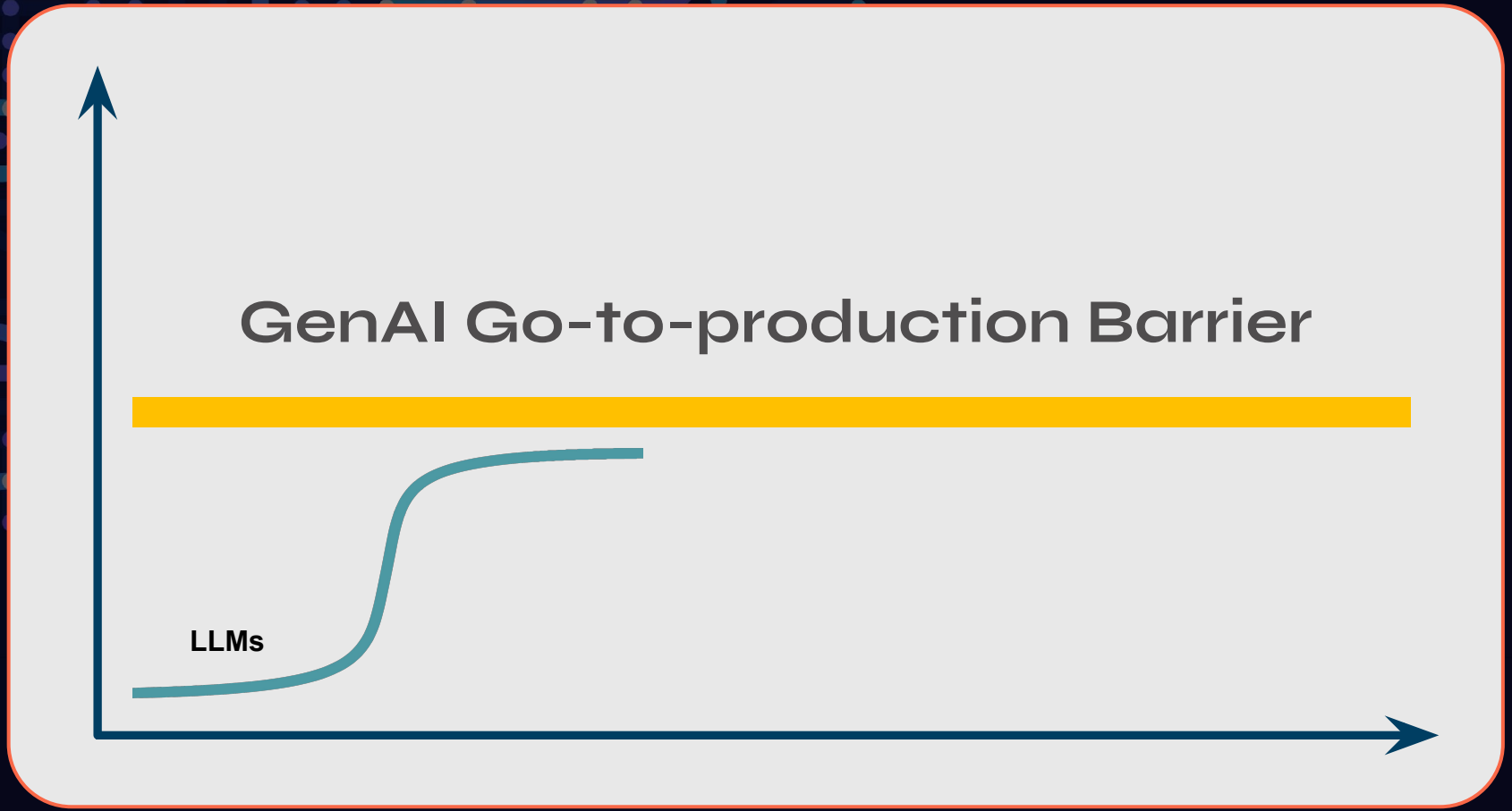


Bloomberg Intelligence

However, 71% of organizations are stuck piloting GenAI projects

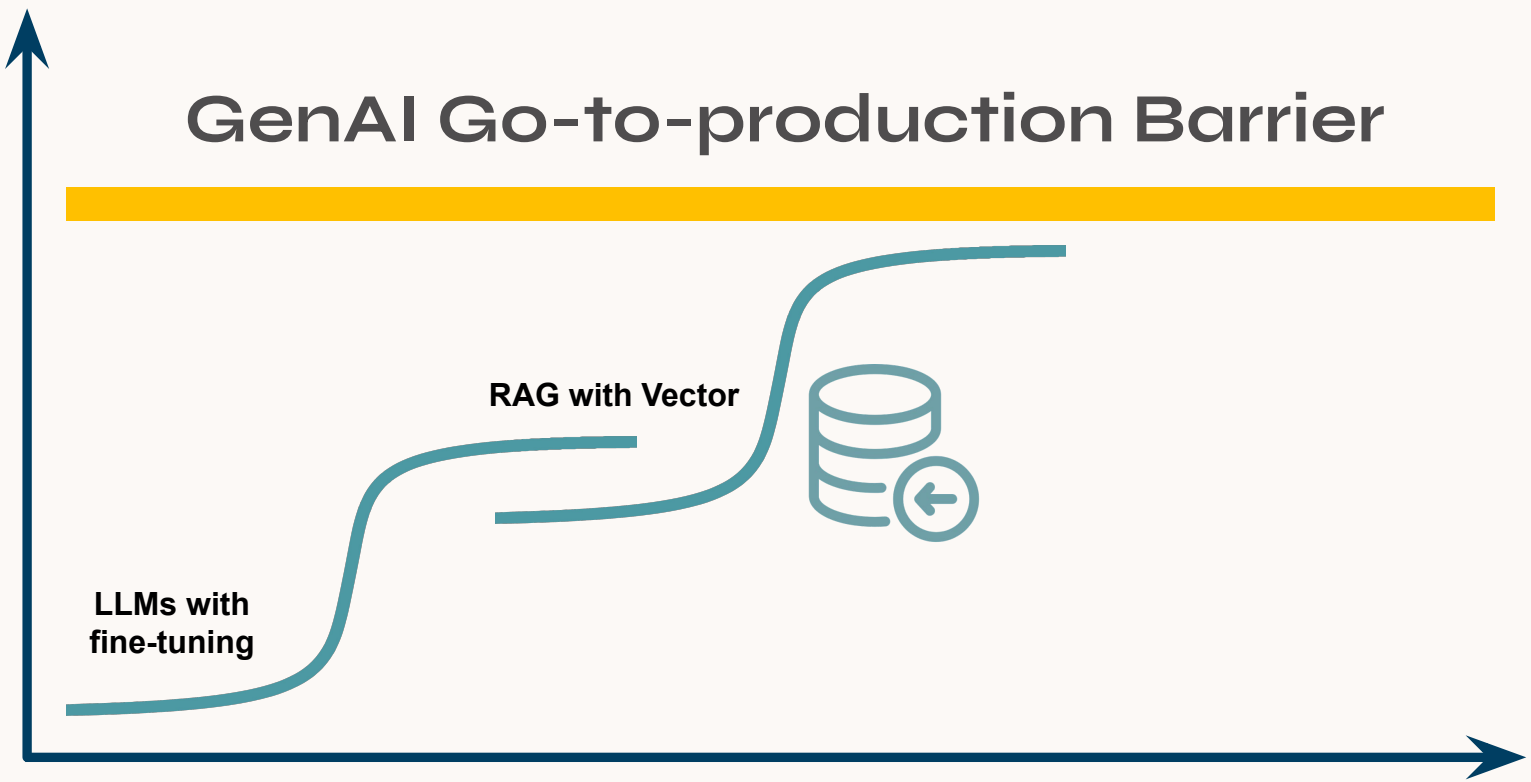
2024 IBM CEO Survey

IBM. All rights reserved 2024



 Neo4j Inc. All rights reserved 2024

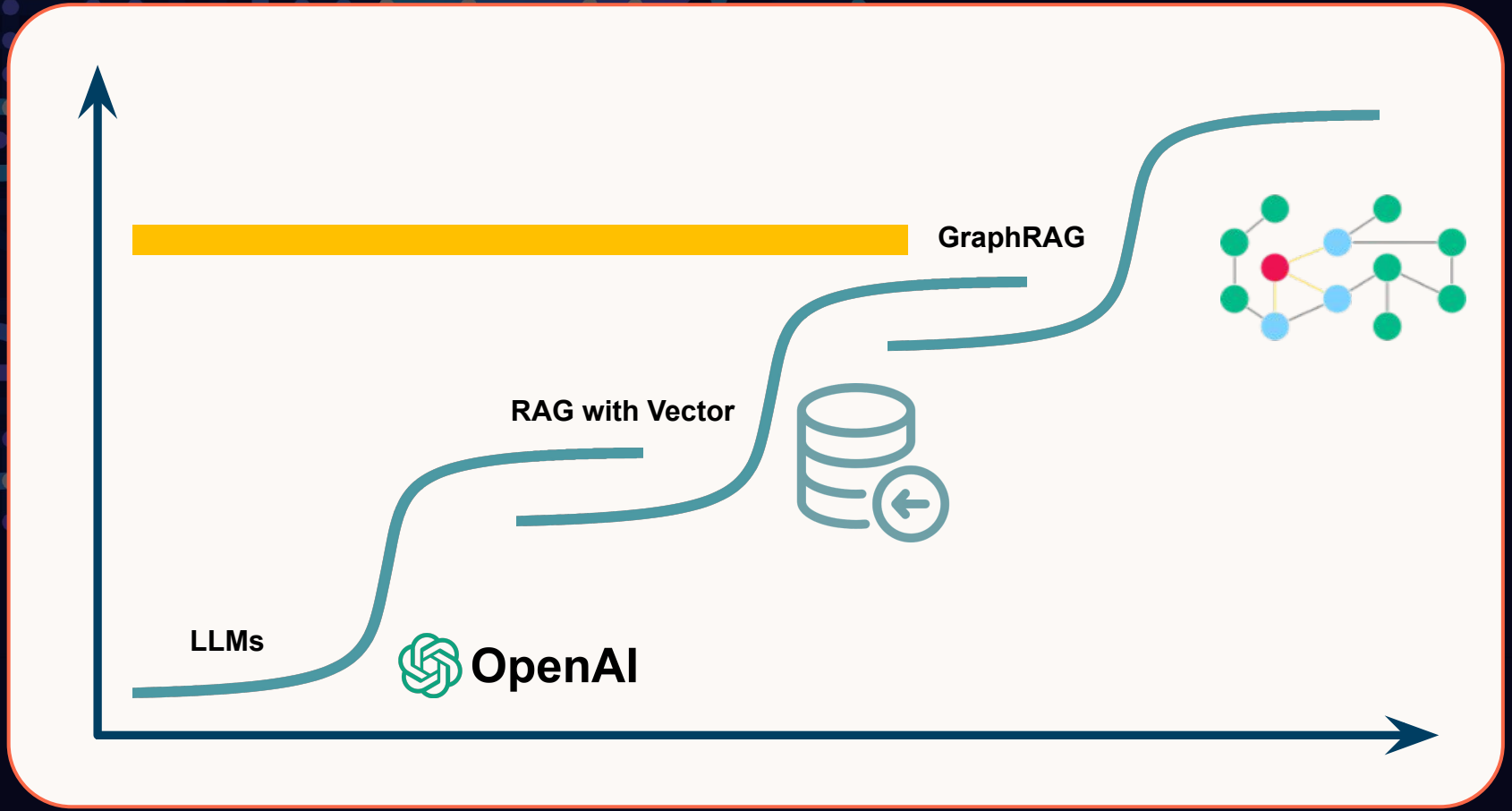
GenAI Go-to-production Barrier



 Neo4j Inc. All rights reserved 2024

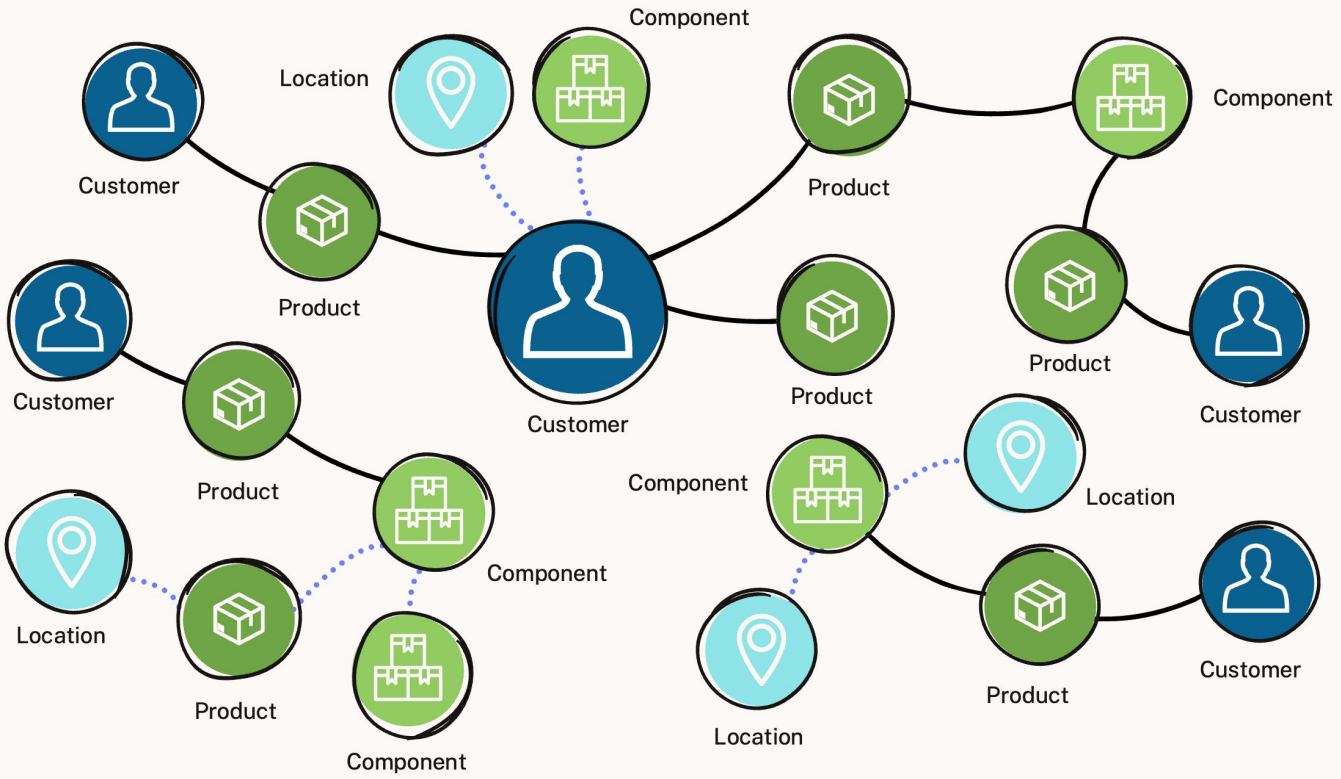
Enter GraphRAG

*GraphRAG is **RAG where the Retrieval path includes a Knowledge Graph**.*



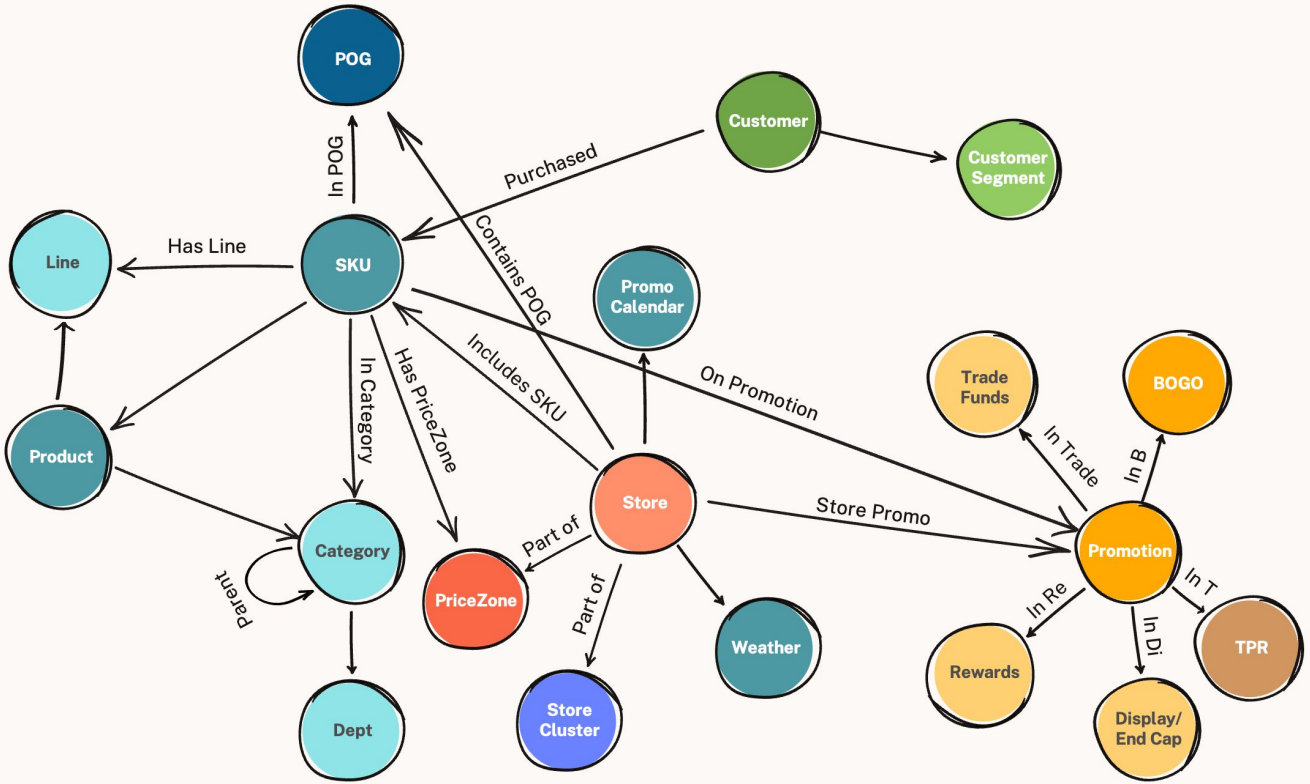
 Neo4j Inc. All rights reserved 2024

Commerce Transactions Graph



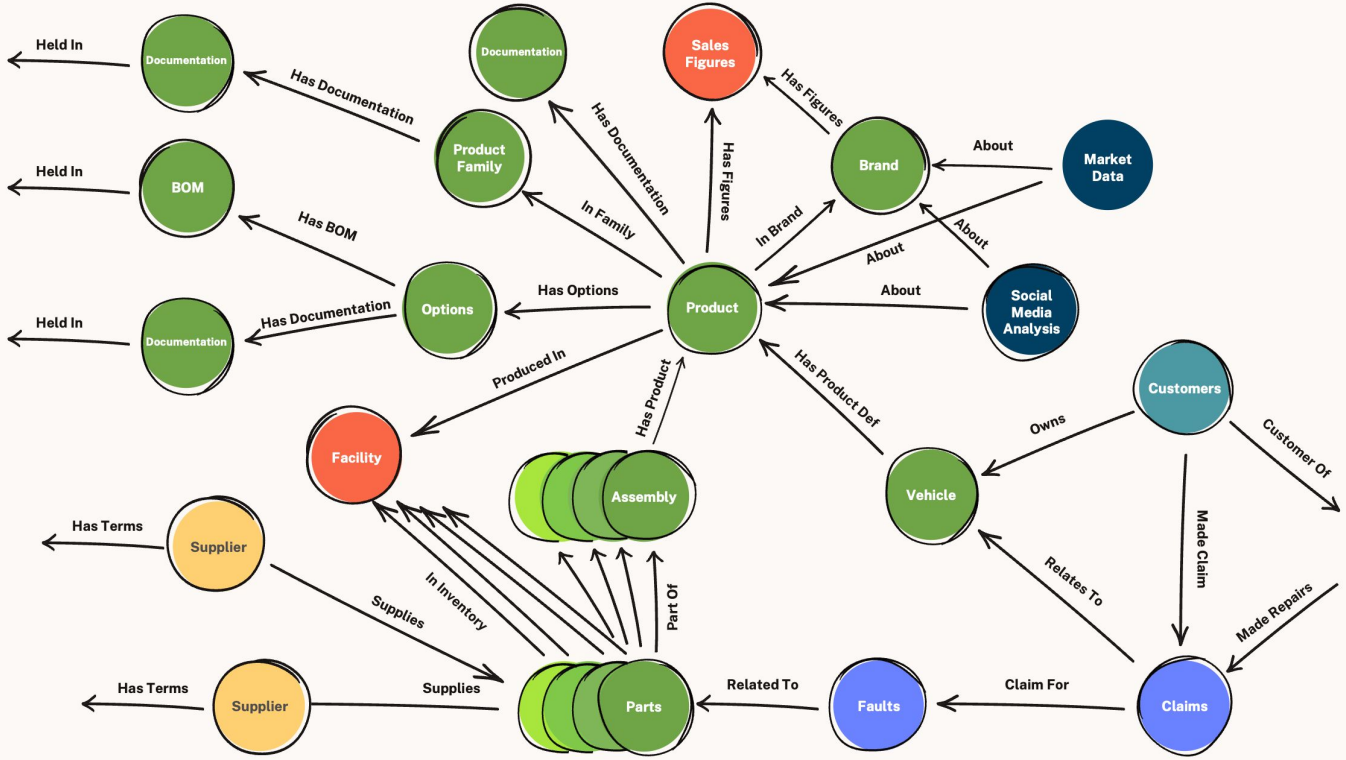
 Neo4j Inc. All rights reserved 2024

Product and Parts Graph



 Neo4j Inc. All rights reserved 2024

Digital Twin Graph

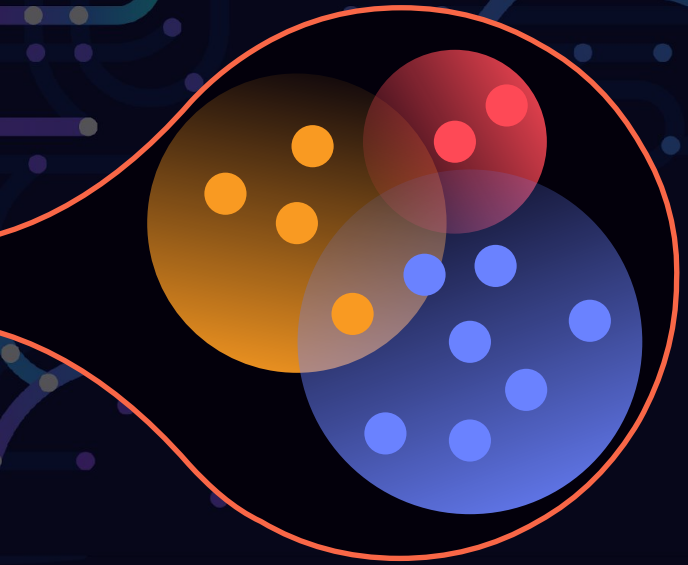


 Neo4j Inc. All rights reserved 2024

 Neo4j Inc. All rights reserved 2024

1. Better Answers

Vector-only RAG



Better accuracy though matching of concepts

Which asset managers are vulnerable to lithium shortage?

Based on the context provided, here is a summary of the asset managers that may be vulnerable to lithium storage:

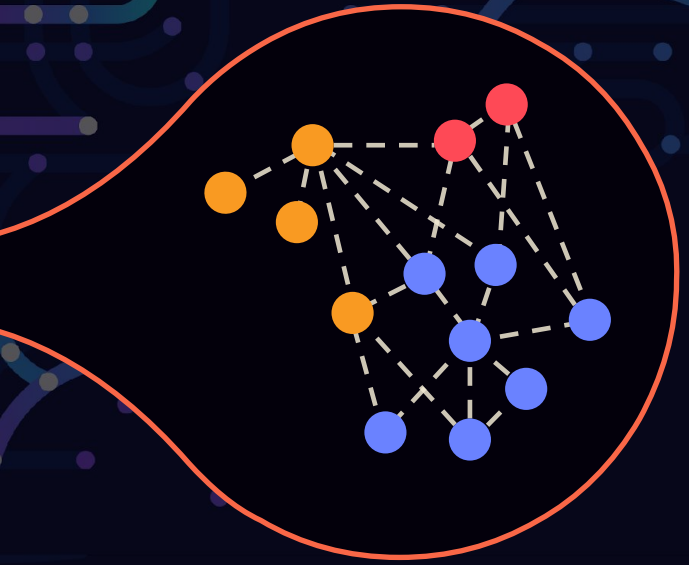
- The company that produces batteries and other products incorporating battery technology is vulnerable, as increased demand for batteries could place capacity constraints on their lithium supply chain.
- The company that sources steel and copper from Chile and Australia is vulnerable, as those countries are expected to experience severe weather due to climate change.
- The company that relies on suppliers for components and systems necessary for their products has some vulnerability. If there are lithium shortages it could result in supply constraints, cost pressures, or component shortages.
- The company uses lithium-ion batteries or other battery technology in their products is vulnerable to lithium shortages which could reduce battery supply, increase

Neo4j Inc. All rights reserved 2024

1. Better Answers

GraphRAG

Vector + Graph



More actionable and concrete answers

Which asset managers are vulnerable to lithium shortage?

Based on the context provided, the asset managers that may be vulnerable to lithium shortage are:

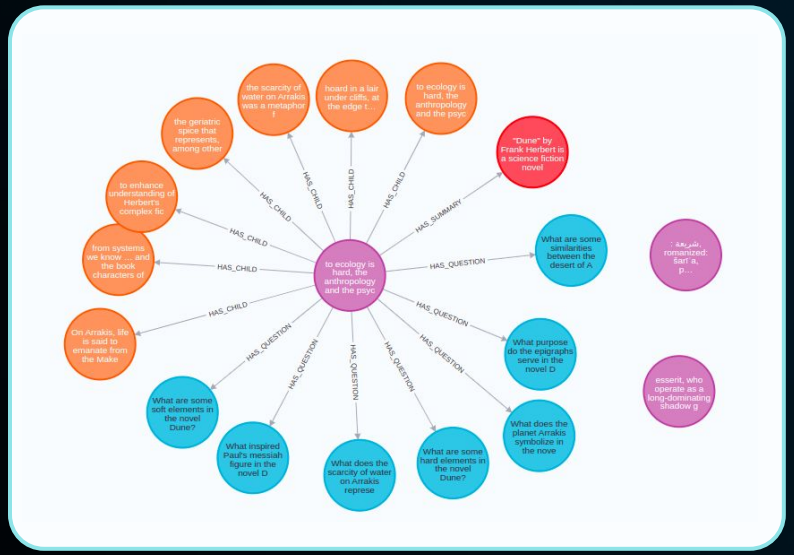
- Quantamental Technologies LLC
- DEPRINCE RACE & ZOLLO INC
- ATRIA Wealth Solutions, Inc.
- CITIGROUP INC
- CHARLES SCHWAB INVESTMENT MANAGEMENT INC
- ROYAL LONDON ASSET MANAGEMENT LTD
- STIFEL FINANCIAL CORP
- TUCKER ASSET MANAGEMENT, LLC
- Patriot Financial Group Insurance Agency, LLC
- SeaCrest Wealth Management, LLC

The context indicates that Stanley Black & Decker uses lithium batteries in many of its products. With increasing demand for lithium, there could be supply constraints and increased costs for lithium. This could negatively impact the profitability of Stanley Black & Decker. The asset managers listed above have holdings in Stanley Black & Decker, and therefore may be vulnerable to the risks of lithium shortage

Neo4j Inc. All rights reserved 2024

2. Easier Development

Transparent & Explainable

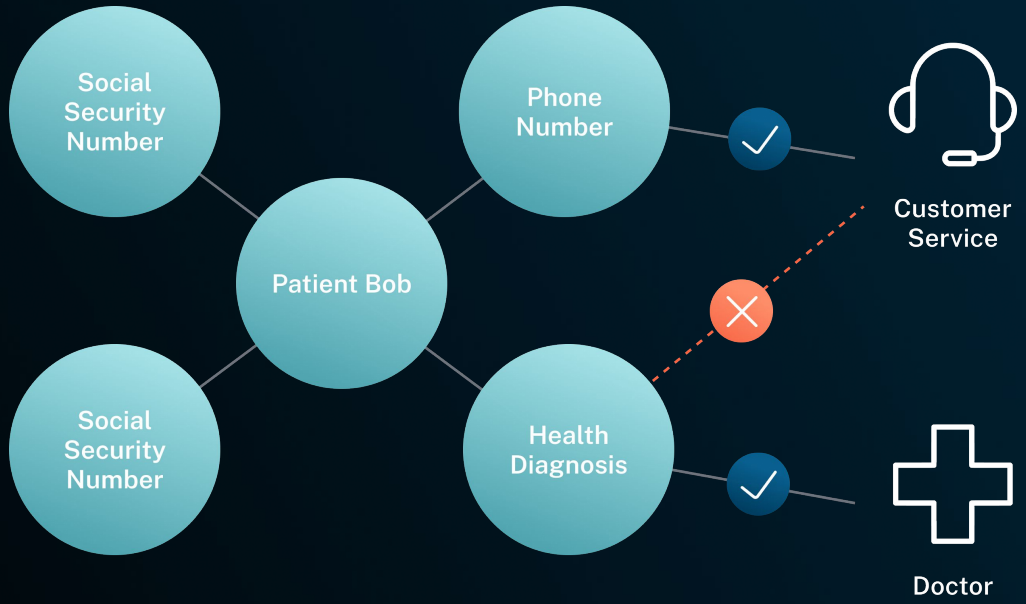


Opaque & Implicit

```
[ 3.9150659e-03  2.6659777e-03  1.0298982e-03 -2.7156321e-03
  1.9977870e-03  3.1204436e-03  1.2055682e-04  1.0450699e-03
 -6.4308796e-04  3.0822519e-03  2.1972554e-03  5.1480172e-05
 -3.7099270e-03  3.9439583e-03  6.8276987e-04  7.7137066e-04
  2.3698520e-03 -7.8547641e-04  6.0383842e-04  4.6370425e-03
 -1.6786088e-03  1.7417425e-03  2.4216413e-03  3.6545738e-03
 -1.9871239e-03  2.9489421e-03 -1.2810023e-03 -4.9174053e-04
 -3.9743204e-03 -2.7023794e-03 -3.0541950e-04 -1.5724347e-03
 -2.1029566e-03 -2.1624754e-03  2.1620055e-04 -1.4000515e-03
 -4.0824865e-03  4.6588355e-04  3.5028579e-03  4.8283348e-03
 -2.8737928e-03 -4.5569306e-03 -7.6568732e-04 -3.3311991e-03
  3.5790715e-03  4.2424244e-03  3.3478225e-03 -7.4140396e-04
  1.0030111e-03 -5.2394503e-04  5.8383477e-04 -4.8430995e-03
  2.6972082e-03 -4.8002079e-03 -2.3011414e-03  8.0388715e-04
  3.1952575e-05 -8.1621204e-04 -3.8127291e-03 -6.7428290e-04
 -1.7713077e-03 -3.0159748e-03  1.7178850e-03 -1.9258332e-03
 -2.4637436e-03  3.3779652e-03  2.7676420e-03  1.8853768e-03
 -2.4718521e-03 -1.9754141e-03  2.6104036e-03 -2.1335895e-03
  2.4405334e-03 -3.2013952e-04  3.9961869e-03  4.0419102e-03
  2.0586823e-03  4.9897884e-03  4.5599132e-03 -1.0976522e-03
  1.5563263e-03  3.9063310e-03 -2.9308300e-03 -4.8254002e-03
 -8.7642738e-06  3.9748671e-03  5.2895391e-04  6.3330121e-04
 -1.2614765e-03 -8.5018738e-04  3.7659388e-03  3.0237564e-03
```

AI Engineer "I actually already fixed a couple of bugs thanks to this!"

3. Explainability & Governance



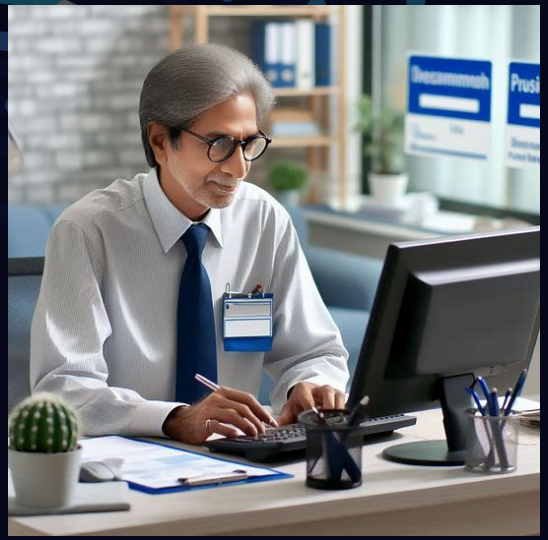
When are Knowledge Graphs *Necessary* for GenAI?

Neo4j Inc. All rights reserved 2024

Higher Stakes = Higher Bar

Low

Business peripheral
No single right answer
Leeway for creativity



Medium

Involves core business
Mitigations in place
(e.g. human in the loop)

High

Wrong answers result in lost $$/ reputation/ fines/ etc.

Neo4j Inc. All rights reserved 2024

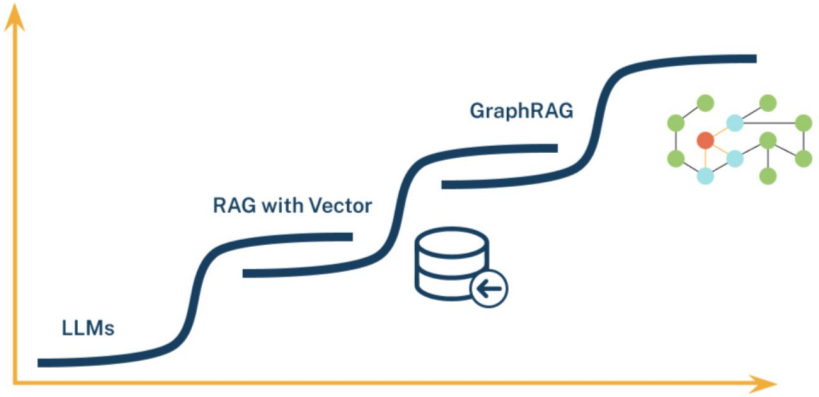
neo4j

The GraphRAG Manifesto: Adding Knowledge to GenAI

Philip Rathle, Chief Technology Officer, Neo4j

Jul 11 · 22 mins read

We're Entering the "Blue Links" Era of RAG



 Neo4j Inc. All rights reserved 2024

Better together

 Neo4j Inc. All rights reserved 2024

neo4j

Thank you

Join Alison Cossette tomorrow @ 2:30 PM
For Getting Started with GraphRAG

DeepLearning.AI

https://www.deeplearning.ai/short-courses/knowledge-graphs-rag/

Neo4j Inc. All rights reserved 2024

neo4j | GraphAcademy

GraphAcademy — Free, Self-Paced, Hands-on Online Training

https://graphacademy.neo4j.com/

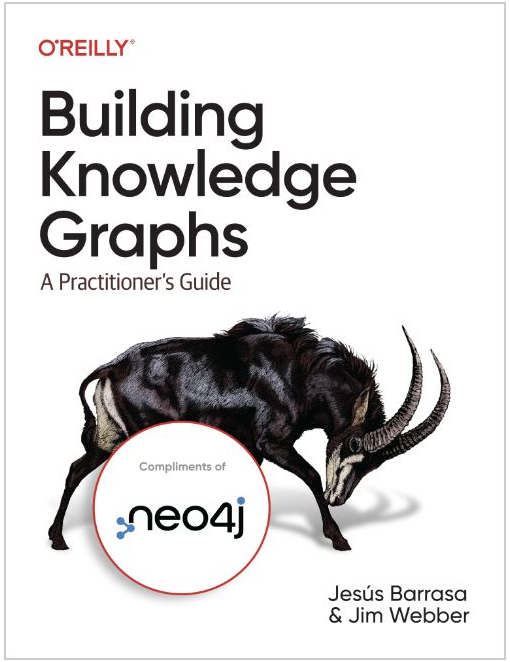
 Neo4j Inc. All rights reserved 2024

Building Knowledge Graphs: A Practitioner's Guide

By Jesús Barrasa & Jim Webber

Publisher: O'Reilly

Available Formats: PDF - EN US



Free e-book!
https://neo4j.com/knowledge-graphs-practitioners-guide/

Neo4j Inc. All rights reserved 2024

What's Next?

Neo4j Inc. All rights reserved 2024

GraphRAG Retrieval Patterns

The core GraphRAG pattern is very simple, yet powerful.

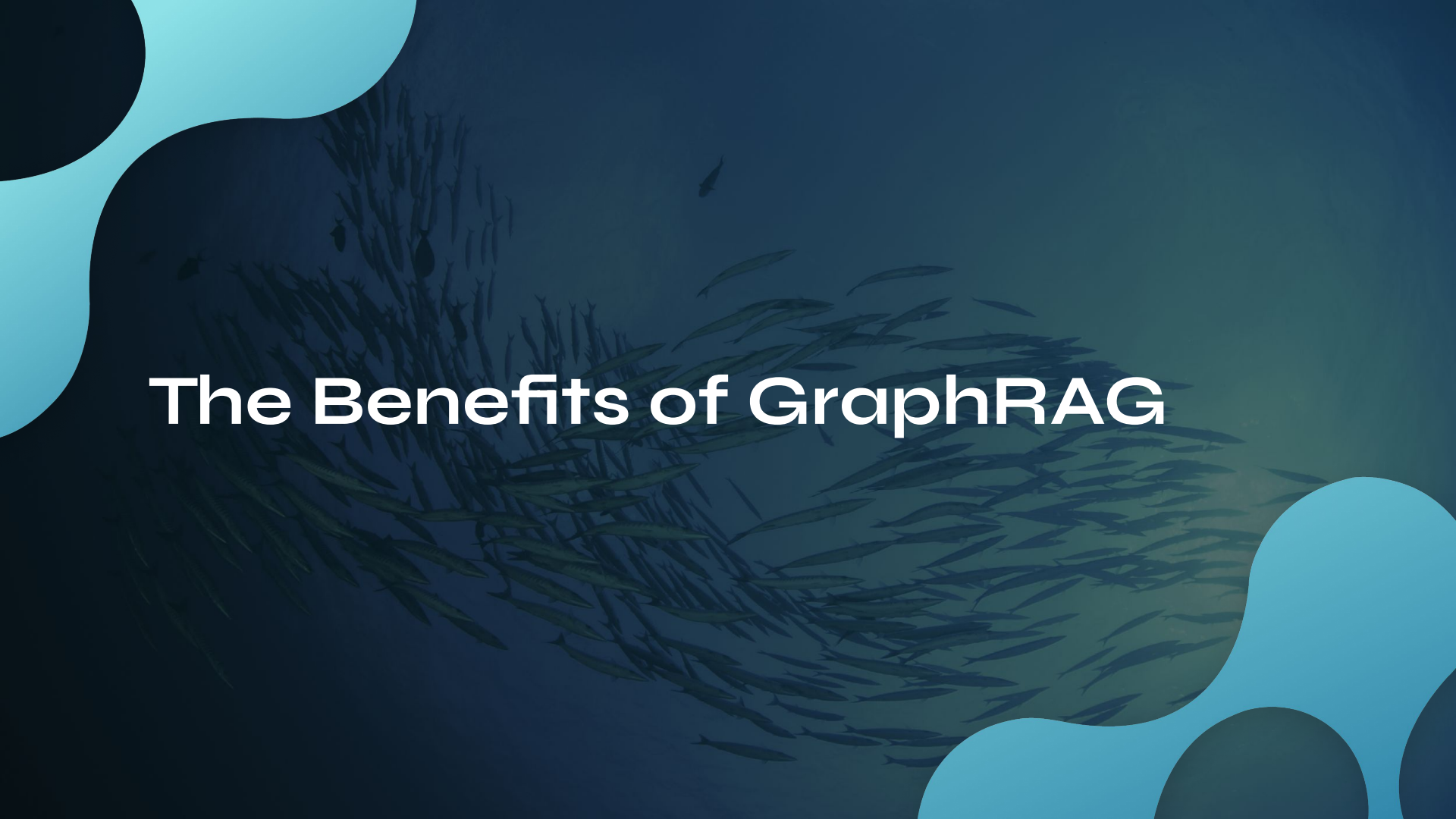
1. Do a vector search to find an initial set of nodes

2. Traverse the graph around those nodes to add context

Optional: Rank the results using the graph and pass the top-k documents to the LLM

The Benefits of GraphRAG

The Benefits of GraphRAG



1. Higher Accuracy

3X

A Knowledge Graph improves the accuracy of LLM responses by 54.2%, an average of 3x

A BENCHMARK TO UNDERSTAND THE ROLE OF KNOWLEDGE GRAPHS ON LARGE LANGUAGE MODEL'S ACCURACY FOR QUESTION ANSWERING ON ENTERPRISE SQL DATABASES

TECHNICAL REPORT

Juan F. Sequeda
data.world
juan@data.world

Dean Allemang
data.world
dean.allemang@data.world

Bryon Jacob
data.world
bryon@data.world

November 14, 2023

ABSTRACT

Enterprise applications of Large Language Models (LLMs) hold promise for question answering on enterprise SQL databases. However, the extent to which LLMs can accurately respond to enterprise questions in such databases remains unclear, given the absence of suitable Text-to-SQL benchmarks tailored to enterprise settings. Additionally, the potential of Knowledge Graphs (KGs) to enhance LLM-based question answering by providing business context is not well understood. This study aims to evaluate the accuracy of LLM-powered question answering systems in the context of enterprise questions and SQL databases, while also exploring the role of knowledge graphs in improving accuracy. To achieve this, we introduce a benchmark comprising an enterprise SQL schema in the insurance domain, a range of enterprise queries encompassing reporting to metrics, and a contextual layer incorporating an ontology and mappings that define a knowledge graph. Our primary finding reveals that question answering using GPT-4, with zero-shot prompts directly on SQL databases, achieves an accuracy of 16%. Notably, this accuracy increases to 54% when questions are posed over a Knowledge Graph representation of the enterprise SQL database. Therefore, investing in Knowledge Graph provides higher accuracy for LLM powered question answering systems.

Keywords Knowledge Graphs · Large Language Models · Question Answering · SQL Databases · Benchmark · Retrieval Augmented Generation (RAG)

1 Introduction

Question answering, the ability to interact with data using natural language questions and obtaining accurate results, has been a long-standing challenge in computer science dating back to the 1960s[8, 9, 20, 10]. The field has advanced throughout the past decades [3, 22, 17], through Text-to-SQL approaches, as a means of facilitating chatting with the data that is stored in SQL databases[13, 12, 16, 1, 6, 19]. With the rise of Generaiive AI and Large Language Models (LLMs), the interest continues to increase. These question answering systems hold tremendous potential for transforming the way data-driven decision making is executed within enterprises.

While question answering systems have shown remarkable performance in several Text-to-SQL benchmarks [4, 2], such as Spider [18], WikiSQL[19], KaggleDBQA[11] their implications relating to enterprise SQL databases remain relatively obscure. [1] We argue that existing Question Answering and Text-to-SQL benchmarks, although valuable, are often misaligned with real-world enterprise settings:

1. these benchmarks typically overlook complex database schemas representing enterprise domains, which likel… comprise hundreds of tables.

[1] A full survey on existing Text-to-SQL benchmarks is ouside the scope of this work. We do believe… beneficial for the community.

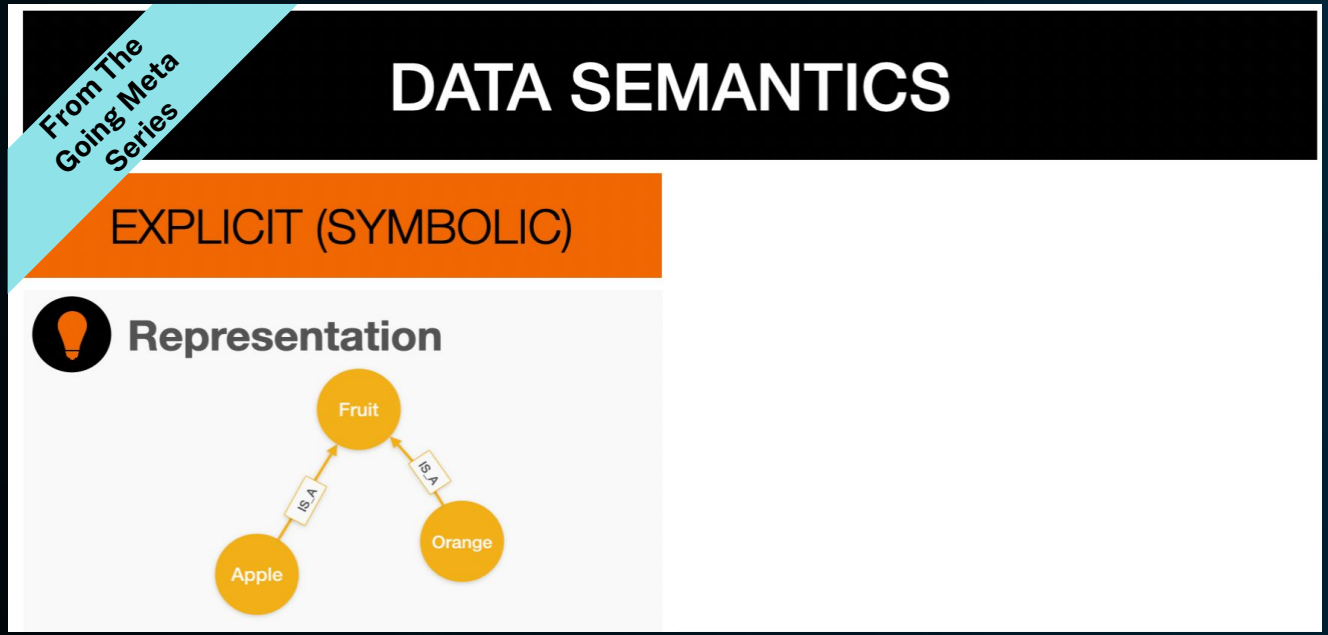
2. Easier Development

Feedback from an AI Engineer

09:05 ***Redacted*** Ciao team exciting news from our caching infrastructure! I've been experimenting with `Neo4J` as a candidate for our cache storage and the results are very promising! Here is the PR with the changes

2. Easier Development: Why

Natural language description: “Apples & oranges are both fruits”



2. Easier Development: Why

Natural language description: “Apples & oranges are both fruits”

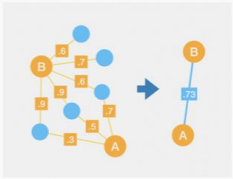
From The Going Meta Series

DATA SEMANTICS

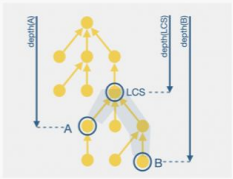
EXPLICIT (SYMBOLIC)

Similarity calculation

Structural
- Node similarity
- Overlap
- Jaccard

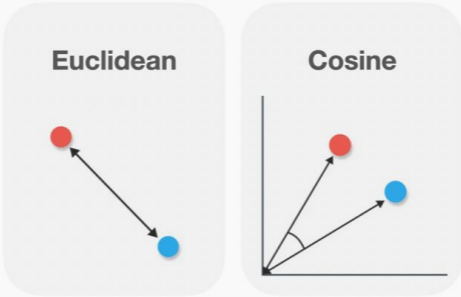


Taxonomy based
- Path
- Leacock-Chodorow
- Wu-Palmer



IMPLICIT (SUB-SYMBOLIC)

Similarity calculation



Vector **distance** based

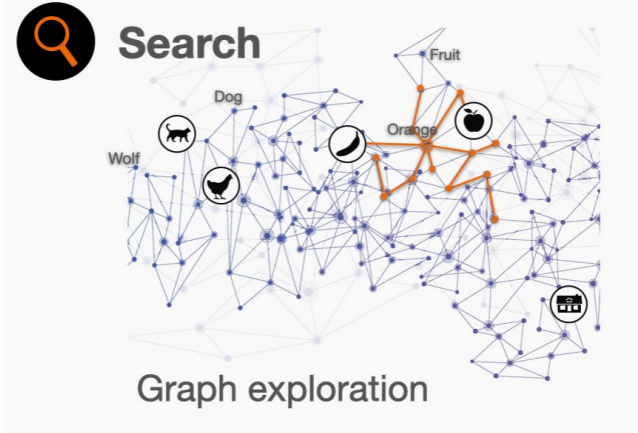
2. Easier Development: Why

Natural language description: “Apples & oranges are both fruits”

From The Going Meta Series

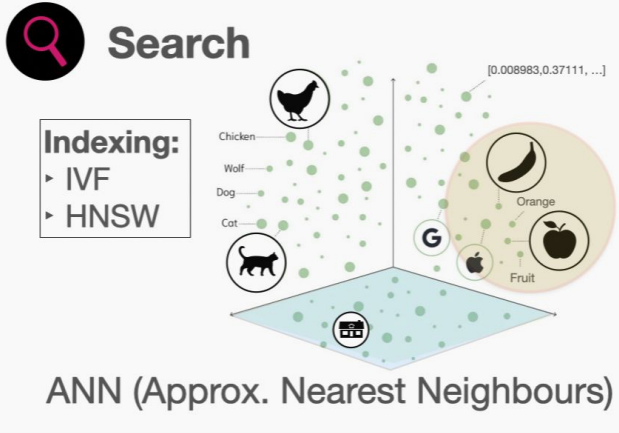
DATA SEMANTICS

EXPLICIT (SYMBOLIC)



Graph exploration

IMPLICIT (SUB-SYMBOLIC)

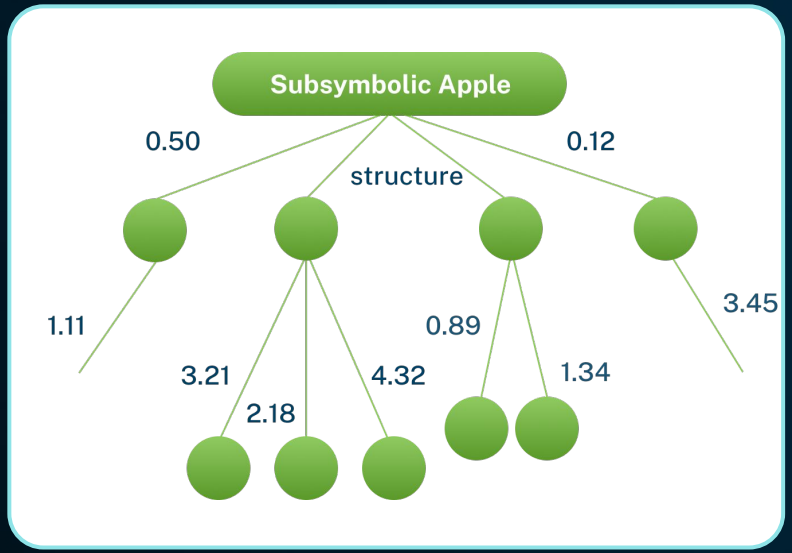


ANN (Approx. Nearest Neighbours)

2. Easier Development: Why

Human View of an Apple

LLM + Vector View of an Apple



Detour! Knowledge Graph Creation

Unstructured data

Typically PDFs or other text documents

Watch This Space!

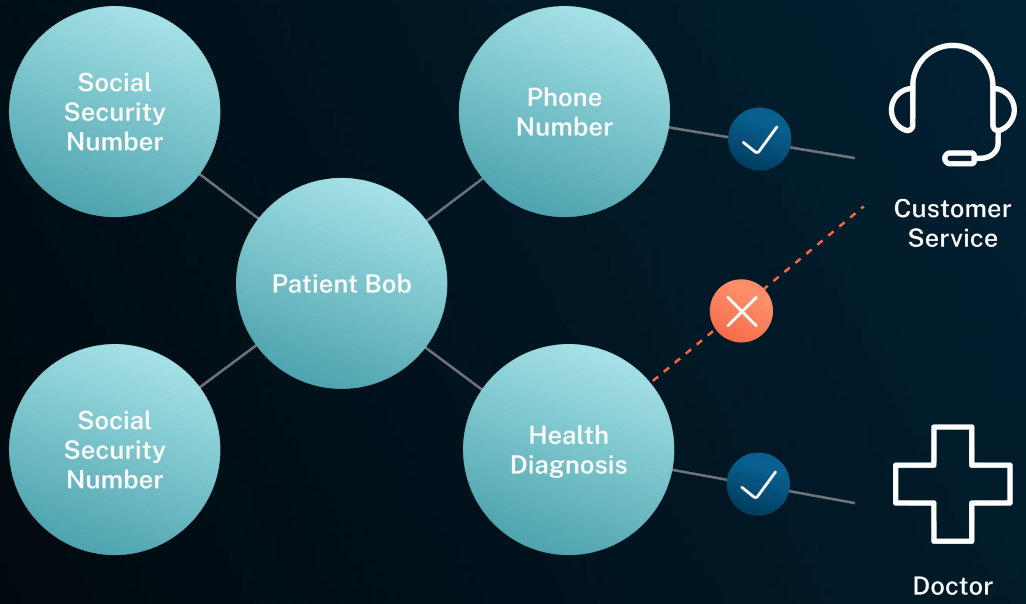
Mix of structured & unstructured data

Structured data with long-form text

Structured data

Structured data with short text values

3. Explainability & Governance



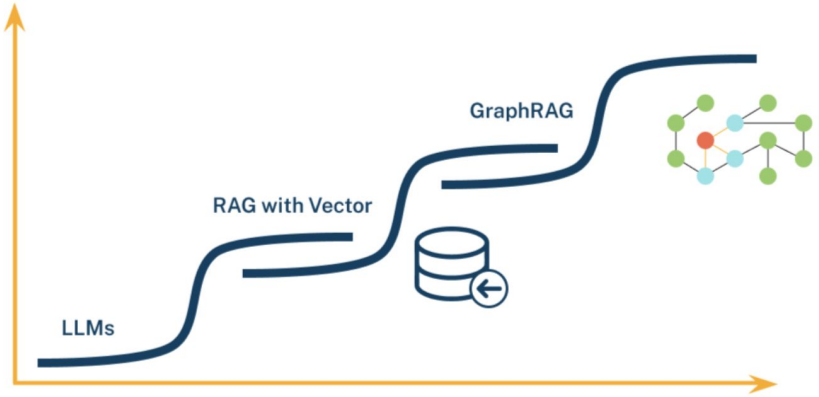
neo4j

The GraphRAG Manifesto: Adding Knowledge to GenAI

Philip Rathle, Chief Technology Officer, Neo4j

Jul 11 · 22 mins read

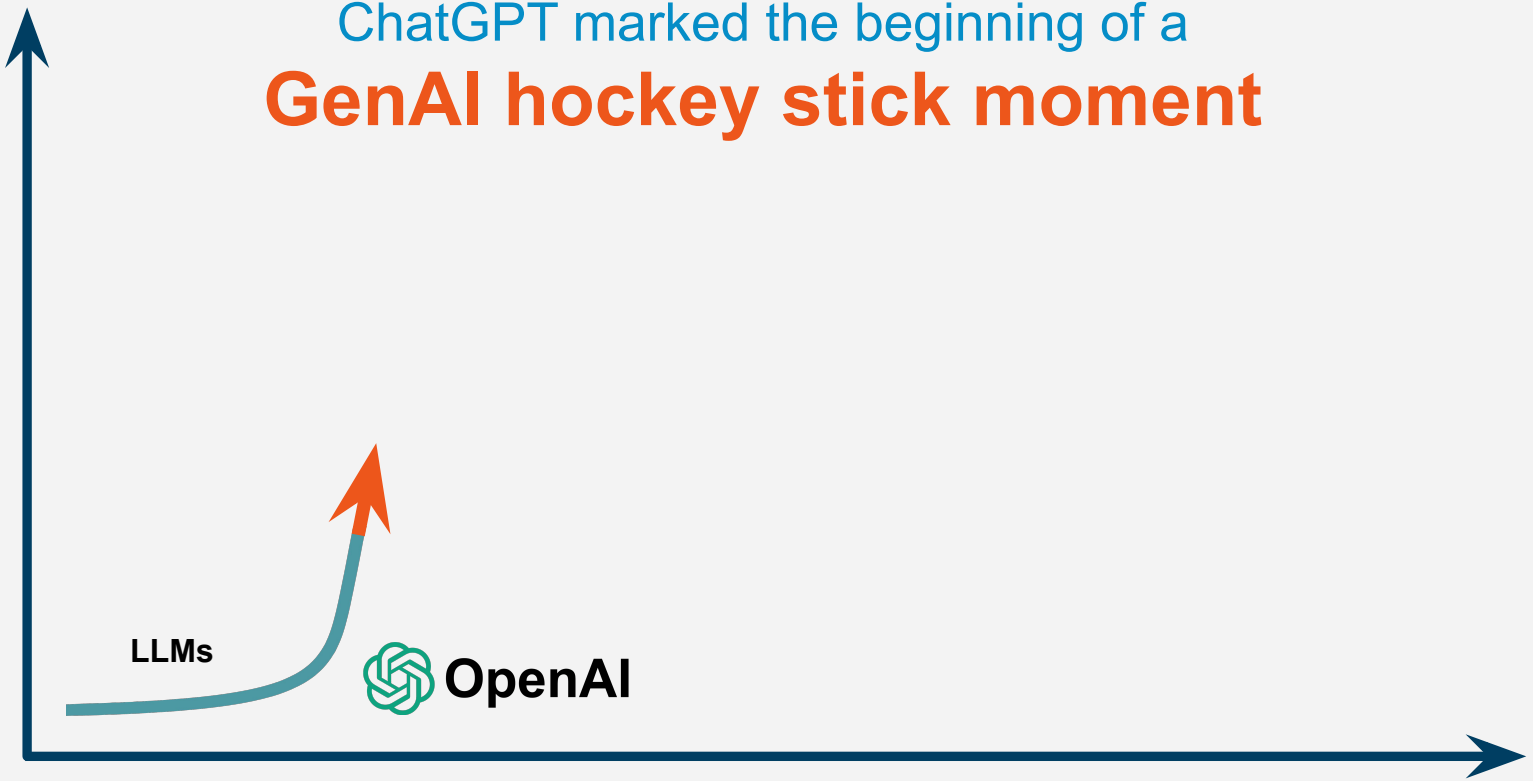
We're Entering the "Blue Links" Era of RAG

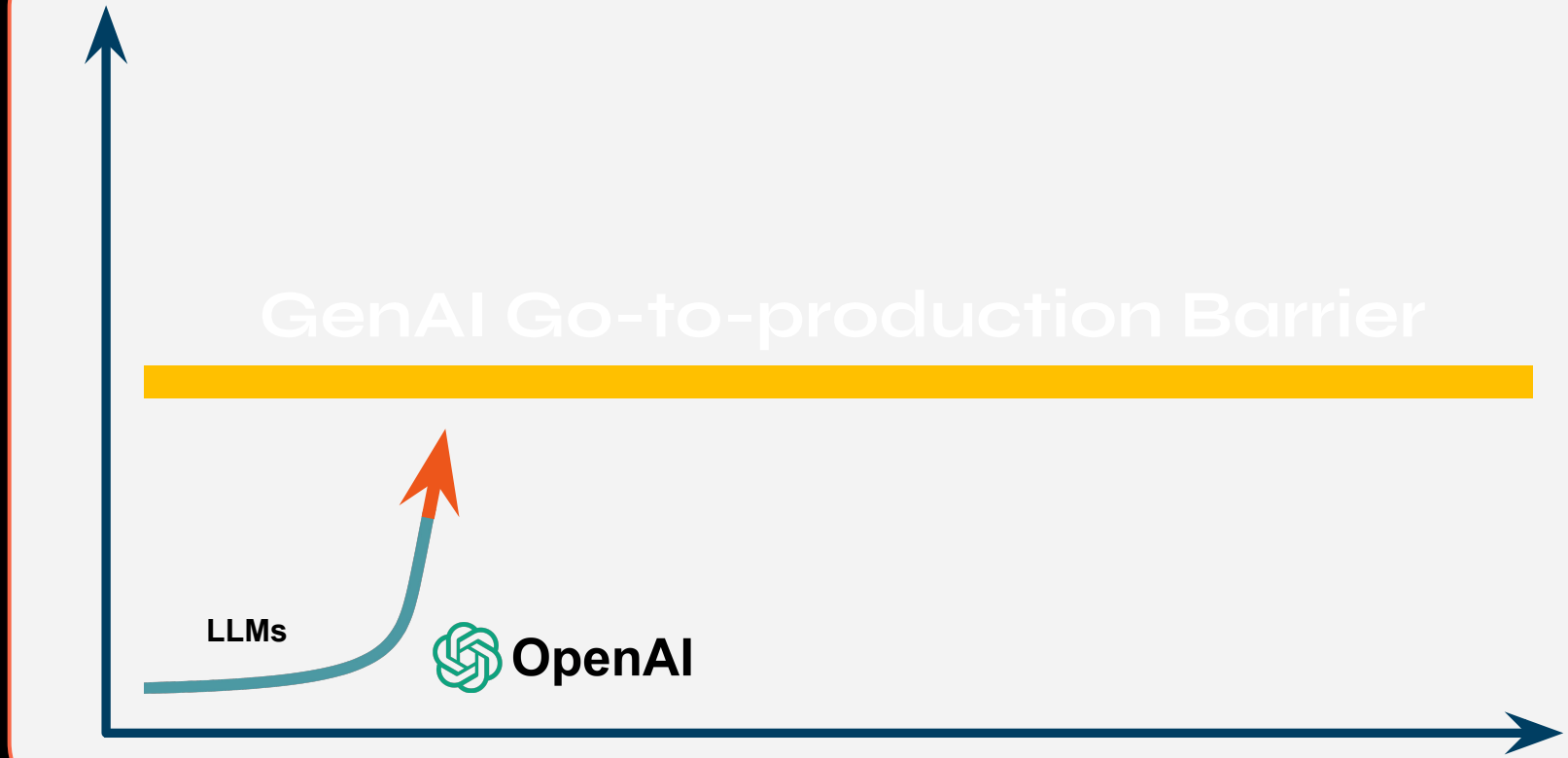


 Neo4j Inc. All rights reserved 2024

ChatGPT marked the beginning of a

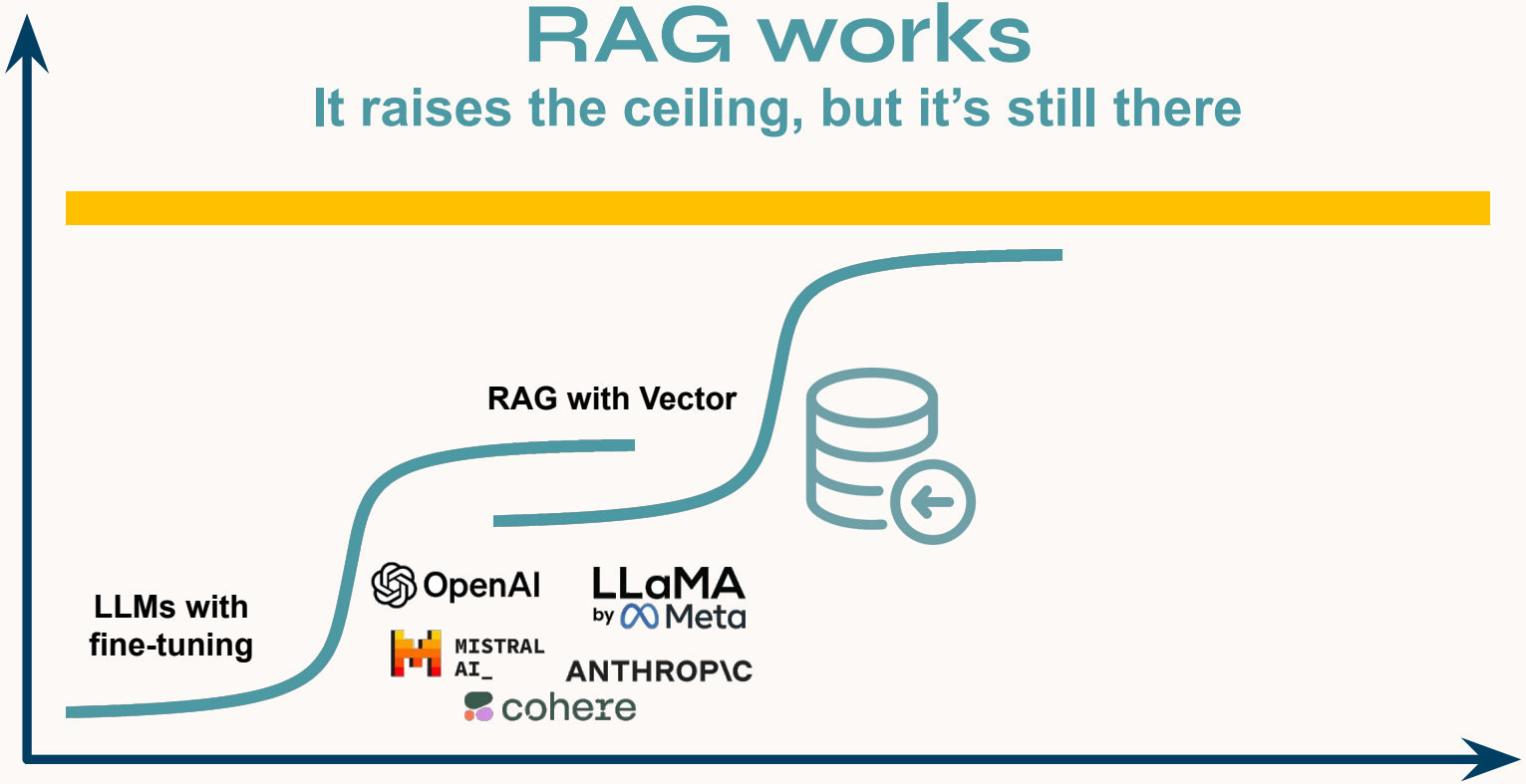
GenAI hockey stick moment




GenAI Go-to-production Barrier
LLMs
OpenAI

RAG works

It raises the ceiling, but it's still there



Neo4j Inc. All rights reserved 2024

Breaking through the next ceiling

Neo4j Inc. All rights reserved 2024

An Analogy

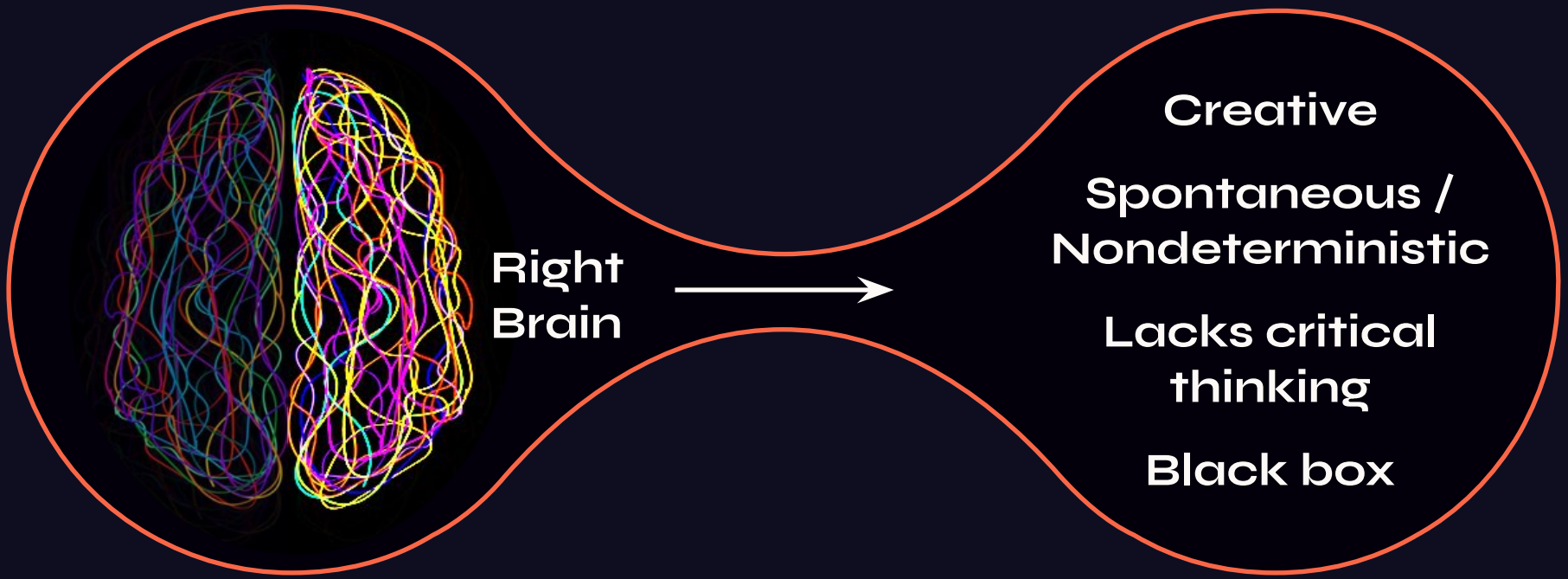
 Neo4j Inc. All rights reserved 2024

Why do knowledge graphs help with GenAI

Neo4j Inc. All rights reserved 2024

LLMs with Vector-Based RAG

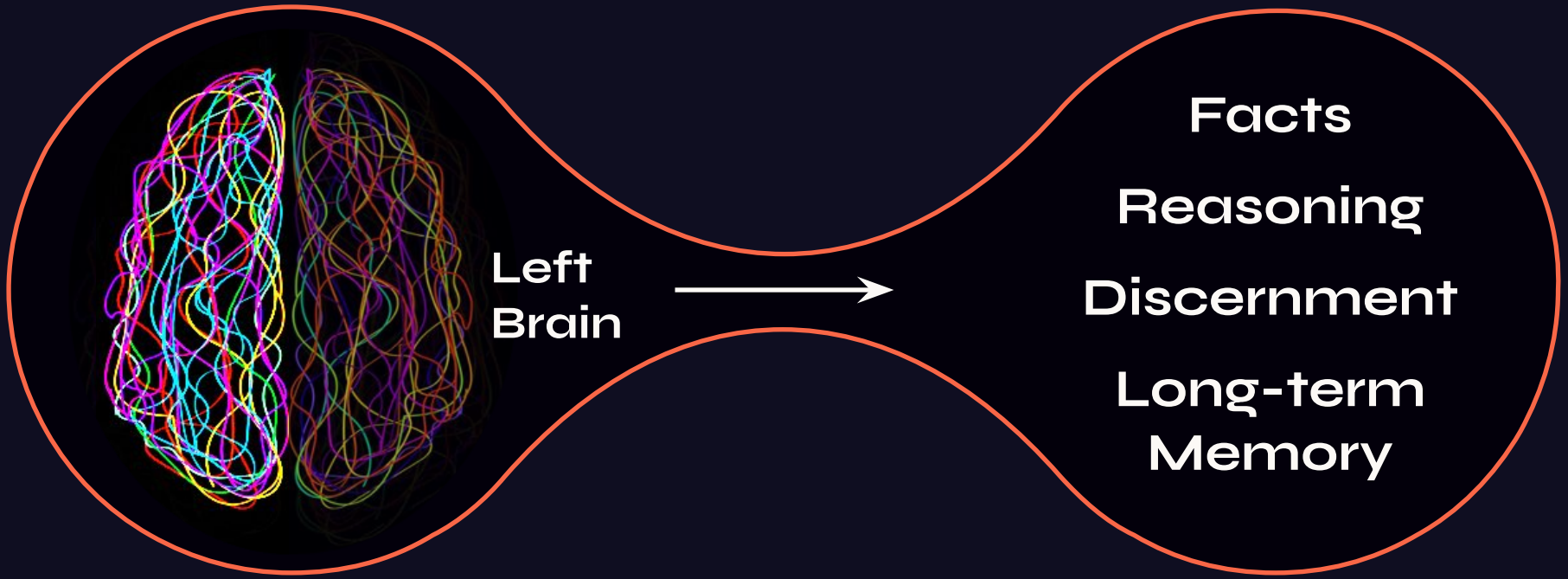
a ‘word model’ view



 Neo4j Inc. All rights reserved 2024

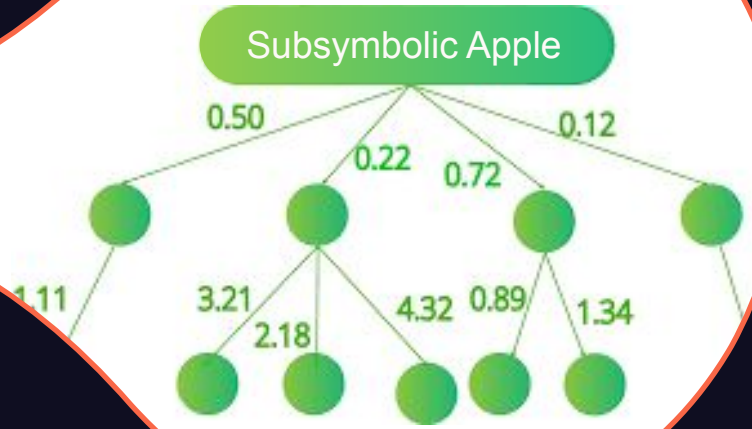
Knowledge Graph

a ‘world model’ view



 Neo4j Inc. All rights reserved 2024

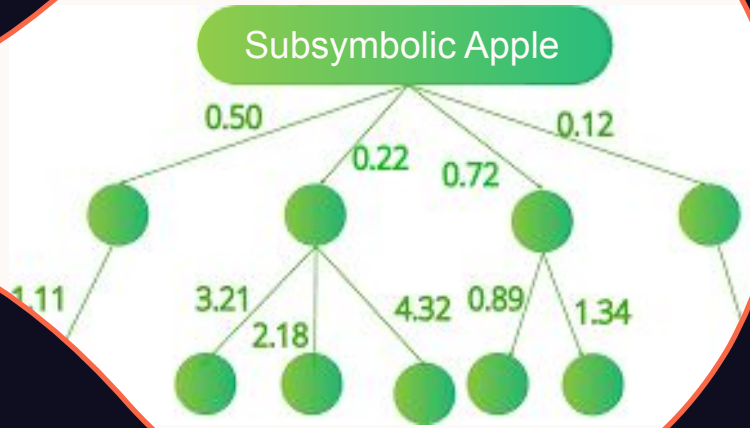
LLM + Vector View of an Apple



Human View of an Apple

Neo4j Inc. All rights reserved 2024

LLM + Vector View of an Apple



Lack explicit domain understanding

Deal in words, not meaning

Answers emerge from language statistics

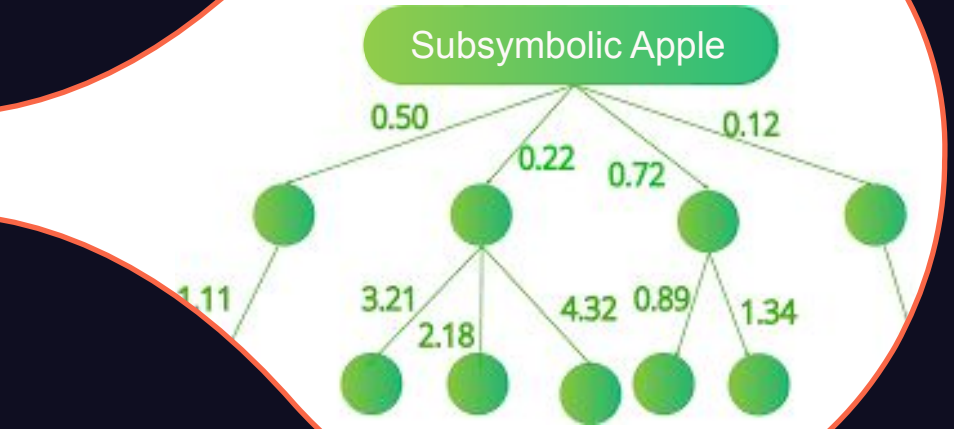
Provide nondeterministic answers

Lack explainability

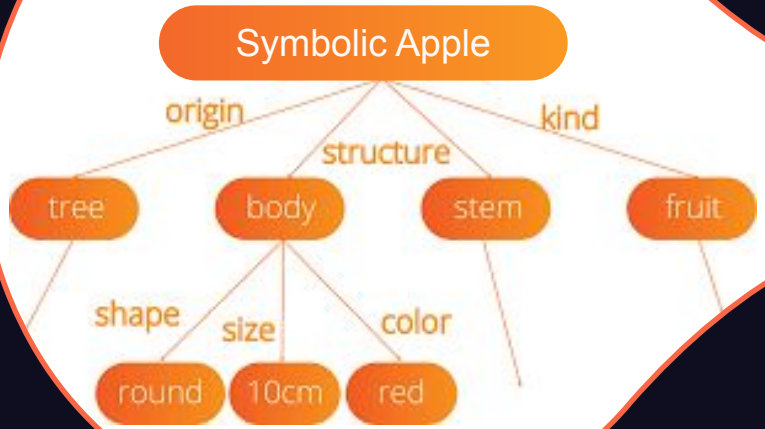
ETHICAL, BIAS, ACCURACY CONCERNS

Neo4j Inc. All rights reserved 2024

LLM + Vector View of an Apple



Knowledge Graph View of an Apple



 Neo4j Inc. All rights reserved 2023

Deal with concepts and meaning

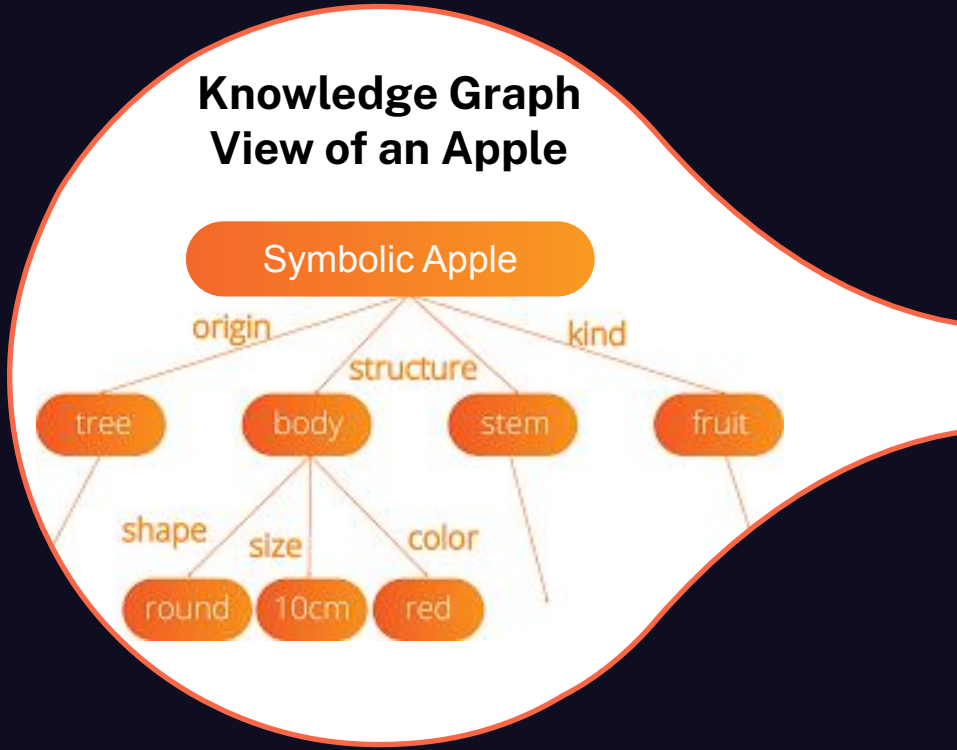
Support human & machine reasoning

Relate concepts, reveals emergent statistics

Answers are deterministic

Support explainability & human understanding

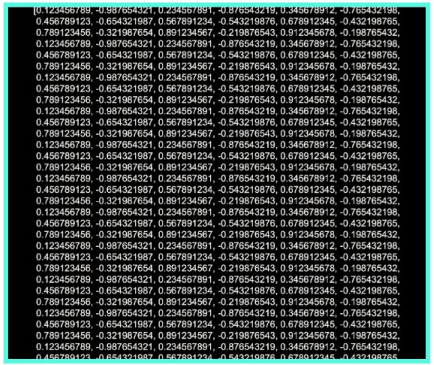
CREATES A SHARED WORLD VIEW



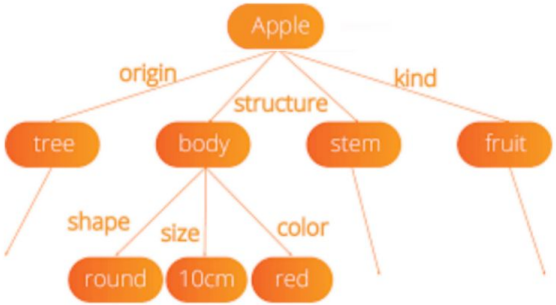
 Neo4j Inc. All rights reserved 2023

Human View of an Apple

Vector View of an Apple



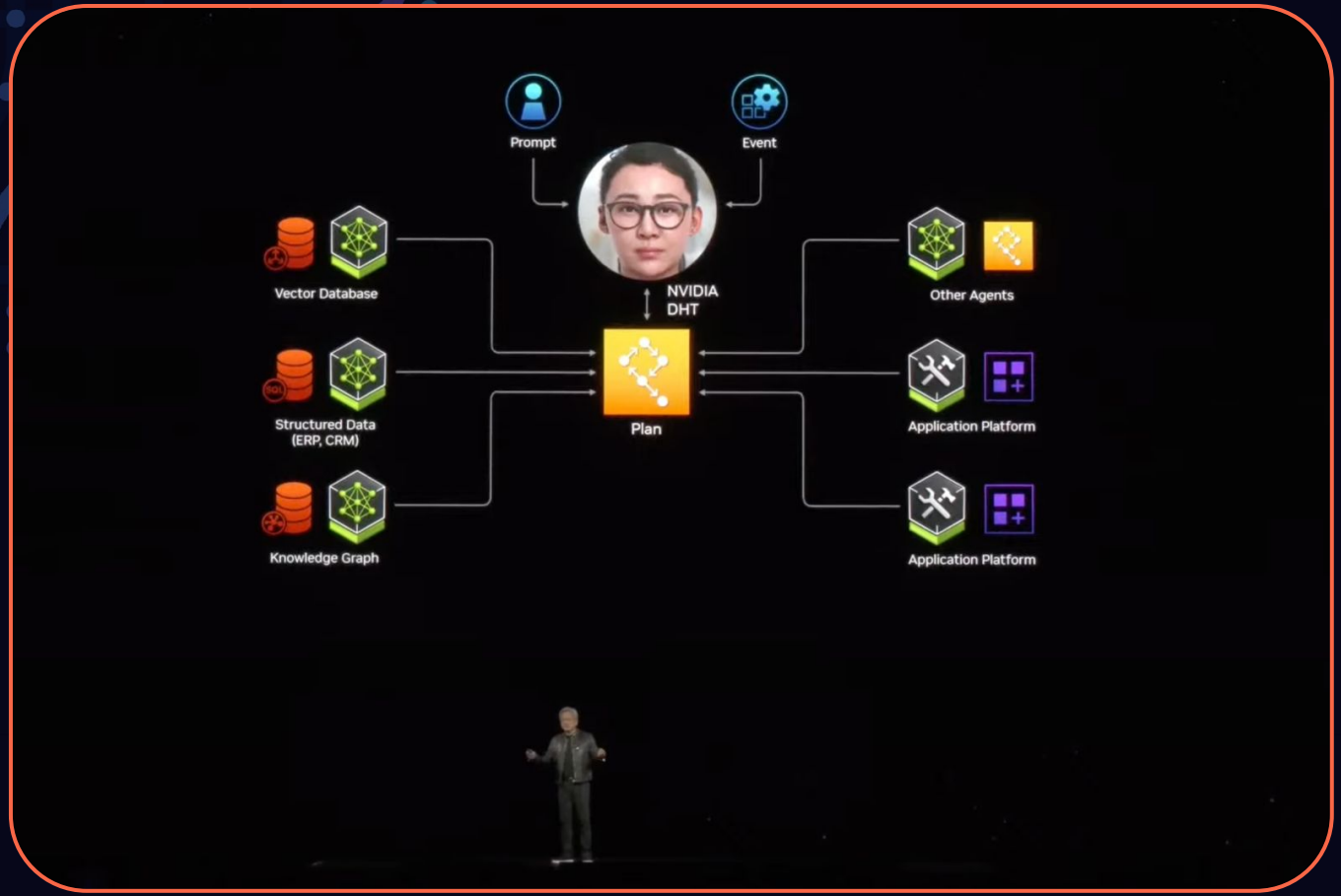
Knowledge Graph View of an Apple



 Neo4j Inc. All rights reserved 2024

GraphRAG @Nvidia

Jensen Huang's Nvidia GTC Keynote



Neo4j Inc. All rights reserved 2024

GraphRAG @Microsoft

teamgraphrag@microsoft.com

 Neo4j Inc. All rights reserved 2024