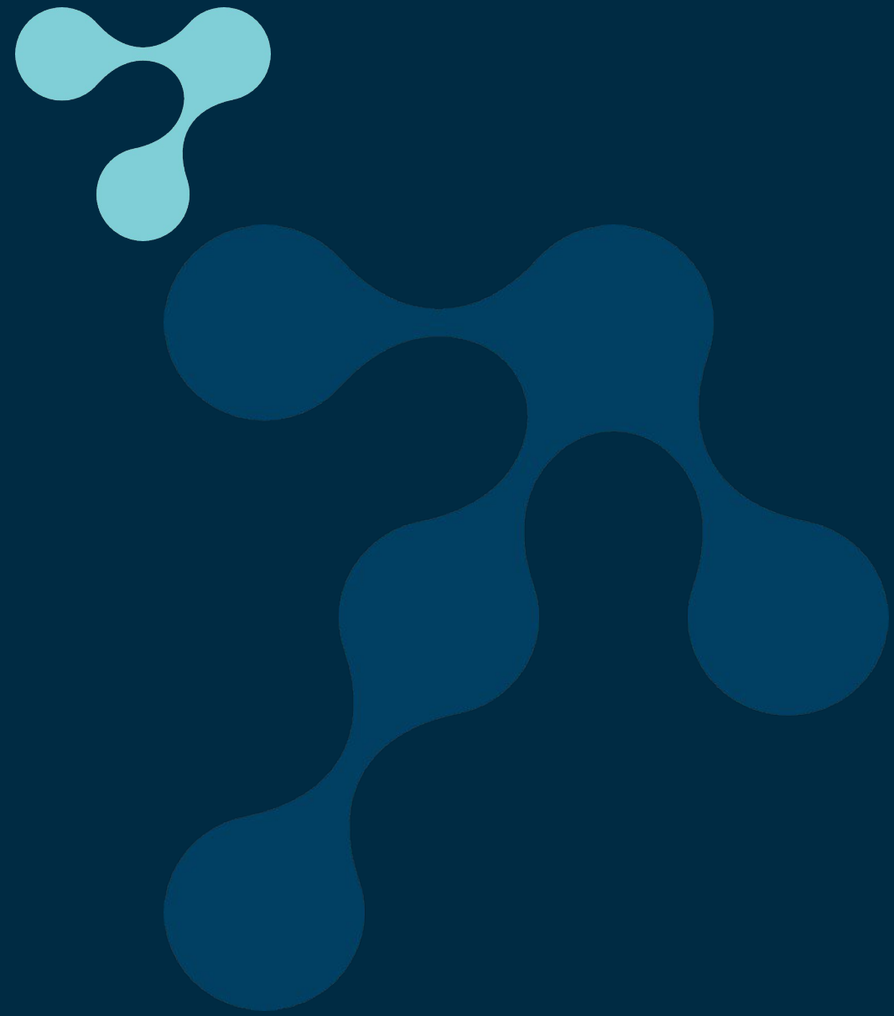
Where does the why go?

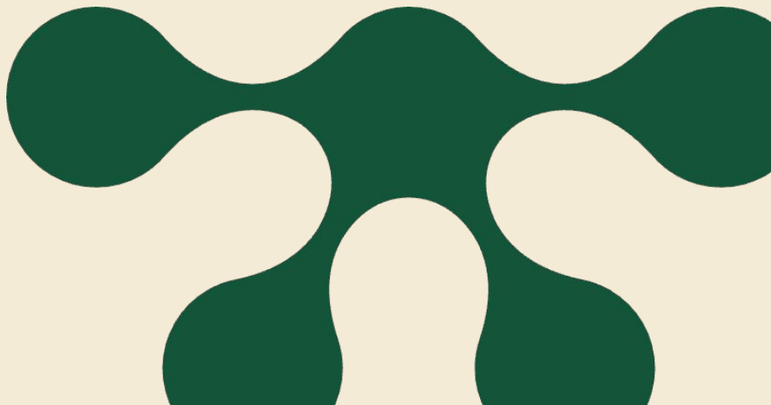
Memory, knowledge graphs, and agents in e-science

Jim Webber, Ph.D.
Chief Scientist | Neo4j



PROLOGUE

Our new robot overlords are just stochastic token generators



WHAT IS THIS?

Has been extensively trained on a wide-range of eclectic subjects

Answers confidently, even when making things up

Opaque about reasoning

Expensive

WHAT IS THIS?



J

If it takes 5 shirts 1 hour to dry in the sun, how long will it take 10 shirts to dry in the sun?



If it takes 5 shirts 1 hour to dry, you can set up a proportion to find out how long it would take for 10 shirts to dry:
(5 shirts / 1 hour) = (10 shirts / x hours)

So, it would take 10 shirts 2 hours to dry in the sun under similar conditions.

J

Who is Emil Eifrem married to?



Emil Eifrem married Johanna Eifrem in 2011.





Could AI be conscious?

Experts believe it's at least possible. We urgently need a plan to navigate the ethical implications



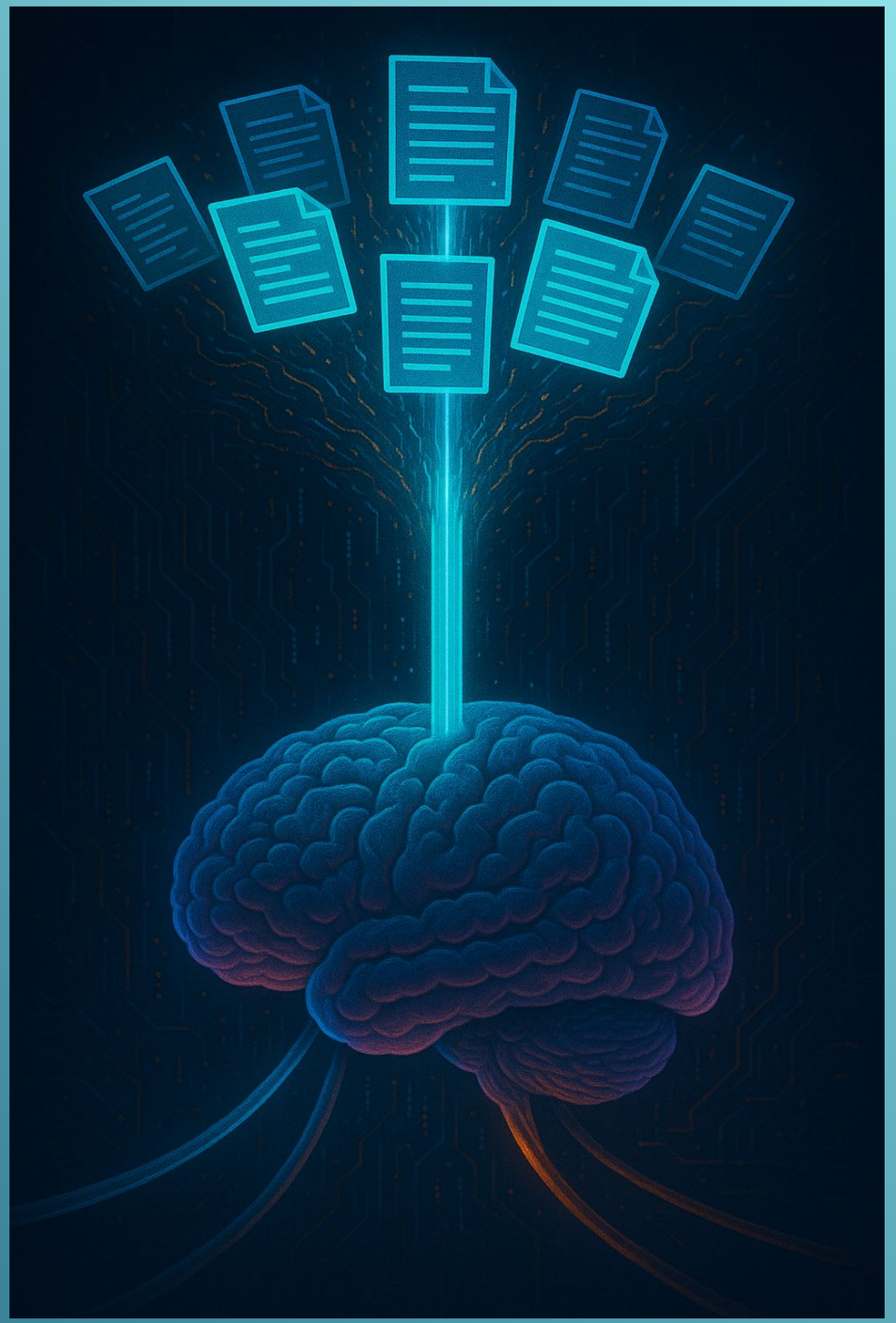
Illustration: Elia Barbieri

https://www.theguardian.com/technology/2026/jul/19/could-ai-be-conscious

No.

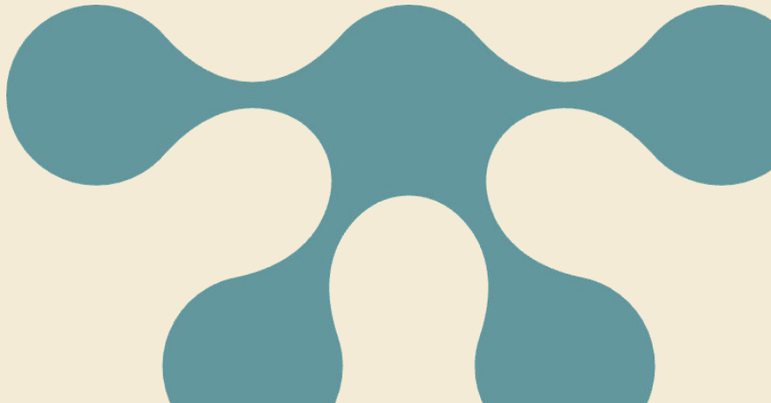
THE TRUTH

Models are context-free
Learning is offline
Mistakes are repeated



PART ONE

How science used to work.



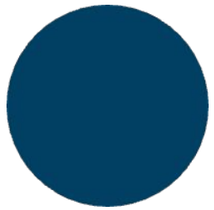
THE RECORD

Three places we wrote it all down.

Lab notebook. Scientific paper. Patents.

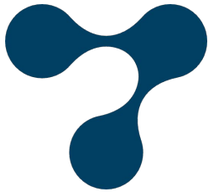
Each one is built to record a **result** — the clean, defensible, finished thing. That's what they're good at.

Where knowledge lived



Lab notebooks

Bench-side, handwritten, one copy, one author.



Papers

Peer-reviewed, polished, the successful runs only.



Patents

Claims and boundaries — written to protect discoveries (not teach them)

Between every step, knowledge leaks out.

Observe — **Experiment** — **Record** — **Publish**

The why

Why this hypothesis, this dose, this cutoff — the reasoning behind the choice.

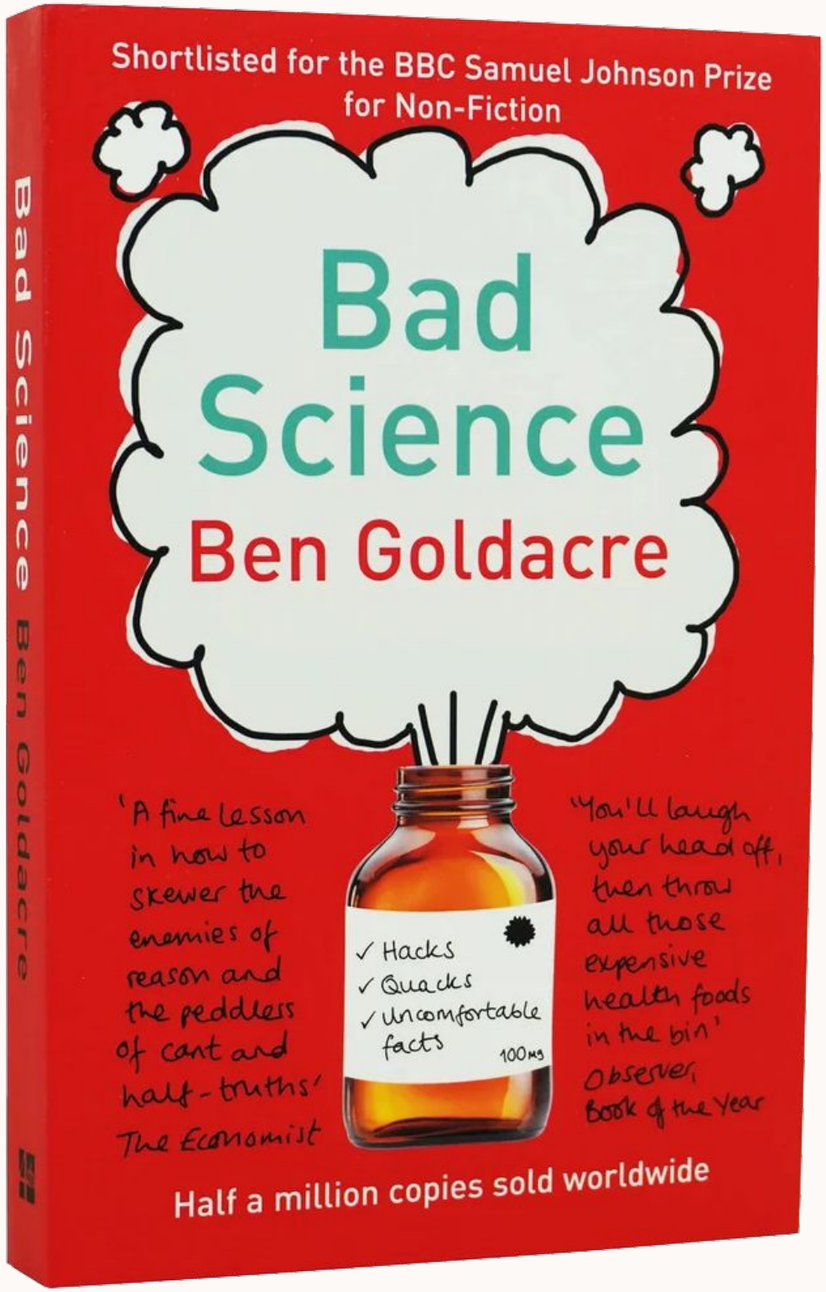
The failures

The runs that didn't work — and the negative results nobody publishes (bad science, good business)

The know-how

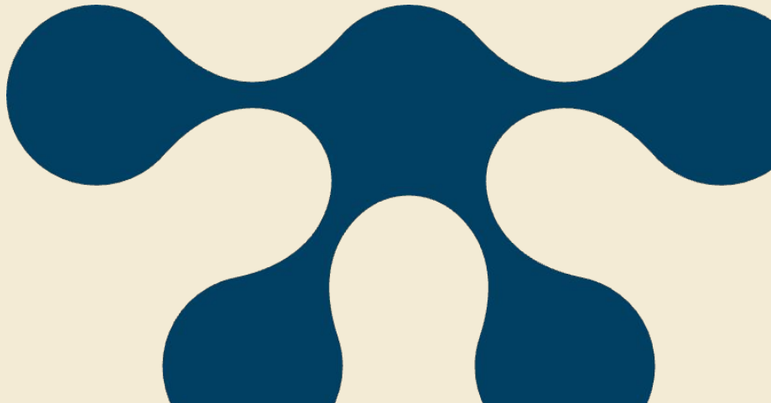
The tacit skill that walks out the door when a researcher leaves the lab.

We are also biased towards positive results!

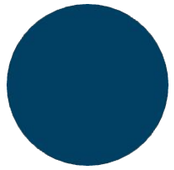


PART TWO

Data as a strategic advantage



We got a lot more sophisticated.



01

Jupyter notebooks

Analysis became executable and reproducible. The *what* now runs on demand.



02

Knowledge graphs

Genes, proteins, diseases, and compounds — finally connected as relationships.



03

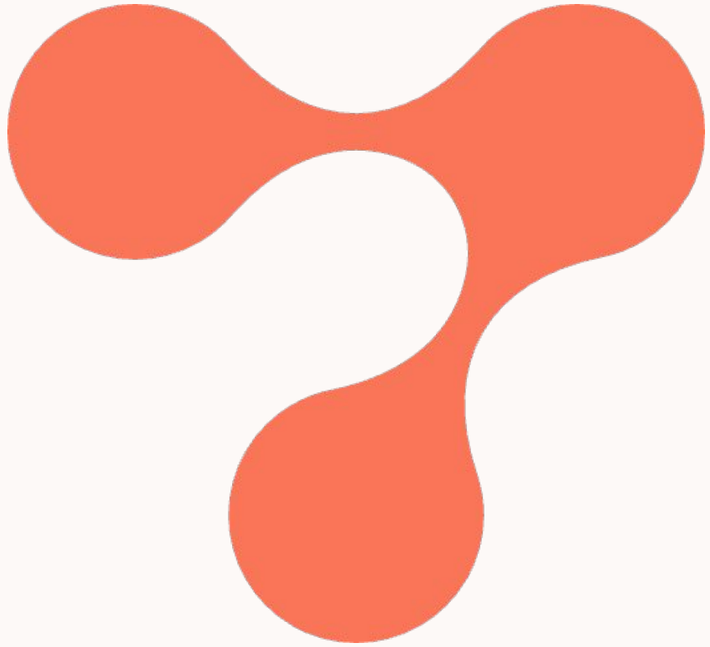
Patents

Still exactly as they were. Plus ca change.

THE SAME STUMBLING BLOCKS

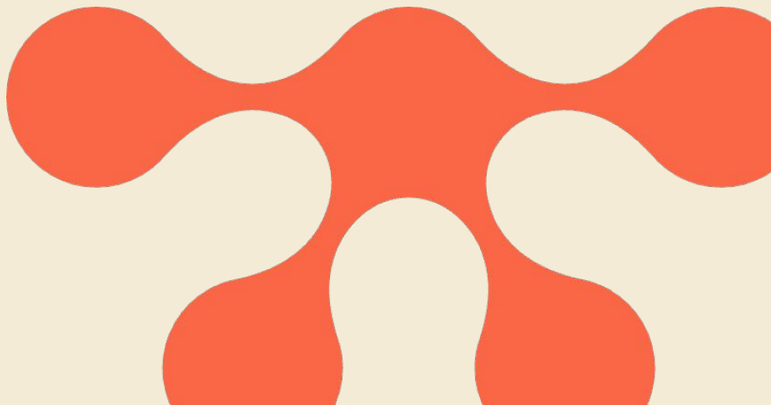
Better tools. Same gap.

We captured more of the **what** than ever. The **why** still falls through — the reasoning and the discarded hypotheses never made it into the data, stayed in our heads as implicit and tribal knowledge.



PART THREE

Then we handed it to the agents.



THE PREMISE

Porting this to an agentic architecture is suboptimal.

THE CORE PROBLEM

Every agent is the greenest postdoc.

Brilliant, tireless, fast. And with **no memory of yesterday.**

On every single invocation it re-derives what the lab already knew, repeats dead ends someone already ruled out, and forgets the moment the context window closes.

A day in the life

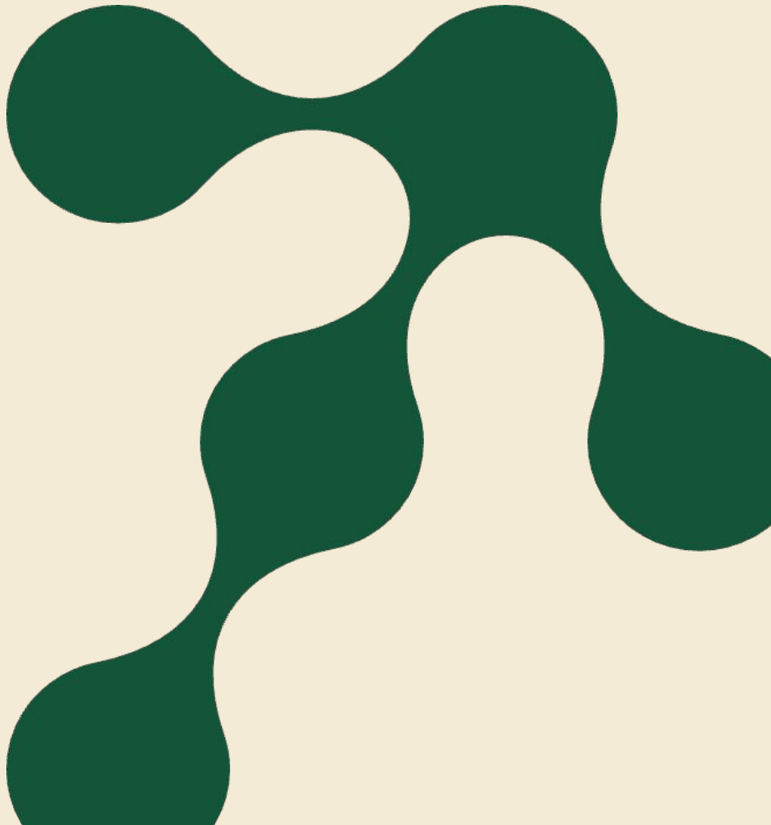
- 09:00 — Onboards. Learns the target biology from scratch.
- 13:00 — Finds a promising lead. Builds real intuition.
- 17:00 — Session ends. Everything it learned is **gone.**
- ↻ — Tomorrow: a brand-new postdoc. Same first day.

Agents inherit our data. They never inherited our experience.

The **what** is queryable. The **why** and the **how** still live in people's heads — and evaporate when a session ends.

PART FOUR

Got: better data. Need: better memory.



Capture the what, the why, and the how.



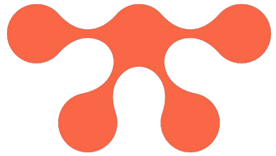
The what

Entities and results — genes, targets, compounds, readouts. The facts, connected.



The why

Hypotheses, evidence, and decisions — **including** the ones that didn't pan out.

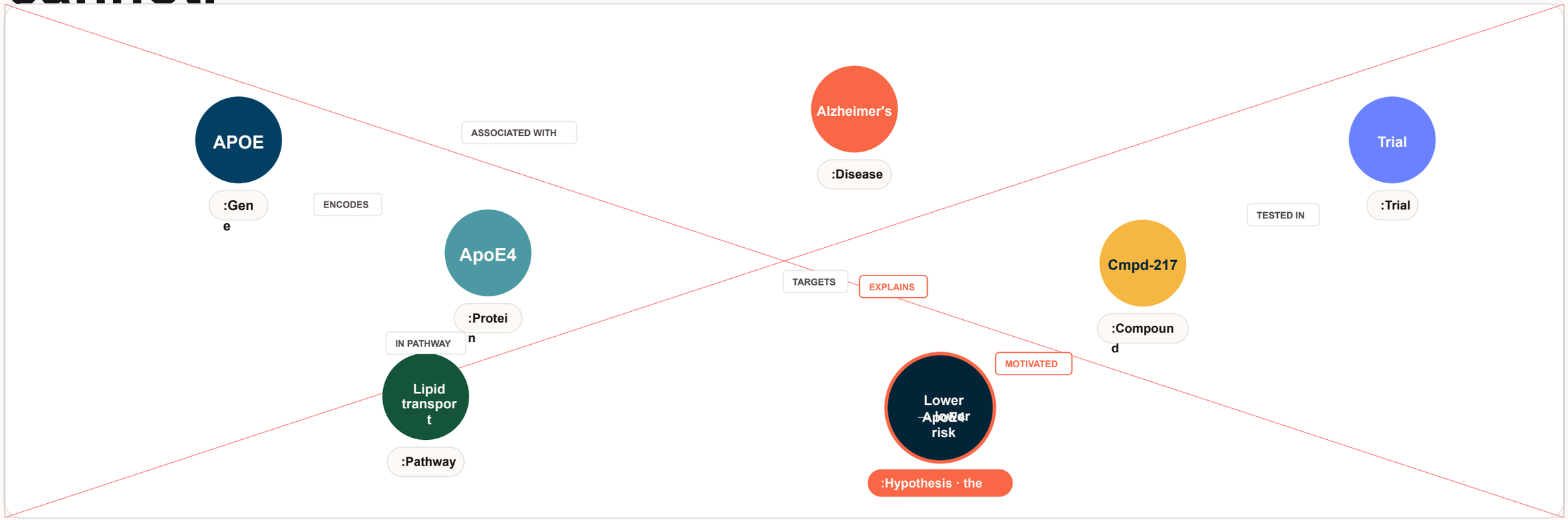


The how

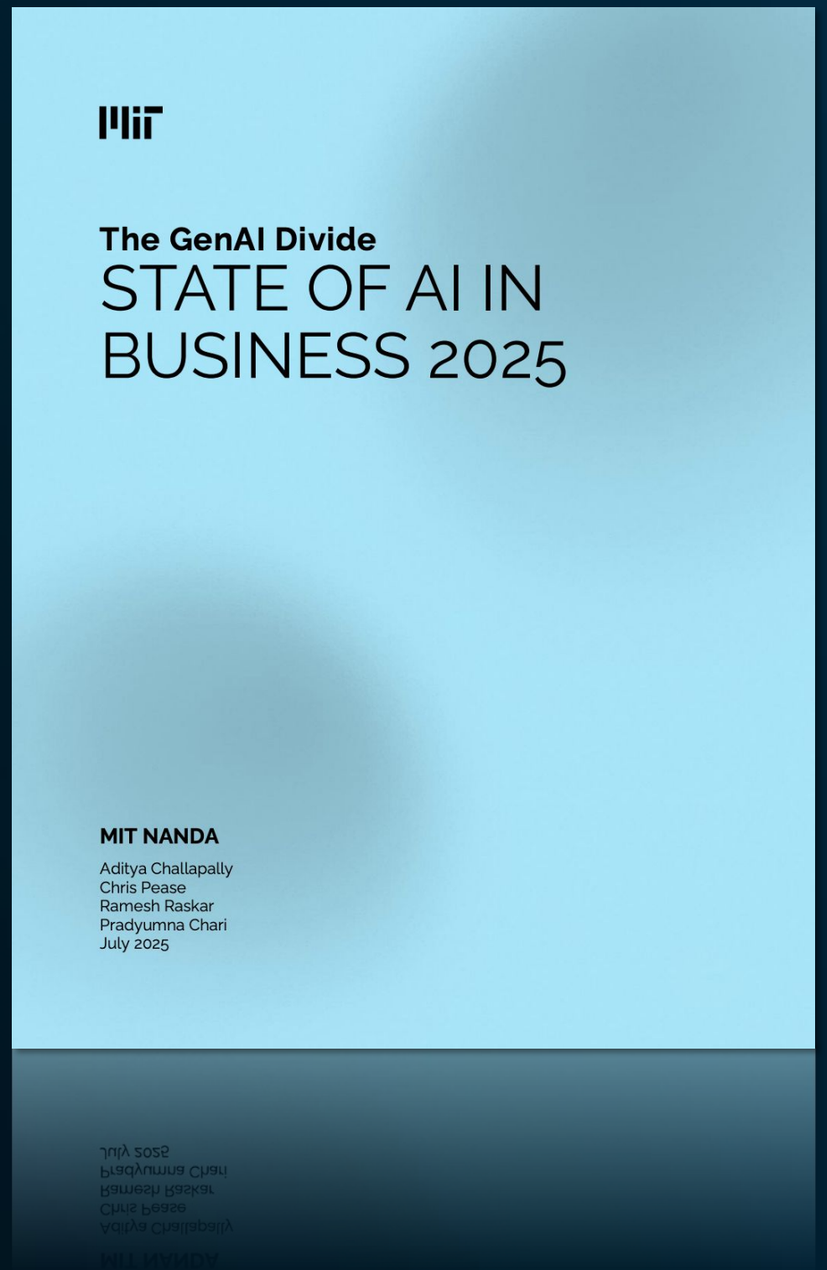
Protocols, parameters, and provenance.

MEMORY, AS A GRAPH

A graph remembers what a paper forgets. And connects facts that notebooks cannot.



“95% of Generative AI efforts fail to deliver value”

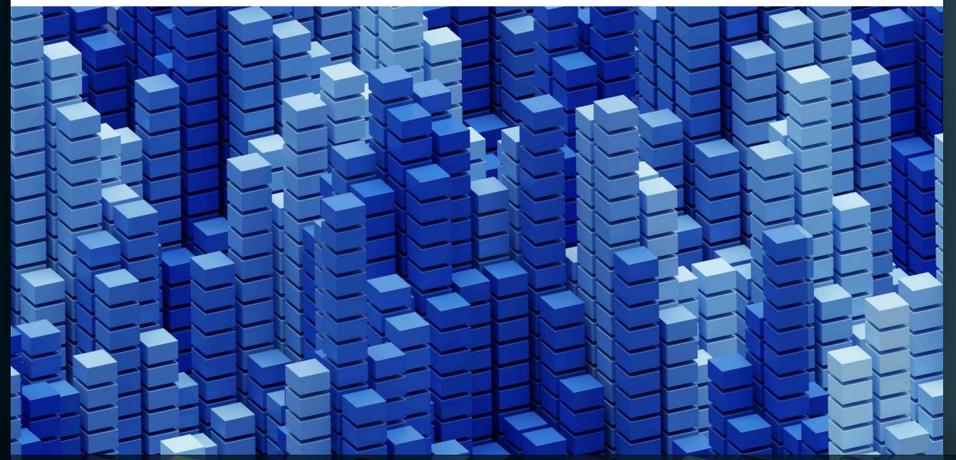


McKinsey Digital

The data dividend: Fueling generative AI

Data leaders should consider seven actions to enable companies to scale their generative AI ambitions.

This article is a collaborative effort by Joe Caserta, Holger Harreis, Kayvaun Rowshankish, Nikhil Srinidhi, and Asin Tavakoli, representing views from McKinsey Digital.



AI And Machine Learning

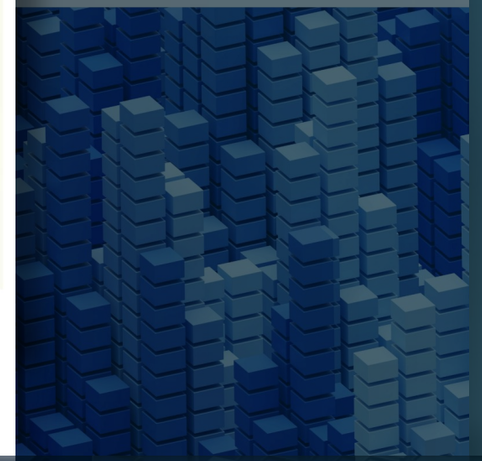
Is Your Company's Data Ready for Generative AI?

by Thomas H. Davenport and Priyanka Tiwari

March 26, 2024



Illustration by Carl Godfrey



Generative AI

To Create Value with AI, Improve the Quality of Your Unstructured Data

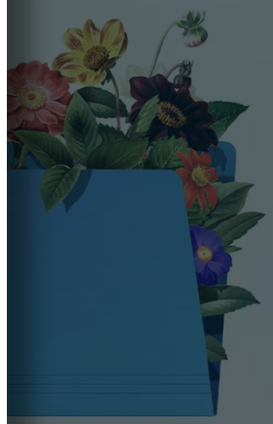
by Thomas H. Davenport, Roger W. Hoerl and Thomas C. Redman

May 28, 2025



Illustration by Rocío Egio

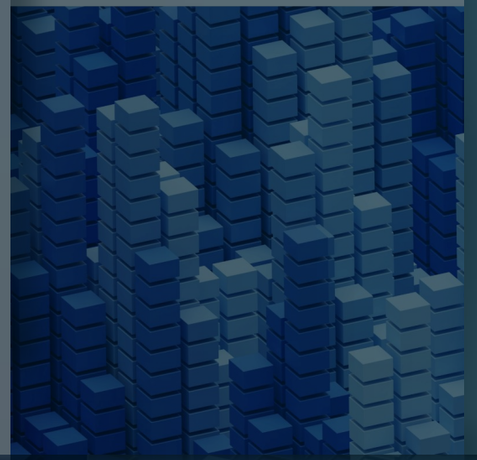
ny's Data
erative AI?
anka Tiwari



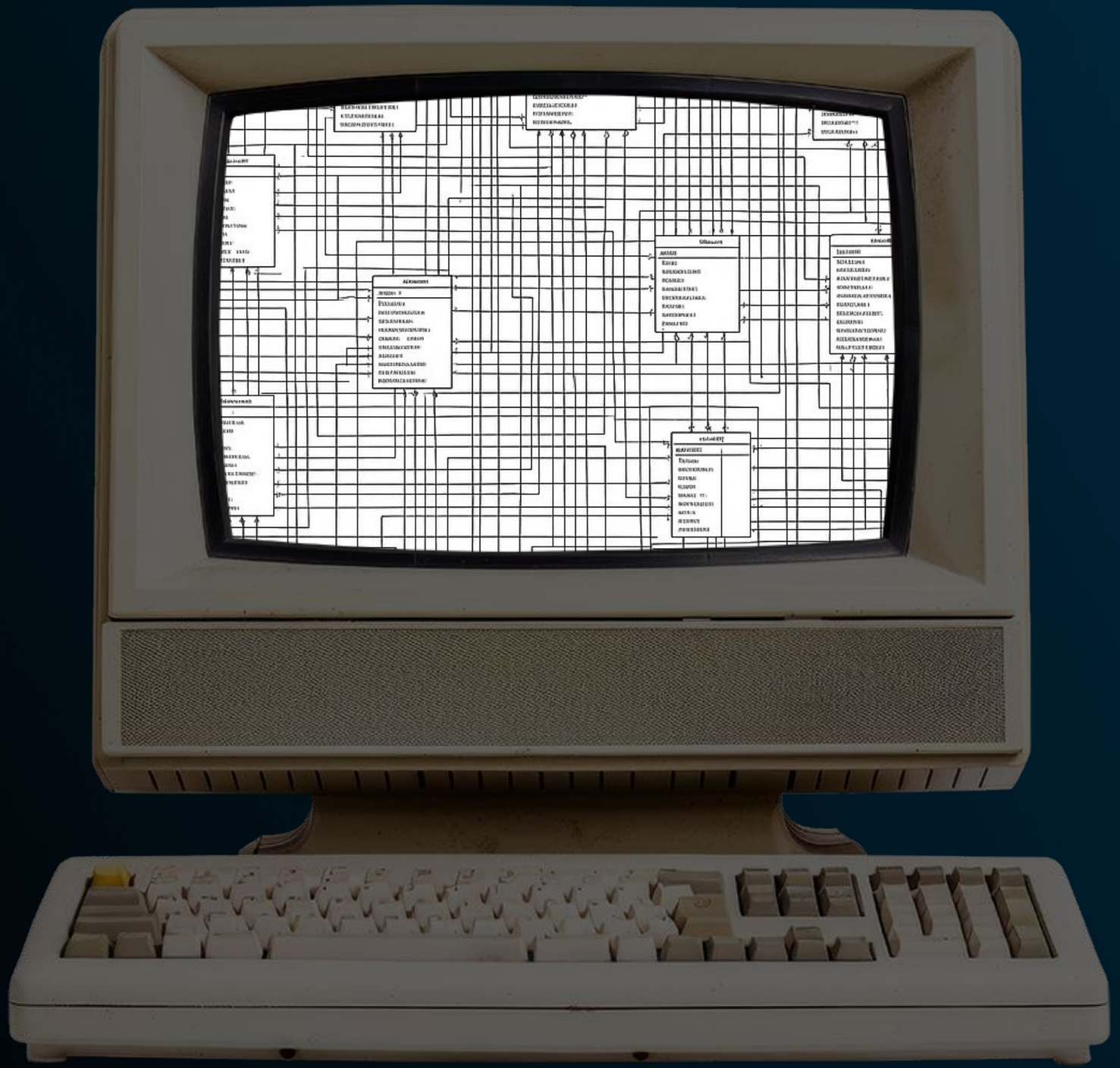
end: Fueling

tions to enable companies to scale

er Harreis, Kayvaun Rowshankish, Nikhil Srinidhi,
gital.

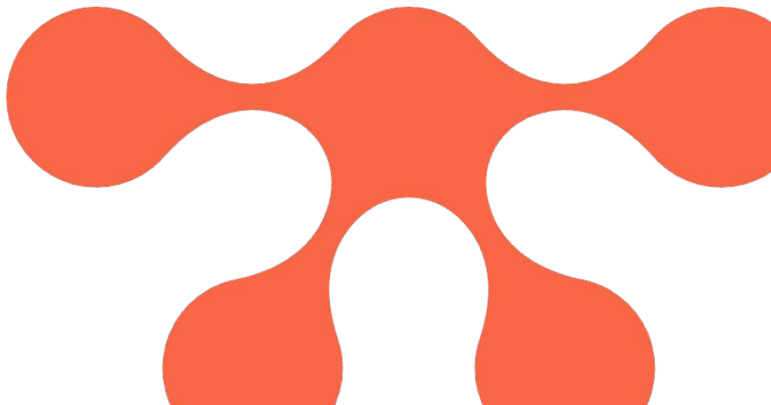


AI IN
S 2025



PART FIVE

All about context and graphs



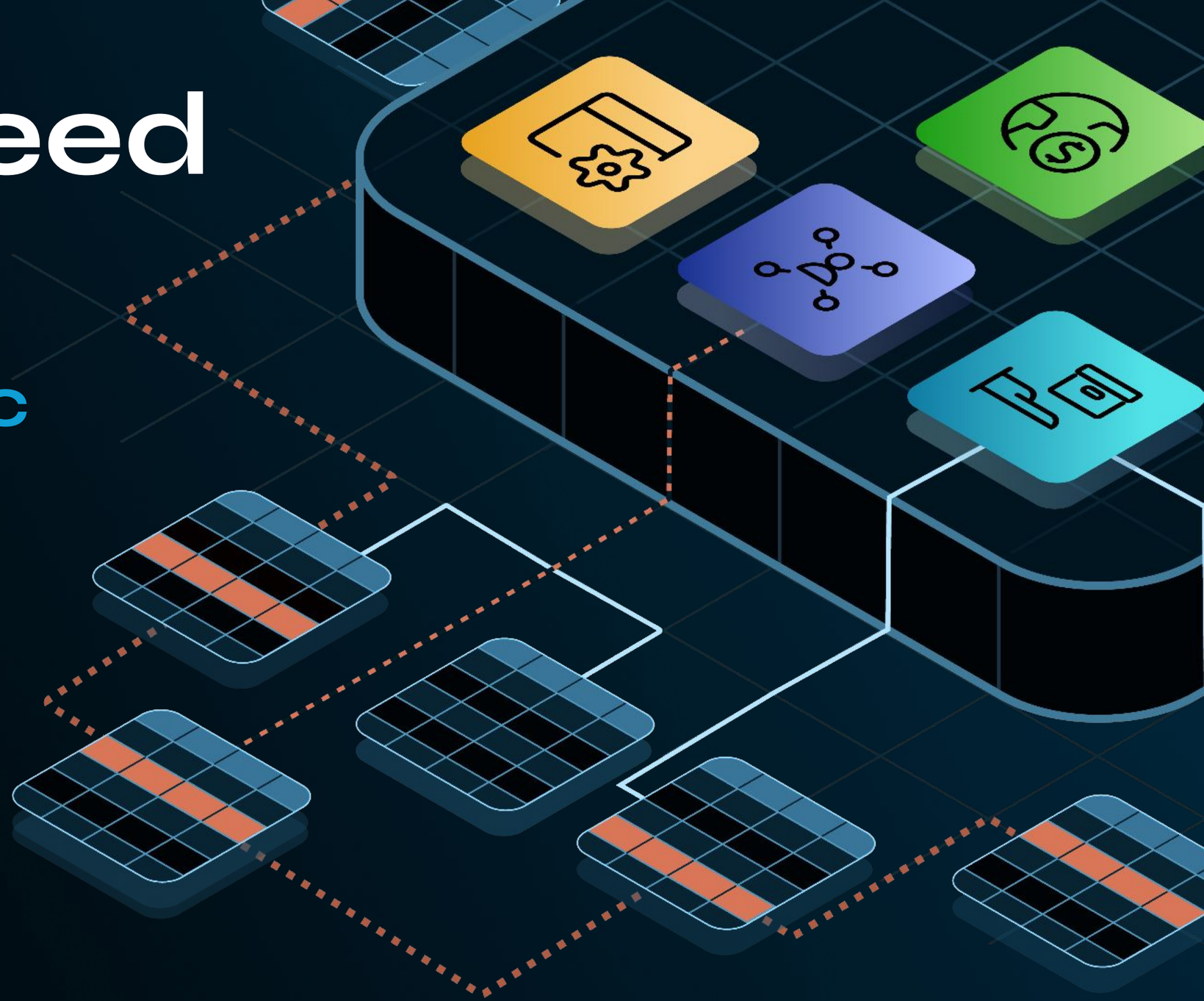
Agents Need Context

The new agentic data model has arrived.

Reasoning Memory

Short-term Memory

Long-term Memory



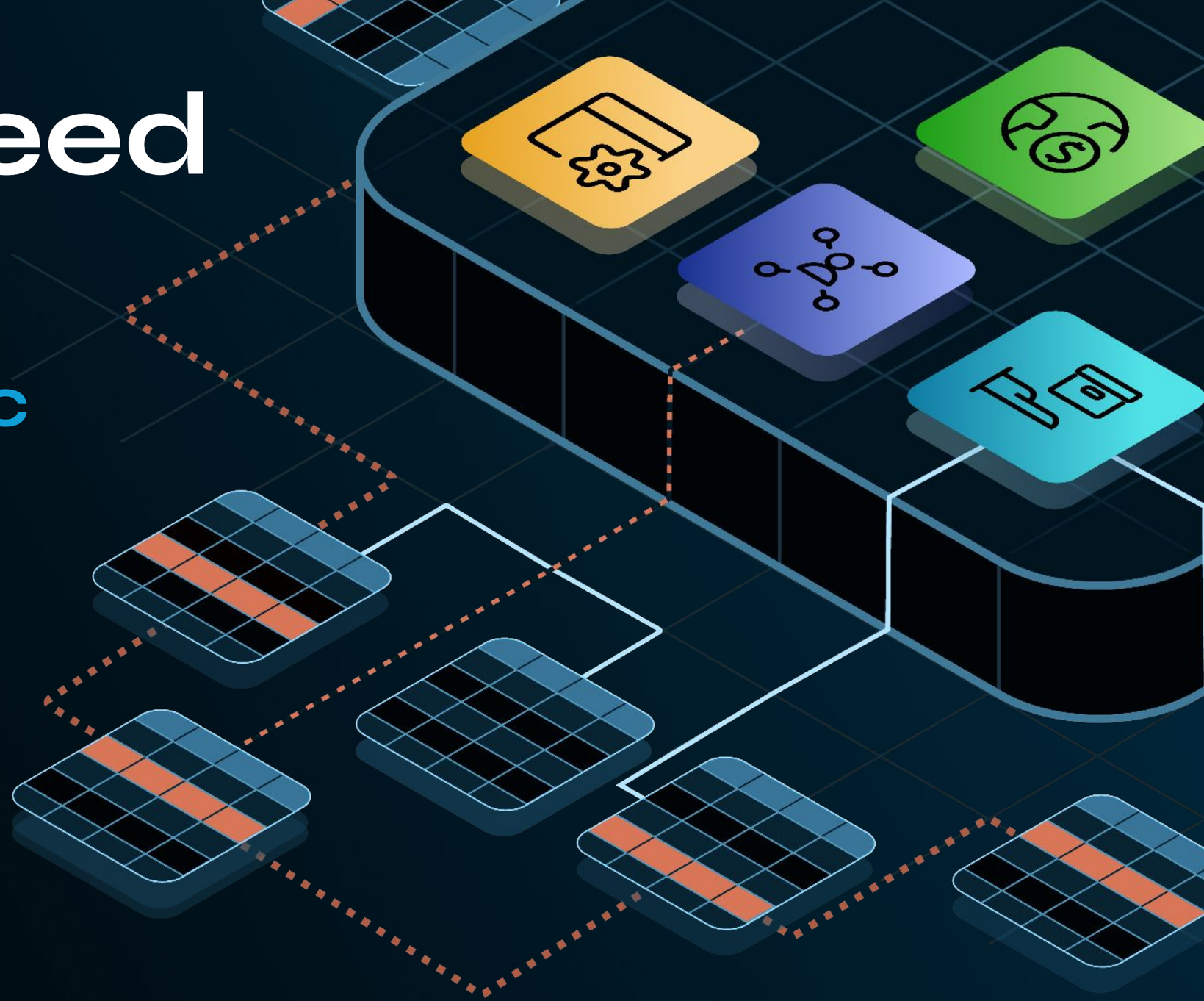
Agents Need Context

The new agentic data model has arrived.

Decision graphs

Conversational graphs

Knowledge graphs



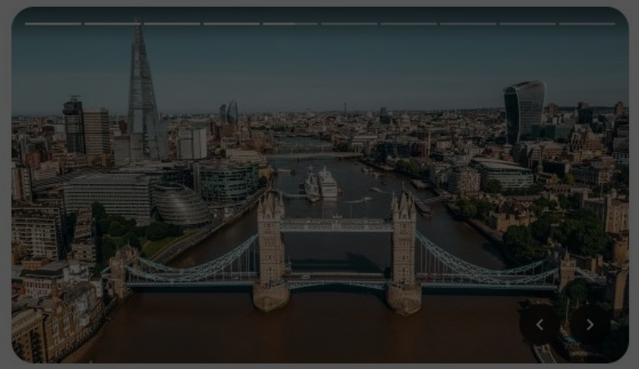
Google

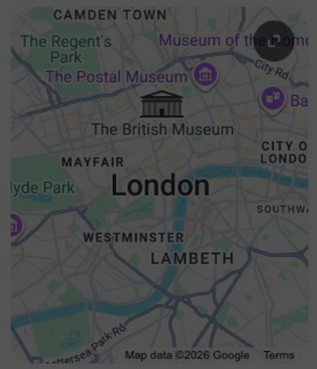
london

AI Mode | All | News | Images | Videos | Maps | Short videos | More | Tools

London

Capital of England





Weather

Tue	Wed	Thu
5°	9°	7°

Google Weather

Directions

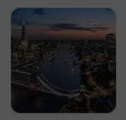
27m from Guildford

1h 12m

Wikipedia
https://en.wikipedia.org › wiki › London

London

London is the capital and largest city of both England and the United Kingdom, with a population of 9.1 million people in 2024. Read more



Things to know



Places to stay
Hotels in London



Where to eat
Restaurants in London

Things to do
Things to do in London

Culture
What is the culture of London like?

About

London, the capital of England and the United Kingdom, is a 21st-century city with history stretching back to Roman times. At its centre stand the imposing Houses of Parliament, the iconic 'Big Ben' clock tower and Westminster Abbey, site of British monarch coronations. Across the Thames River, the London Eye observation wheel provides panoramic views of the South Bank cultural complex, and the entire city. — Google

Population: 8.866 million (2024)
GSS code: E12000007 (region); E61000001 (GLA)
Local government: 32 London boroughs, the City of London and Queen's Park parish
Postcode areas: show 22 areas: E, EC, N, NW, SE, SW, W, WC, BR, CM, CR, DA, EN, HA, IG, KT, RM, SM,UB, WD, TN, TW
Region: London
Settled by Romans: AD 47; 1978 years ago, as Londinium
Sovereign state: United Kingdom

Feedback

Google

london

AI Mode All

About

London, the capital of England and the United Kingdom, is a 21st-century city with history stretching back to Roman times. At its centre stand the imposing Houses of Parliament, the iconic ‘Big Ben’ clock tower and Westminster Abbey, site of British monarch coronations. Across the Thames River, the London Eye observation wheel provides panoramic views of the South Bank cultural complex, and the entire city. — Google

Population: 8.866 million (2024)
GSS code: E12000007 (region); E61000001 (GLA)
Local government: 32 London boroughs, the City of London and Queen's Park parish
Postcode areas: show 22 areas: E, EC, N, NW, SE, SW, W, WC, BR, CM, CR, DA, EN, HA, IG, KT, RM, SM,UB, WD, TN, TW
Region: London
Settled by Romans: AD 47; 1978 years ago, as Londinium
Sovereign state: United Kingdom

Feedback

London
Capital of Engl

Wiki
https://en.

London
London is the
population of 9

Things to

Pla
Hot

Wh
Res

Things to d
Things to do

Culture
What is the culture of London like?

Wed 9°
Thu 7°
Weather

1h 12m
ldford

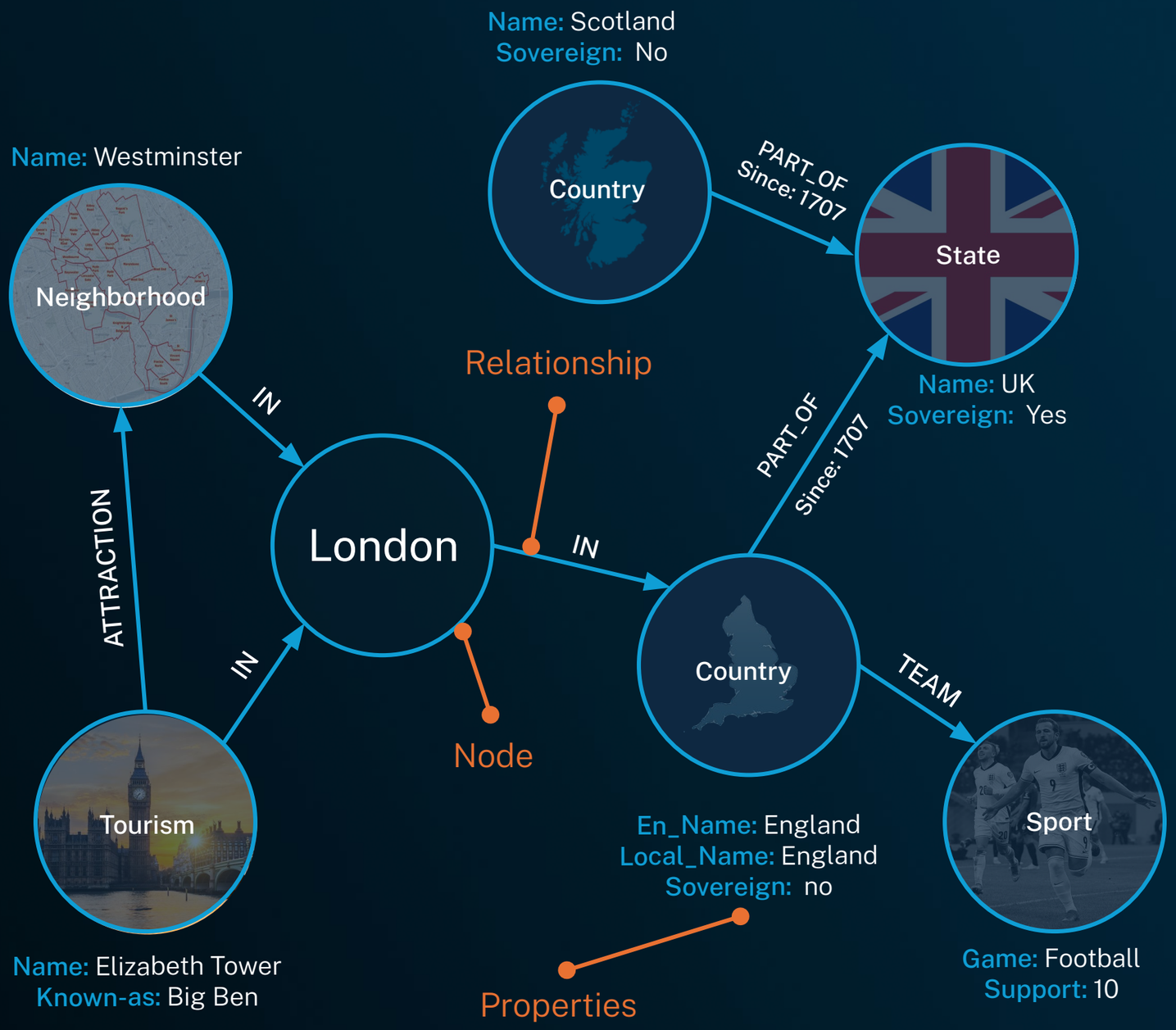
and and the United Kingdom, is
ory stretching back to Roman
e imposing Houses of
Ben' clock tower and
British monarch coronations.
he London Eye observation
iews of the South Bank cultural
. — Google
2024)
gion); E61000001 (GLA)
don boroughs, the City of
parish
areas: E, EC, N, NW, SE, SW,
N, HA, IG, KT, RM, SM,UB,
; 1979 years ago, as Londinium
Sovereign state: United Kingdom
Feedback

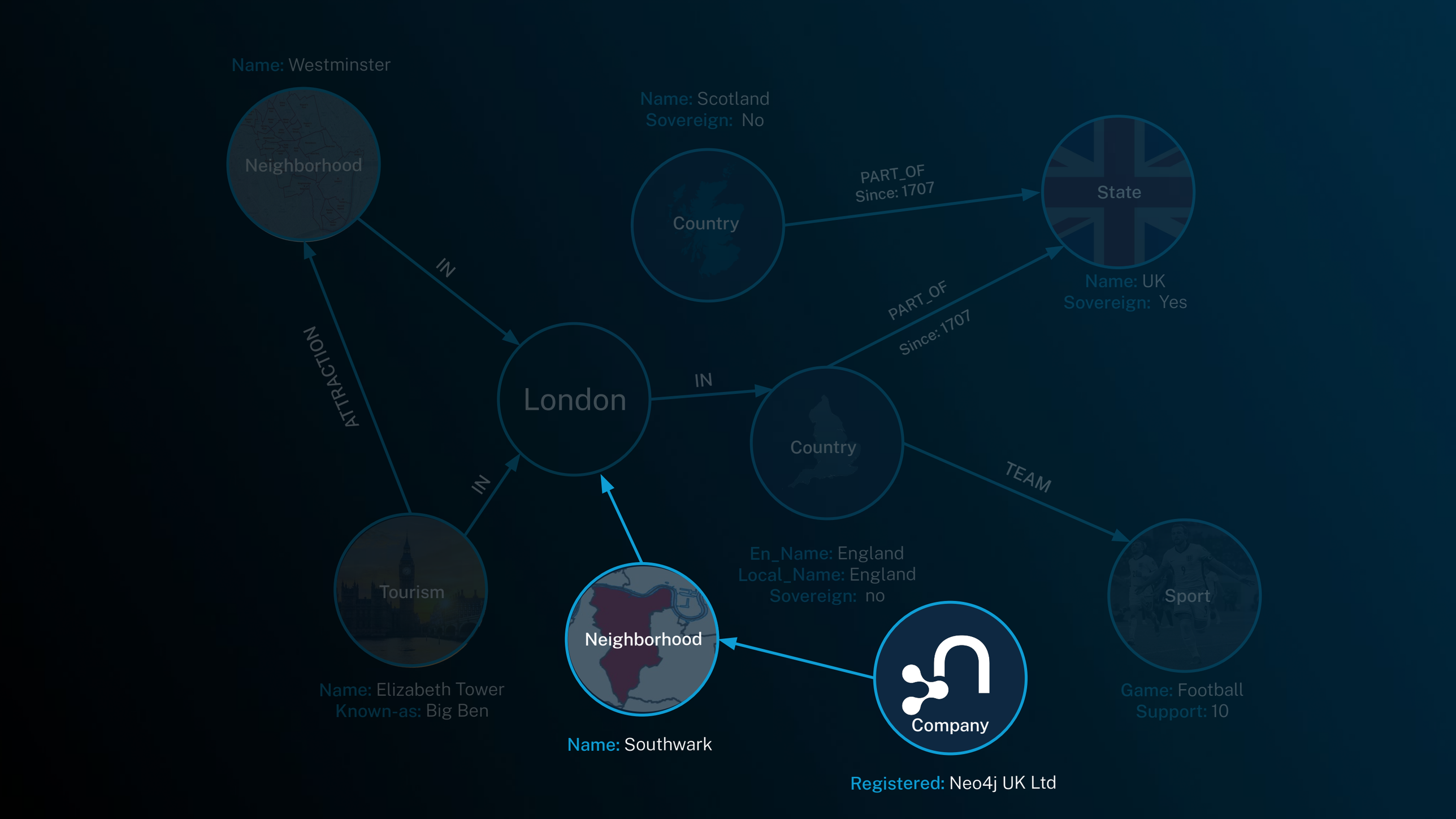
About

London, the capital of England and the United Kingdom, is a 21st-century city with history stretching back to Roman times. At its centre stand the imposing Houses of Parliament, the iconic 'Big Ben' clock tower and Westminster Abbey, site of British monarch coronations. Across the Thames River, the London Eye observation wheel provides panoramic views of the South Bank cultural complex, and the entire city. — Google

Population: 8.866 million (2024)
GSS code: E12000007 (region); E61000001 (GLA)
Local government: 32 London boroughs, the City of London and Queen's Park parish
Postcode areas: show 22 areas: E, EC, N, NW, SE, SW, W, WC, BR, CM, CR, DA, EN, HA, IG, KT, RM, SM,UB, WD, TN, TW
Region: London
Settled by Romans: AD 47; 1978 years ago, as Londinium
Sovereign state: United Kingdom

Feedback

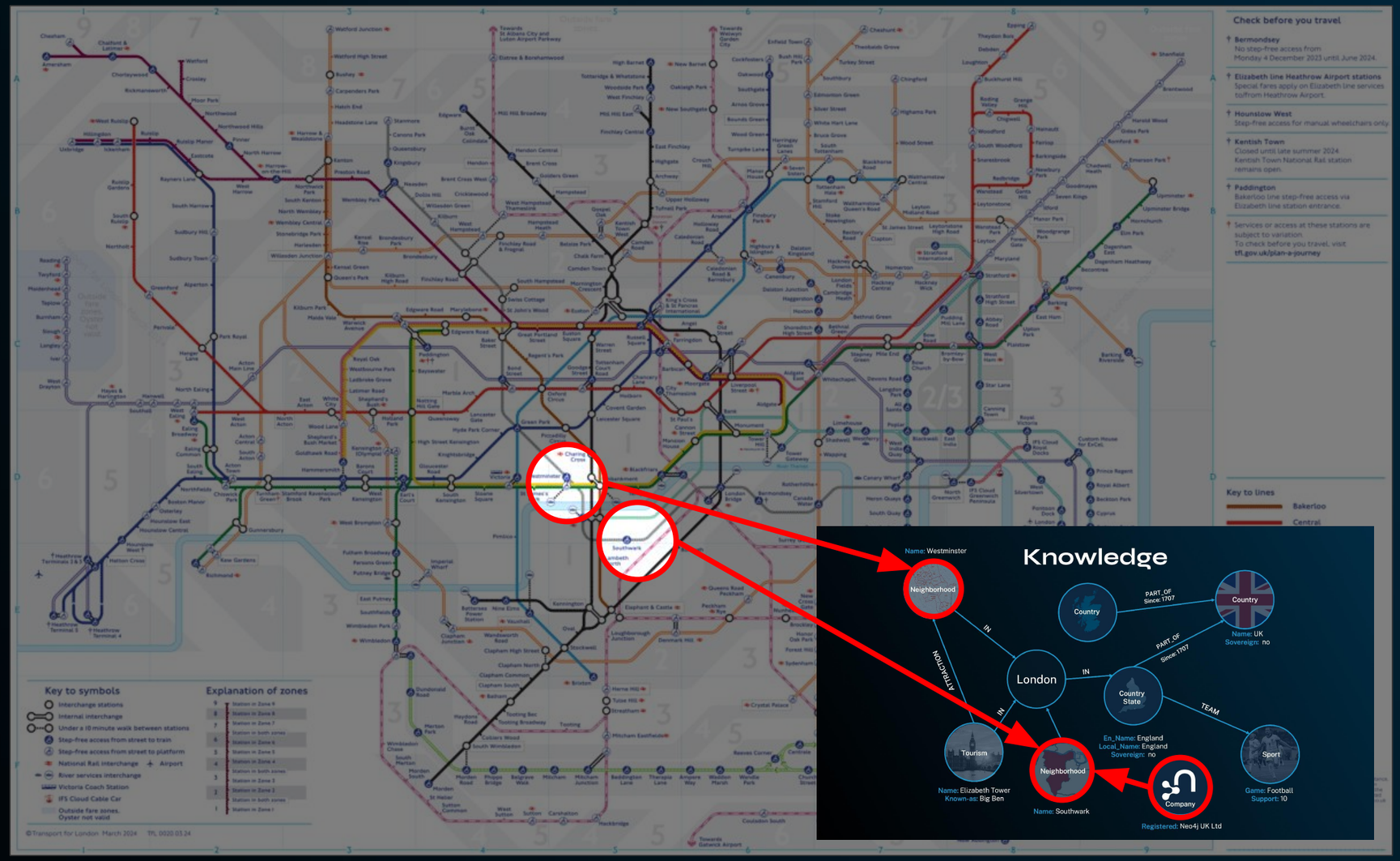



Name: Westminster
Neighborhood
IN
ATTRACTION
Name: Scotland
Sovereign: No
Country
PART_OF
Since: 1707
State
Name: UK
Sovereign: Yes
PART_OF
Since: 1707
London
IN
Country
TEAM
Tourism
IN
Name: Elizabeth Tower
Known-as: Big Ben
En_Name: England
Local_Name: England
Sovereign: no
Neighborhood
Name: Southwark
Company
Registered: Neo4j UK Ltd
Sport
Game: Football
Support: 10

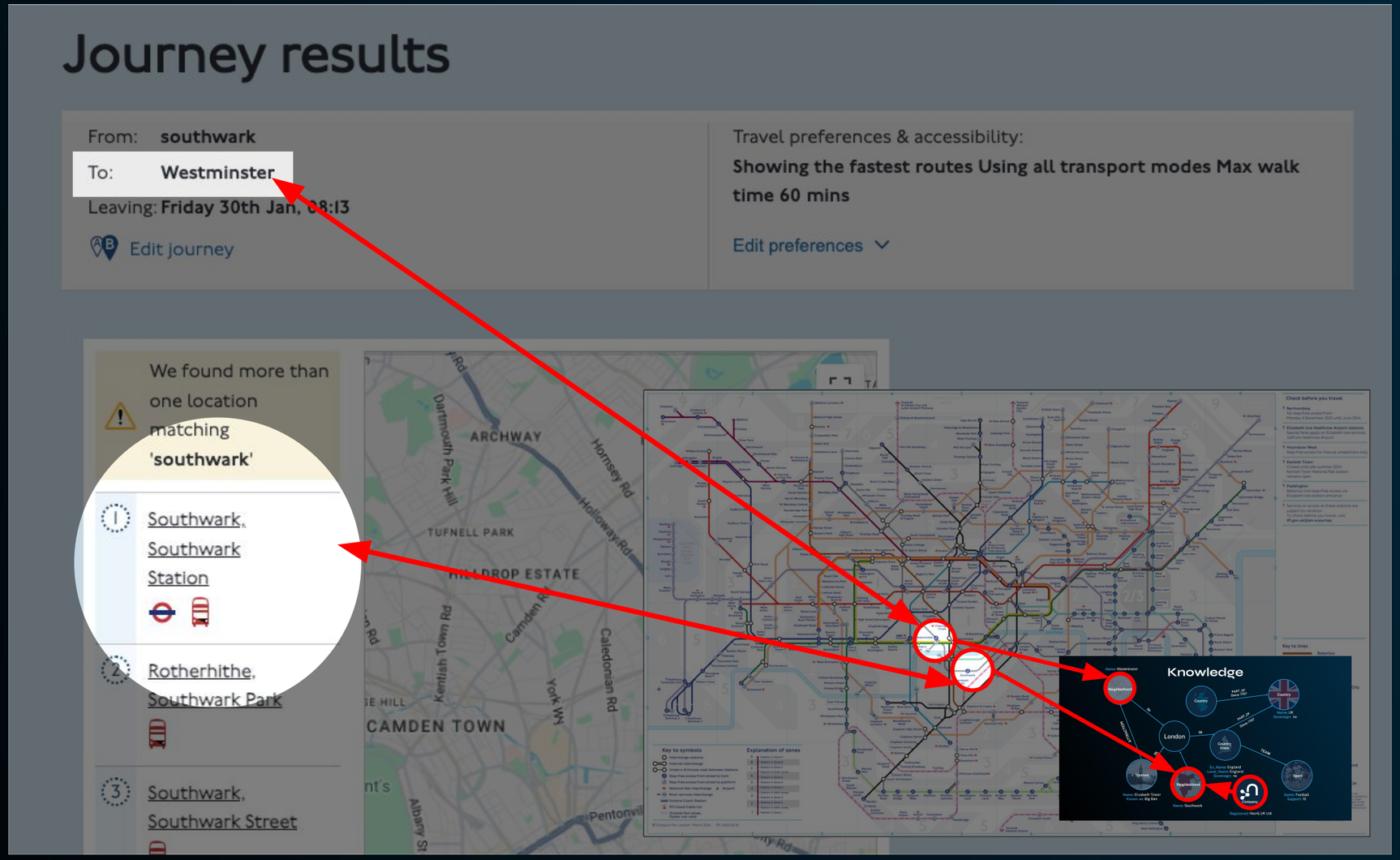
Knowledge


Name: Westminster
Neighborhood
IN
ATTRACTION
Name: Scotland
Sovereign: No
Country
PART_OF
Since: 1707
State
Name: UK
Sovereign: Yes
PART_OF
Since: 1707
London
IN
State
TEAM
IN
Tourism
Name: Elizabeth Tower
Known-as: Big Ben
Neighborhood
Name: Southwark
En_Name: England
Local_Name: England
Sovereign: no
Company
Registered: Neo4j UK Ltd
Sport
Game: Football
Support: 10

Conversation

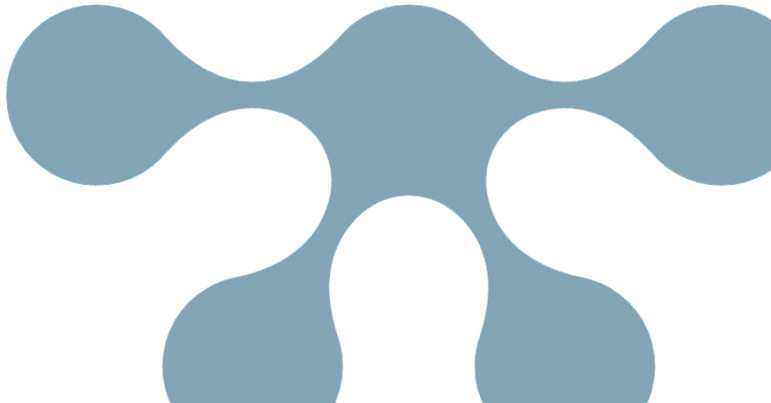


Decision


Journey results
From: southwark
To: Westminster
Leaving: Friday 30th Jan, 08:13
Edit journey
Travel preferences & accessibility:
Showing the fastest routes Using all transport modes Max walk time 60 mins
Edit preferences
We found more than one location matching 'southwark'
1 Southwark, Southwark Station
2 Rotherhithe, Southwark Park
3 Southwark, Southwark Street
ARCHWAY
TUFNELL PARK
CAMDEN TOWN
Knowledge
London

PART SIX

Real-world example: Merck



Merck: Directions, Challenges, and Platforms

Digitalization

Merck’s management team have a strong focus on digitizing all data and processes. Their data assets are a competitive advantage, but utilizing them can be hard.

Complex data gathering

Disparate data sources (including Excel and email), differing access permissions to actual databases, low awareness of data sources leading to meetings! Frustrating, time consuming, wasteful.

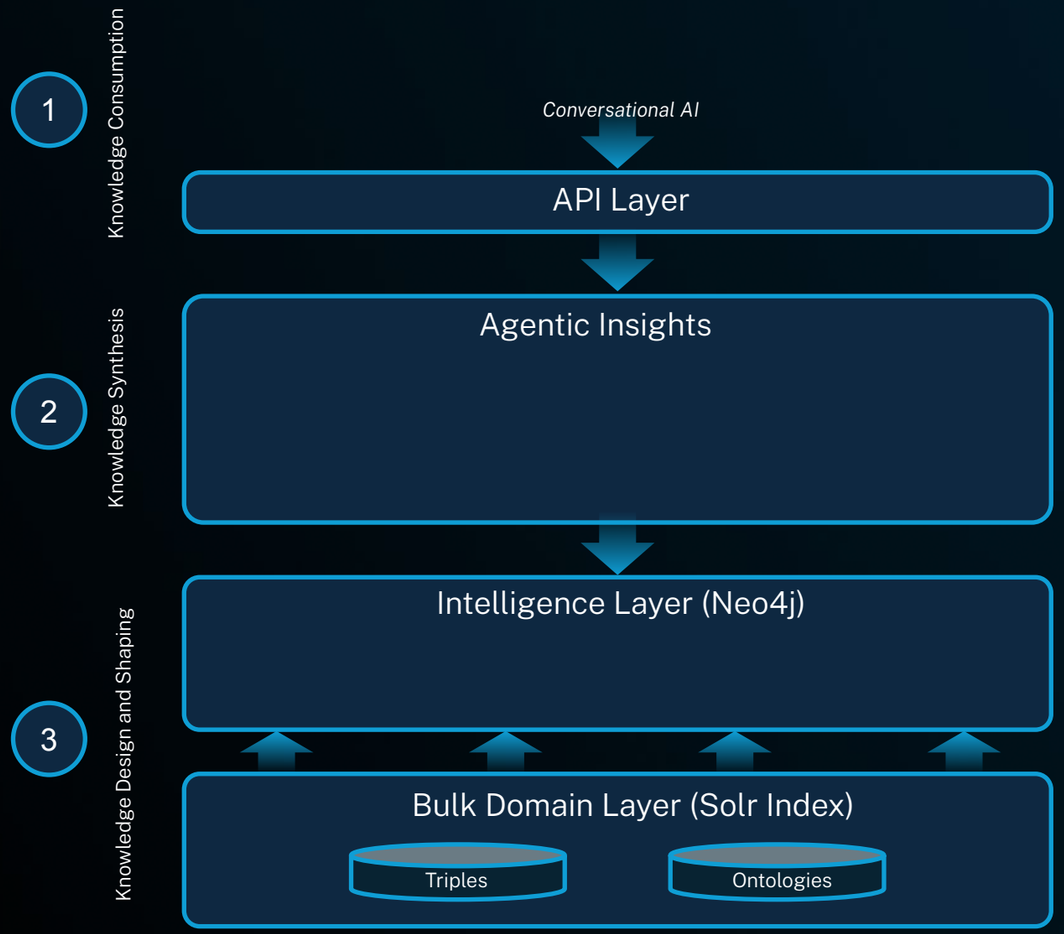
Synaptix

Merck’s platform for sharing knowledge across R&D. Defines a knowledge graph that is the connective fabric across molecular, clinical, operational, and regulatory domains. Serves roles from discovery to launch, and from scientific reasoning to compliance traceability

Merck Architecture

Agentic Insights | Analytics | Bulk Data

Three core parts of the solution



Knowledge Consumption

Enable users of the system to have conversation with the agentic systems to discover data insights (or let them self-serve with APIs or raw graphs)

Knowledge Synthesis

A group of agents, working collaboratively to discover new therapeutics. Agents and critics together we access to other information systems

Knowledge Design and Shaping

Bulk data (Solr index) feeds Neo4j databases for advanced analytics and agents. Scalable system for ingestion and exploitation of data.

IN PRODUCTION

Neo4j runs the agents.

The graph is the memory

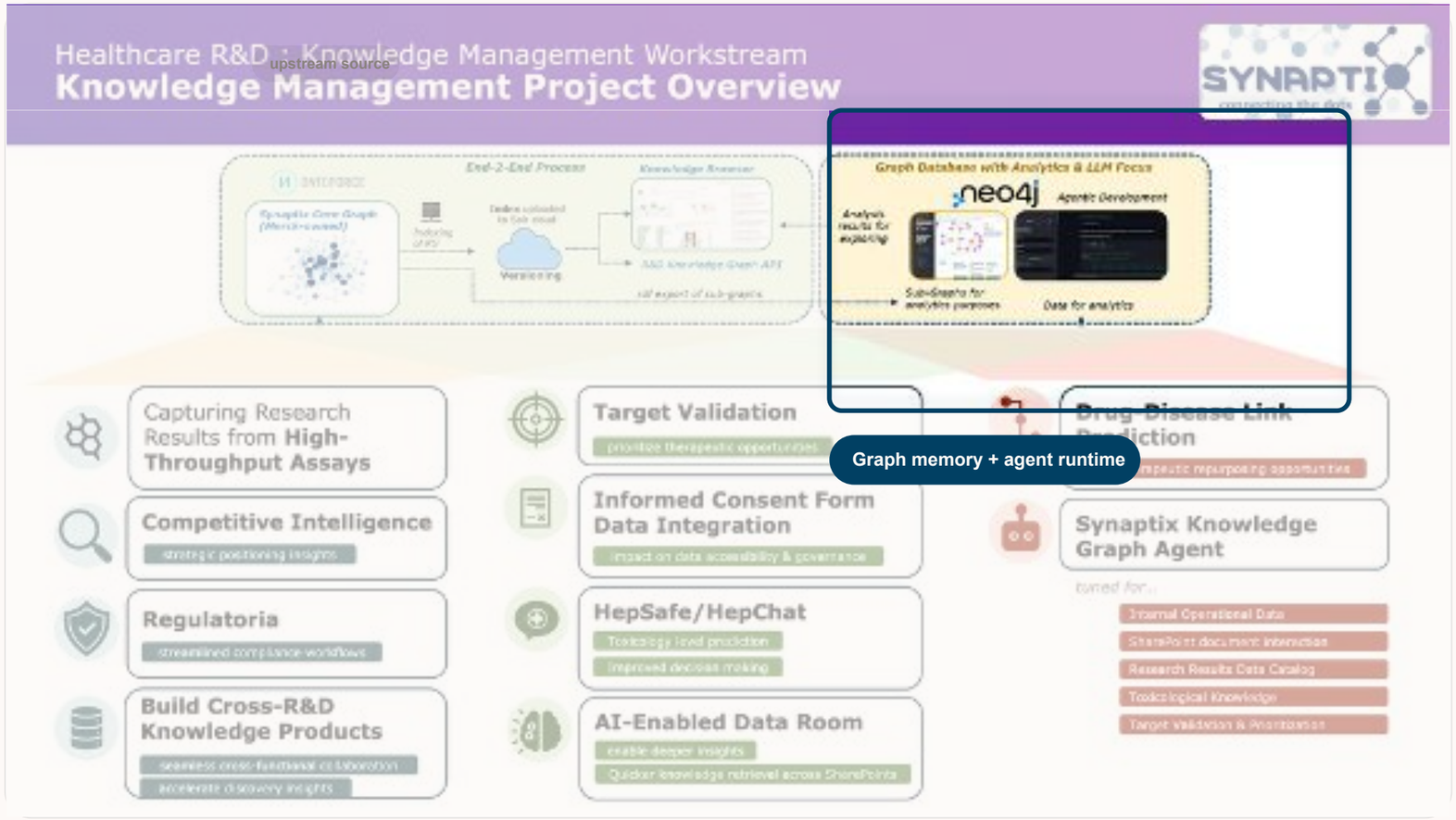
The what, why, and how — held as relationships, not scattered files.

Analytics & retrieval, in place

LLM-focused queries run against the graph — not a stale copy of it.

Where agents operate

Agentic development plugs straight in — durable context on every invocation.



Upstream sources feed the Neo4j graph.

PROVEN WINS · MERCK

Winning with graph + agents

> Put context graph layers under the agents and the system gets **grounded, connected, and accountable** — not just fast.

Five wins, from Merck's Healthcare R&D team.

1. **Grounding agentic behavior in trusted context**
   Agents act on facts you can vouch for — not plausible guesses.

2. **Dynamic decision-making needs relationships**
   Decisions follow the edges between entities, in real time.

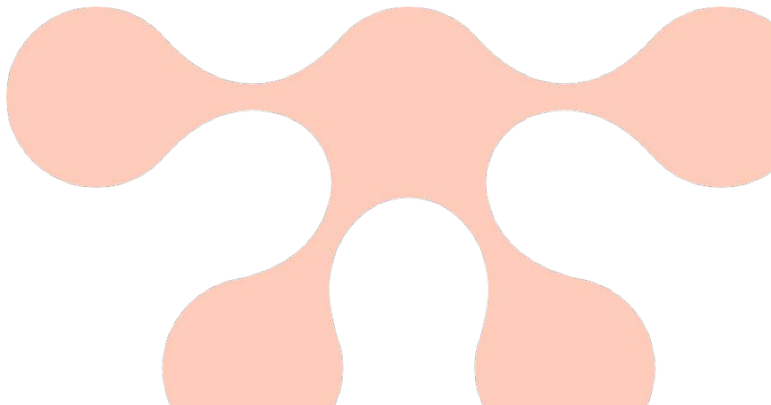
3. **Memory + context + reusability**
   Capture it once; every agent and workflow reuses it.

4. **A data integration backbone**
   One connected layer across otherwise siloed R&D sources.

5. **Transparency + traceability**
   Every answer traces back to the evidence behind it.

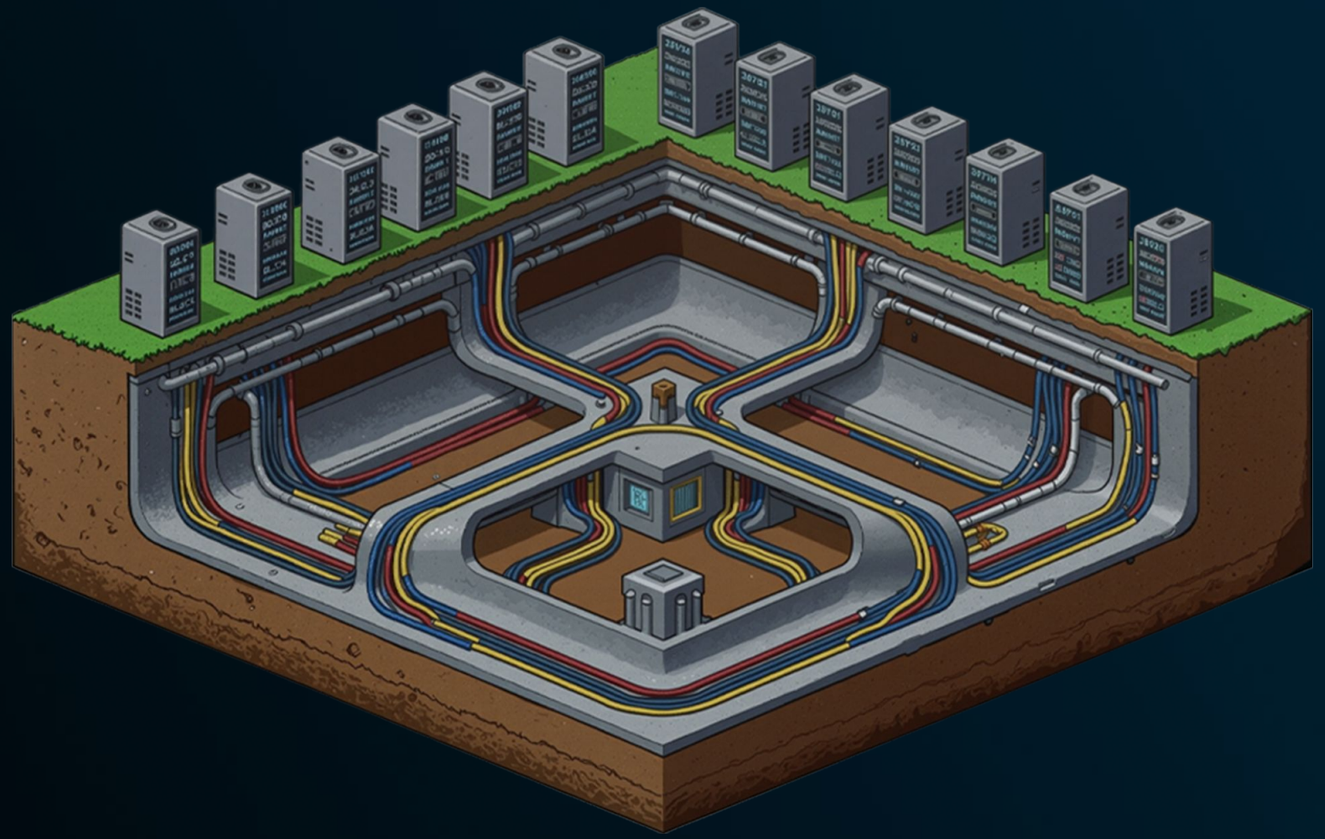
PART SEVEN

Thinking simply, in layers



Base Layer:

Enterprise Knowledge Graph



Mid Layer:
Conversational Graph

Base Layer:
Enterprise Knowledge Graph



Top Layer:
Decision Graph

Mid Layer:
Conversational Graph

Base Layer:
Enterprise Knowledge Graph



Context Graph

Top Layer:
Decision Graph

Mid Layer:
Conversational Graph

Base Layer:
Enterprise Knowledge Graph

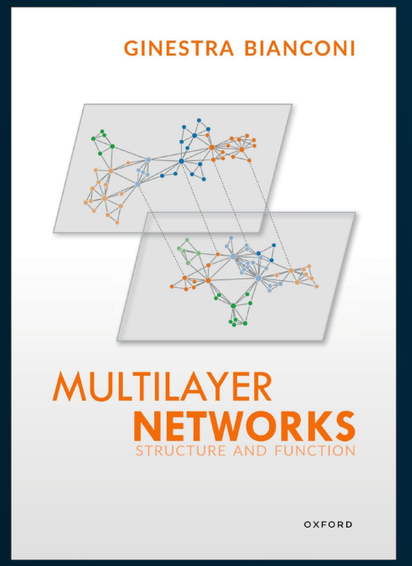


Top Layer:
Decision Graph

Mid Layer:
Conversational Graph

Base Layer:
Enterprise Knowledge Graph

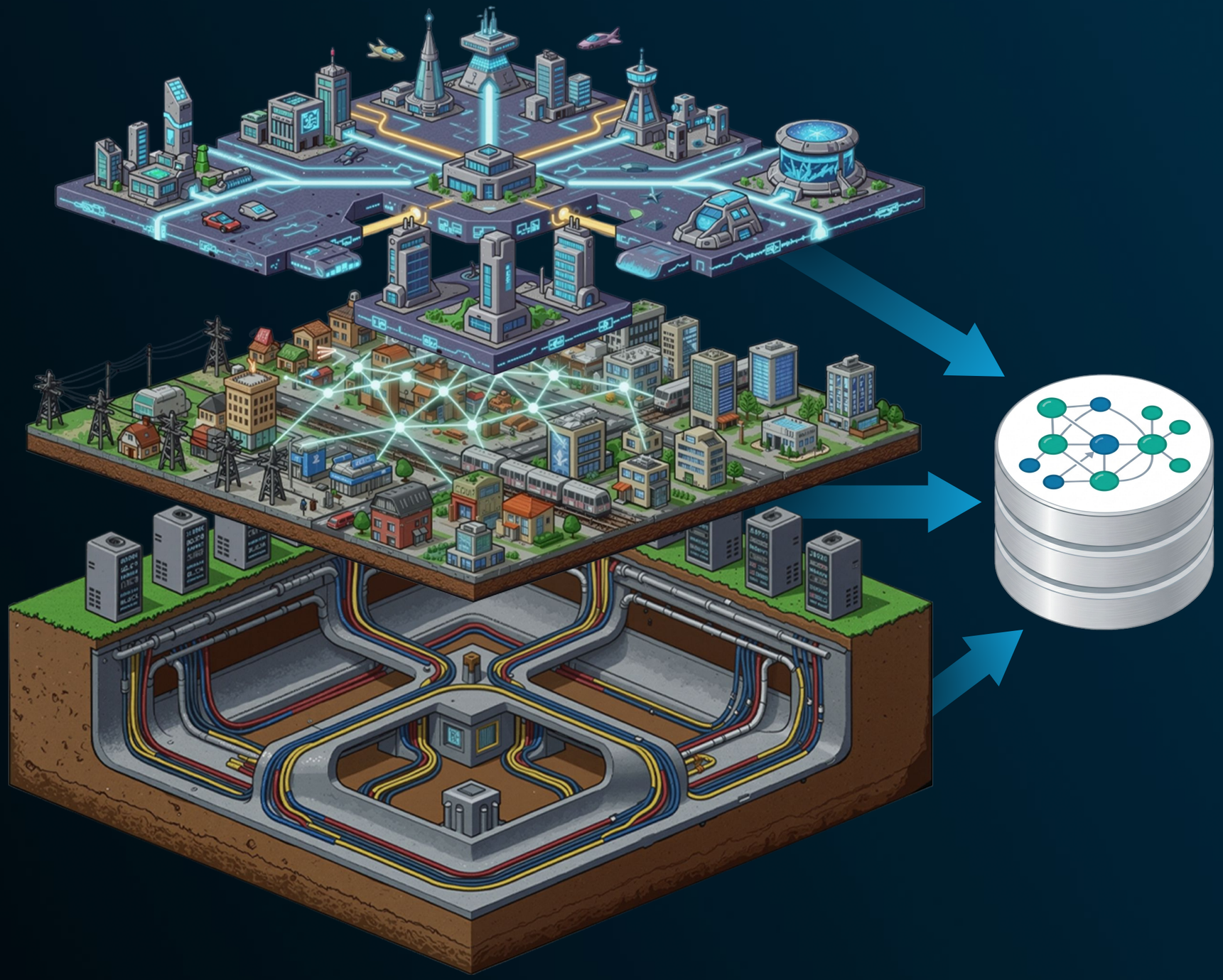




Top Layer:
Decision Graph

Mid Layer:
Conversational Graph

Base Layer:
Enterprise Knowledge Graph



PART EIGHT

Summary

Graph RAG



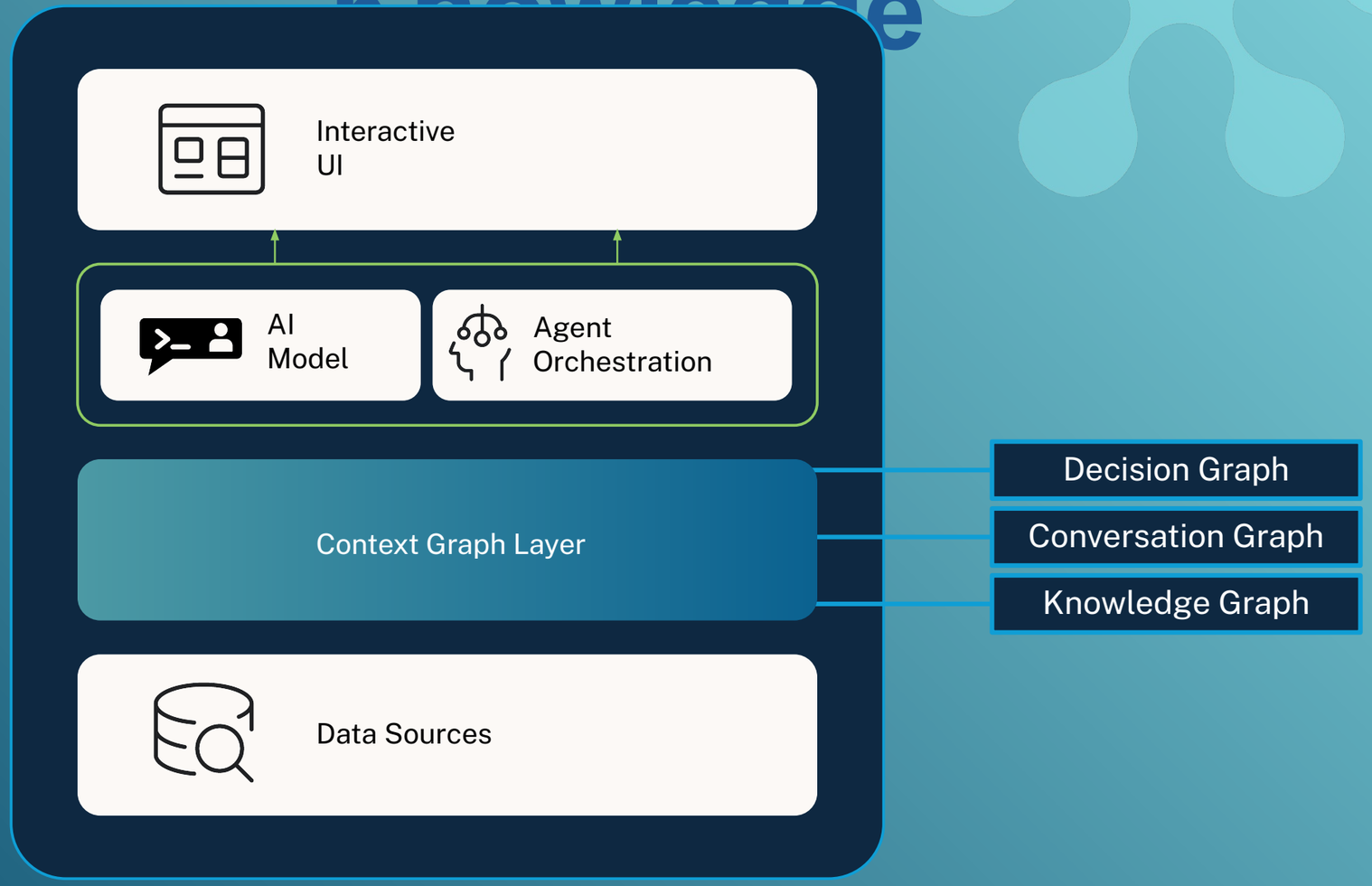
- LLMs are a lossy compression of the internet as it was
- Knowledge graphs provide facts to address the loss

Context is King



- Knowledge graphs provide facts to address the loss
- Conversation graphs provide information you care about now
- Decision graphs record what was done and why, enriching future context and knowledge

The critical layer in the AI Enterprise is Contextual Knowledge



Resources

Create Context Graph Source



github.com/neo4j-labs/create-context-graph

Create Context Graph Demo

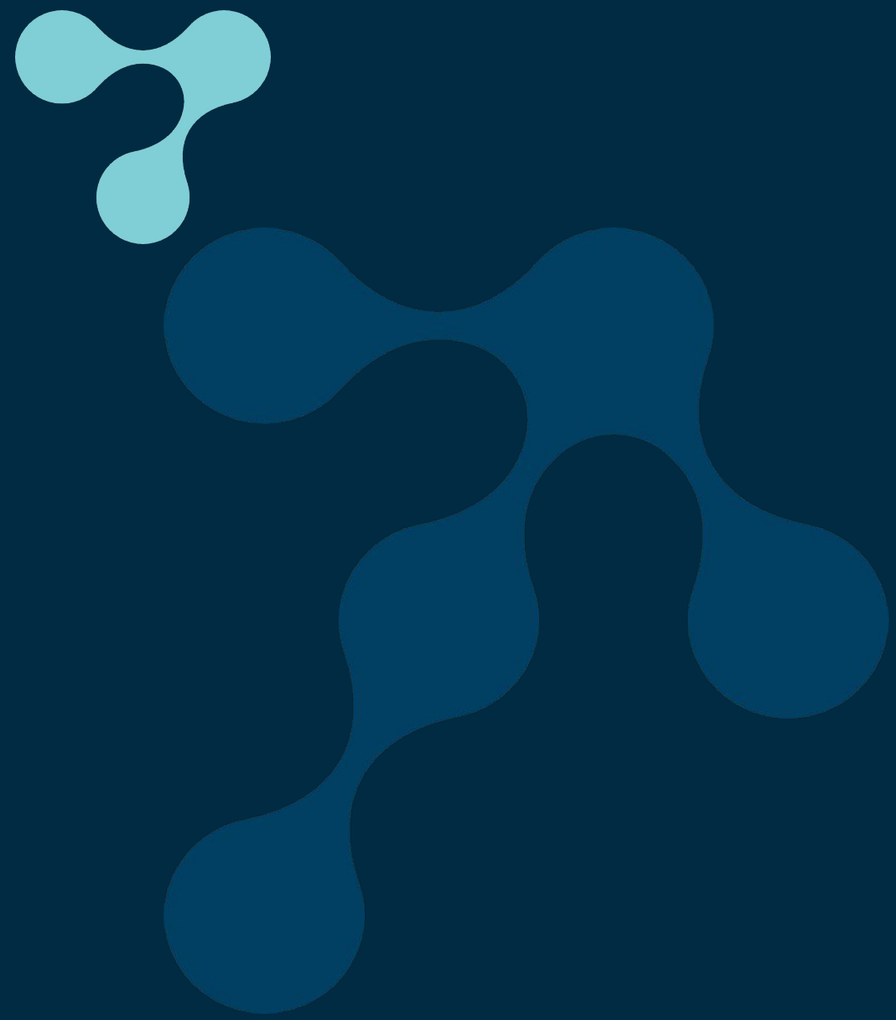


create-context-graph.dev

Neo4j Agent memory



github.com/neo4j-labs/agent-memory



Thank you for listening

Enjoy your day with us.

Jim Webber, Ph.D.
Chief Scientist | Neo4j