



"The AI Said So?": How to Build Auditable AI Agents Using Context Graphs

Nyah Macklin

Sr. Developer Advocate, Artificial Intelligence



Nyah Macklin

Developer Relations @ Neo4j



- Bootcamp graduate - resilientcoders.org | 100devs
- Graph Theory, Graph Algorithms, & Context Engineering
- Ethics, Governance, Safety, Responsible A.I.
- Taught at universities, high schools, and nonprofits

Learn more with



- **Completely Free**
- **Hands-on Courses**
teaching Neo4j Fundamentals, Cypher, Drivers and Graph Data Science
- **Curated Learning Paths** catering for everyone from beginners to experts
- **Free Certifications**

graphacademy.neo4j.com



Introduction to Neo4j Graph Data Science

Gain a high-level technical understanding of the Neo4j Graph Data Science (GDS) library

30 minutes



Neo4j Graph Data Science Fundamentals

Learn all you need to know about Graph Algorithms and Machine Learning Pipelines

1 hour



Path Finding with

Learn how to find the shortest paths between pairs of nodes in the graph

1 hour



Building Neo4j Applications with .NET

Learn how to interact with Neo4j from .NET using the Neo4j .NET Driver



Building Neo4j Applications with Go

Learn how to interact with Neo4j from your Go application using the Neo4j Go Driver

Building Neo4j Applications with TypeScript

Learn how to interact with Neo4j in your TypeScript project using the Neo4j JavaScript Driver

What do I need?

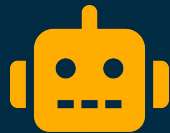


Workshop materials:

- graphacademy.neo4j.com account
- graphacademy.neo4j.com/courses/genai-fundamentals



Question



What happens
when *dozens* of AI Agents
have to work together?

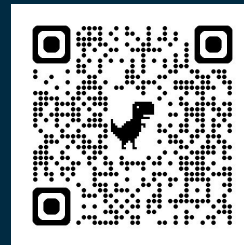




95%
AI pilots
fail



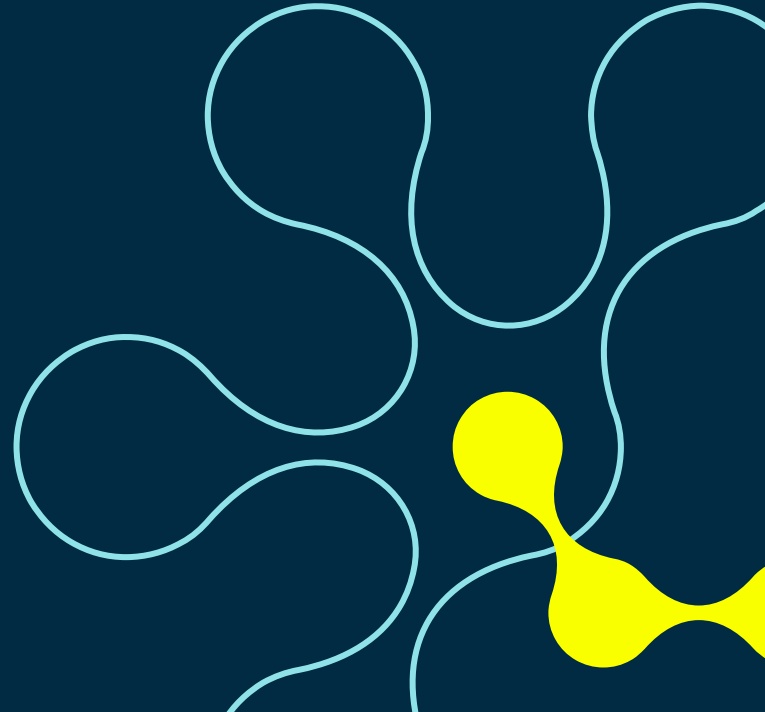
Massachusetts
Institute of
Technology





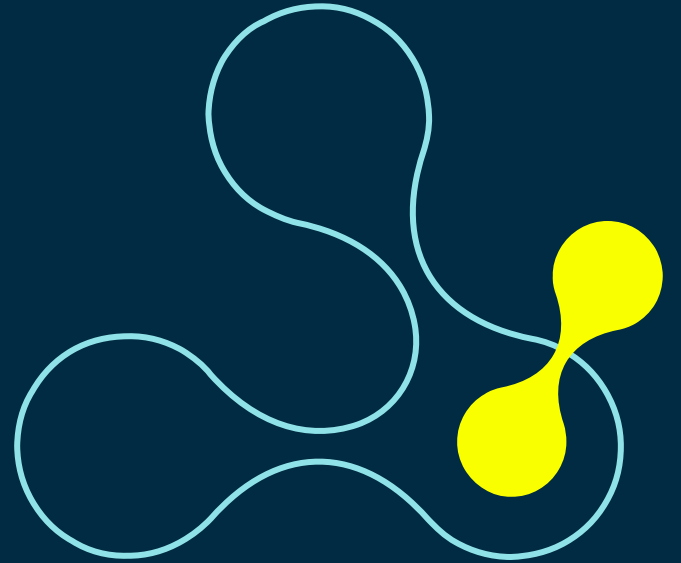


The past





Context Engineering



Context Engineering



Knowledge
Graph-Augmented
Context

RAG + Hybrid
Search



Structure and
Order of Context

Memory
Management



Tools and Function
Calling as Context

RAG + Hybrid Search



HOW DO YOU DO, FELLOW KIDS?



Memory Management



Structure and Order Context



Tools and Function Calling

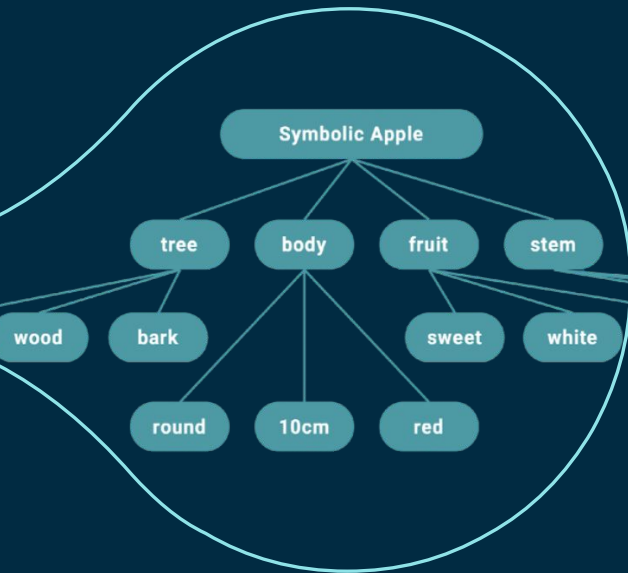


Knowledge Graph Augmented Context

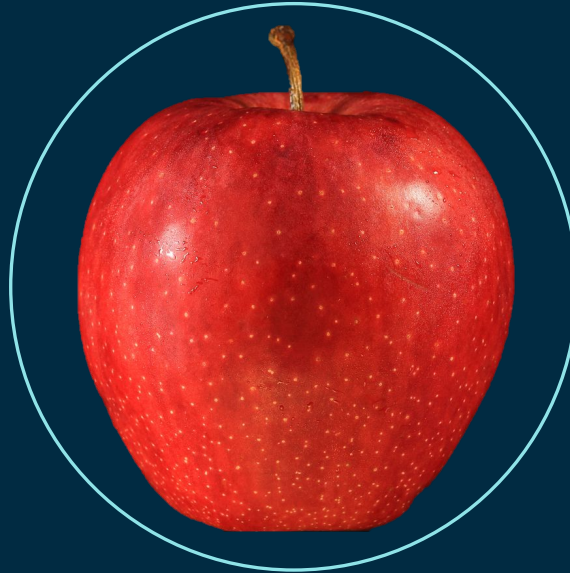


To Overcome AI's Black Box Problem *We Need Knowledge to be Transparent*

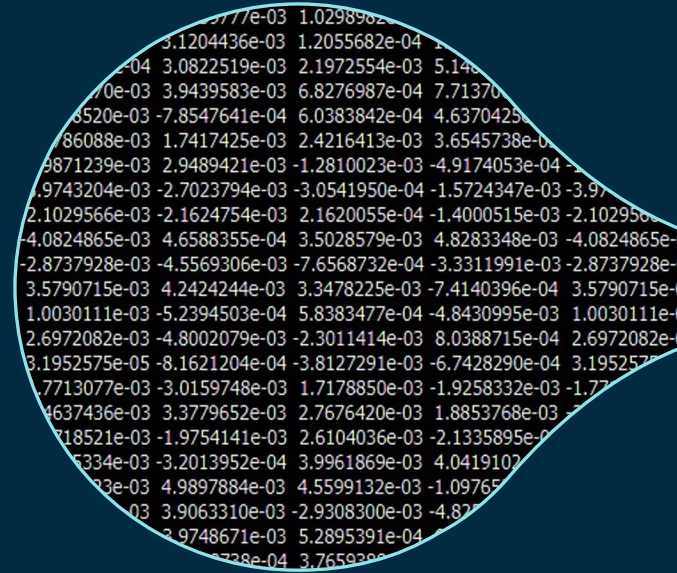
Knowledge Graph View
of an Apple



Human View
of an Apple



Vector View
of an Apple

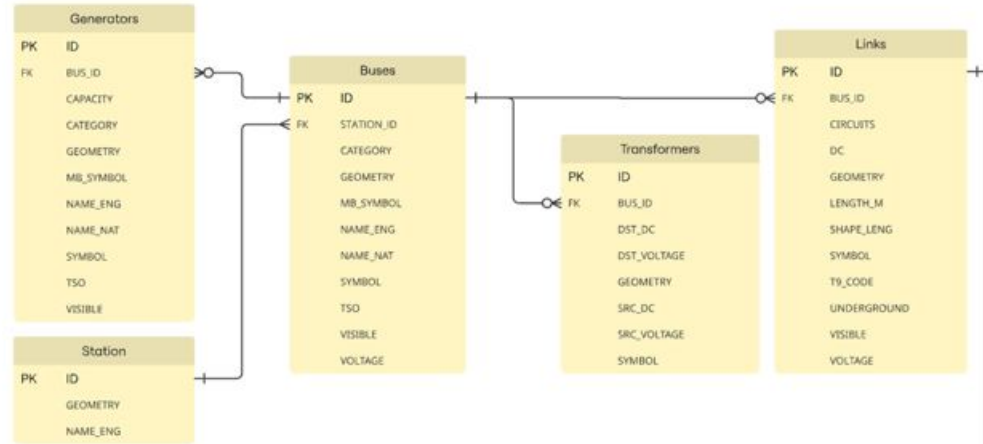


Network Data

Unstructured PDF & JSON

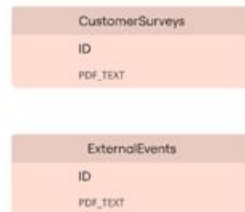


Structured (SQL)

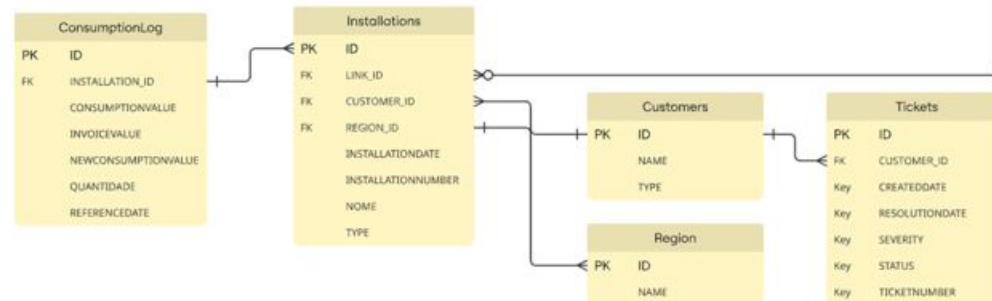


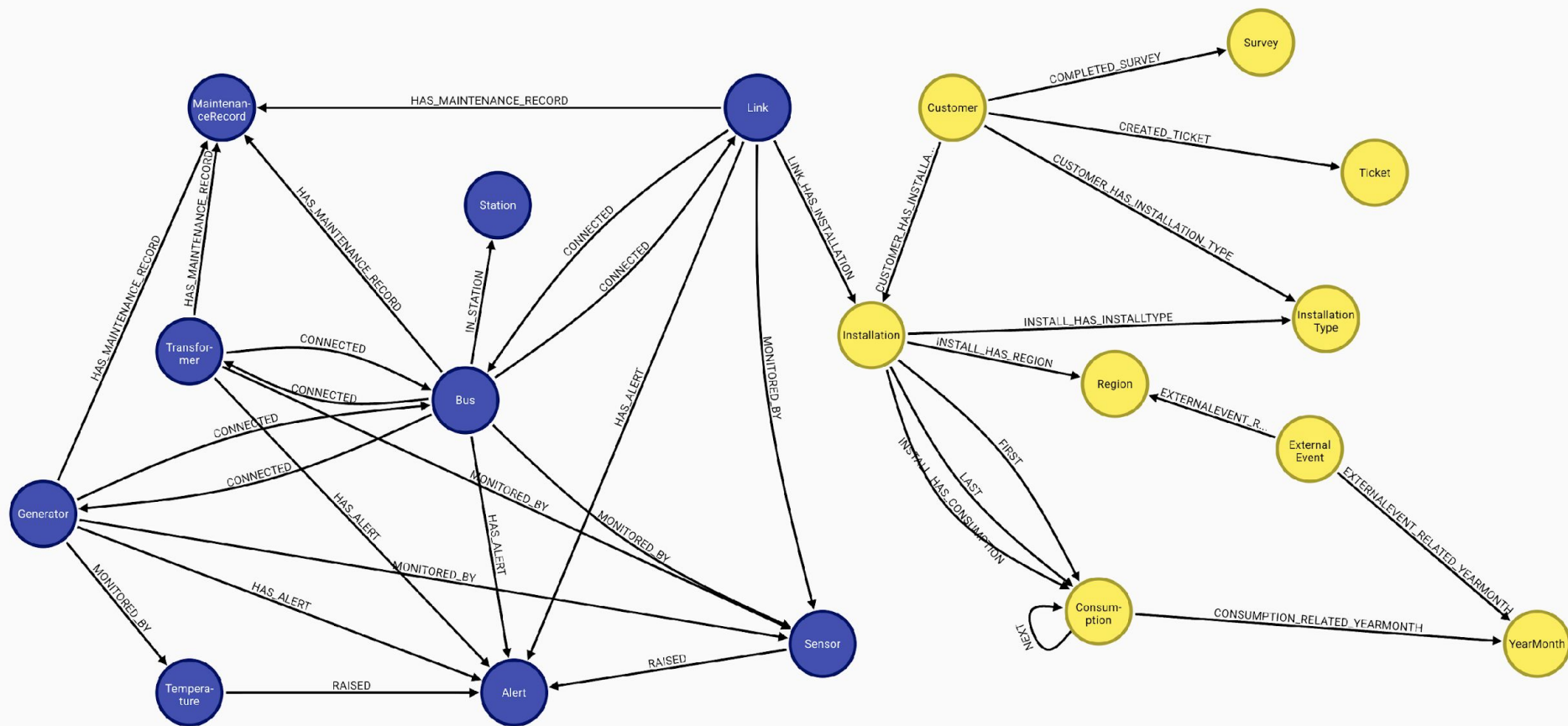
Customer Data

Unstructured PDF

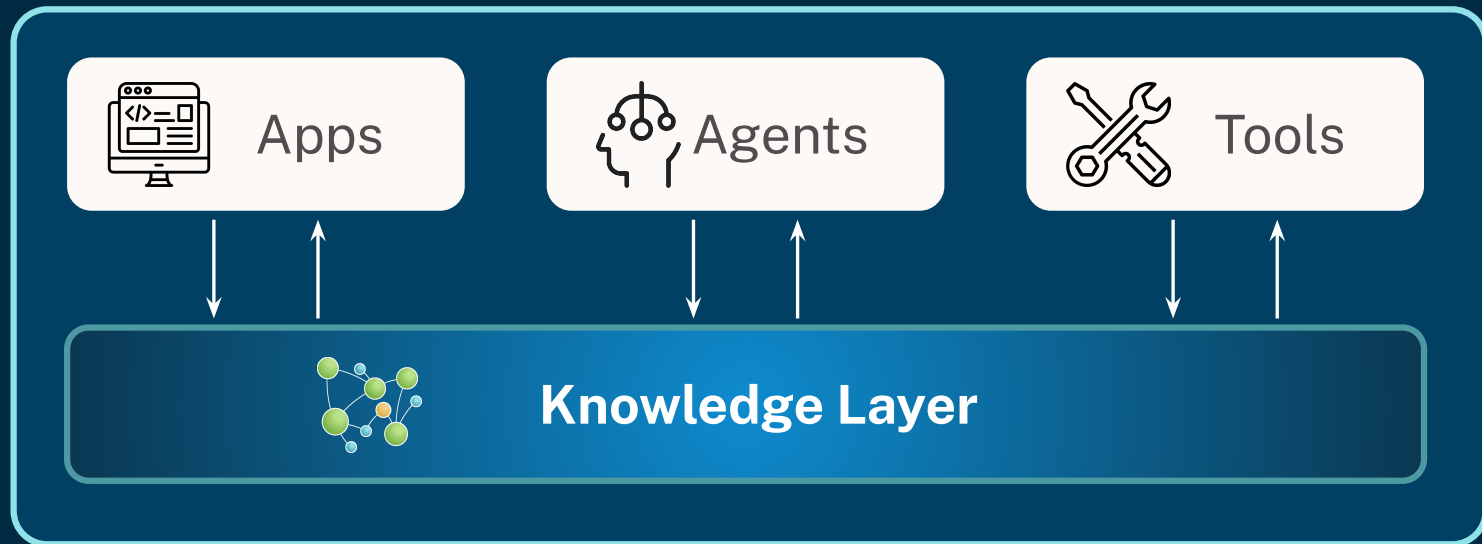


Structured (SQL)



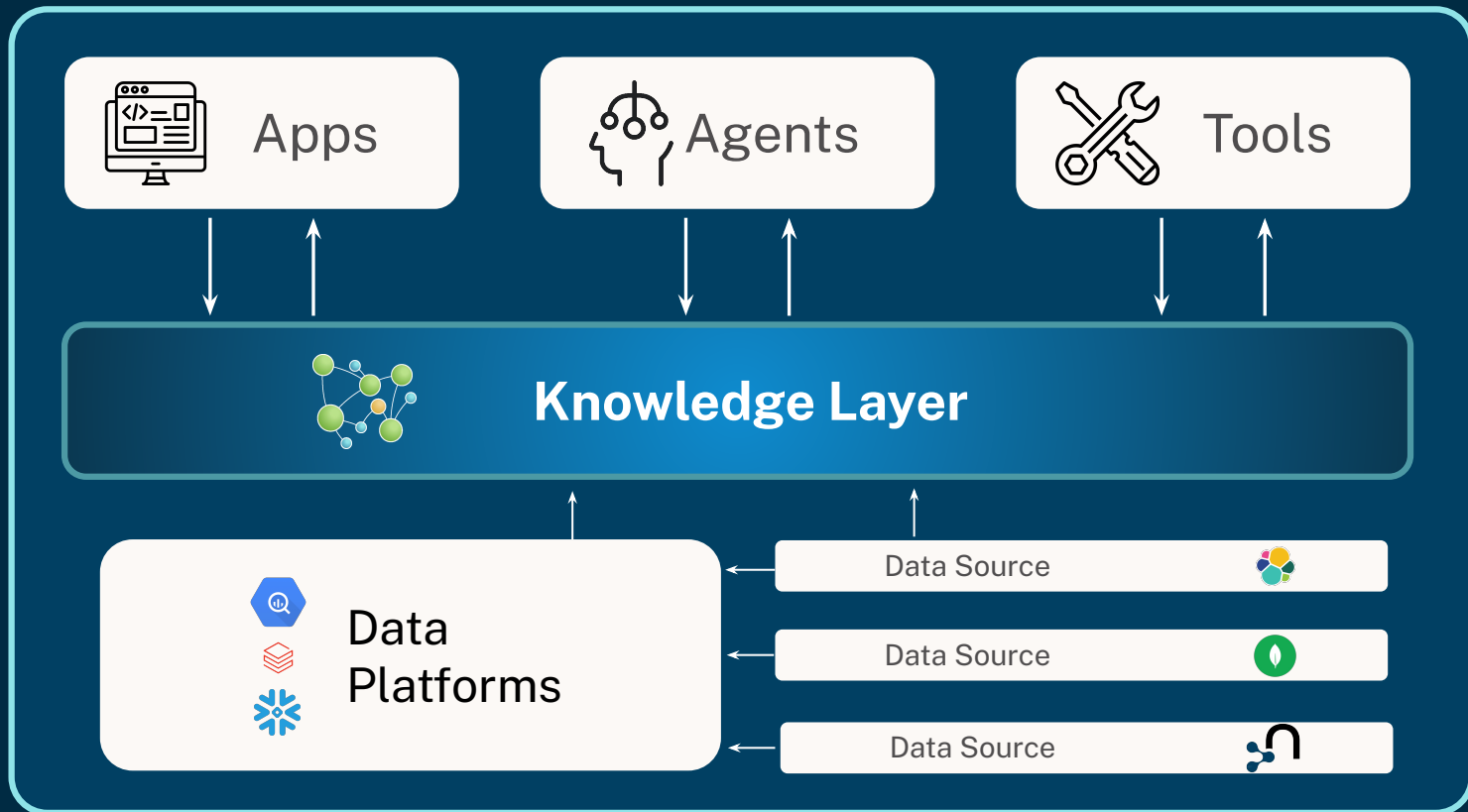


The Next Layer in the AI Stack: Knowledge



- Holds structure, context, and meaning
- Bridges human and machine understanding
- Uniform memory & retrieval layer for agents
- Links and unifies data sources and objects

The Next Layer in the AI Stack: Knowledge





Hands on Demo

What do I need?



Link!

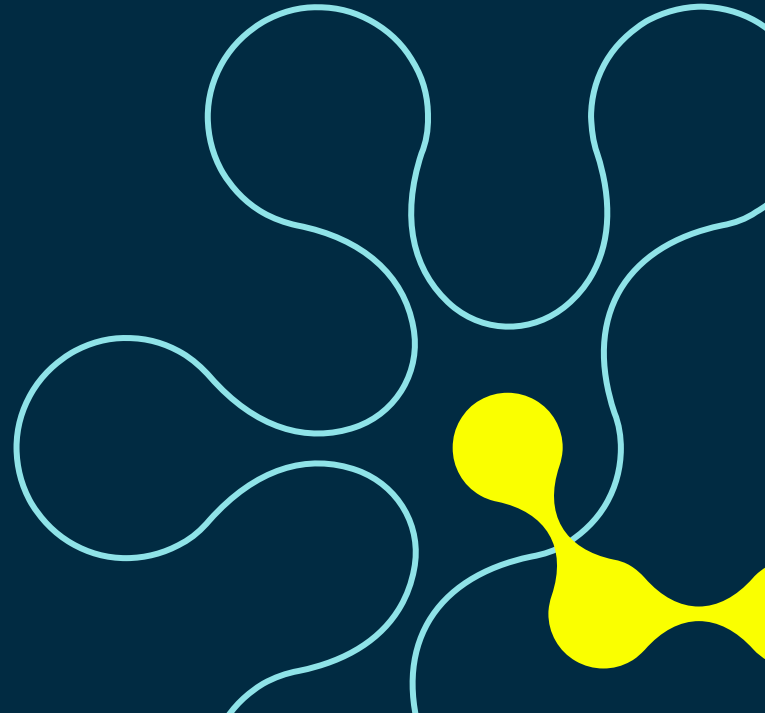
- <https://dev.neo4j.com/workshop-genai-fundamentals-notes>

Workshop materials:

- graphacademy.neo4j.com account
- graphacademy.neo4j.com/courses/genai-fundamentals



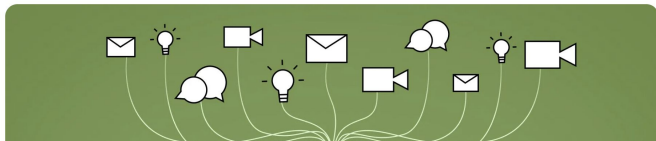
The present + future





AI's trillion-dollar opportunity: Context graphs

12.22.2025 | By: Jaya Gupta, Ashu Gang



<https://foundationcapital.com/context-graphs-ais-trillion-dollar-opportunity/>

arXiv:2406.11160v3 [cs.AI] 28 Jun 2024

Context Graph

Chengjin Xu¹, Cehao Yang¹, Xuhui Jiang^{2,3}, Luningyuan Tang¹, Yiyao Qi¹,
Jian Guo^{1*}

1. IDEA Research, International Digital Economy Academy

2. Department of Computer Science and Engineering, The Chinese University of Hong Kong

3. CAS Key Laboratory of AI Safety, Institute of Computing Technology, CAS
{xuchengjin, linuzhi, yangcehao, jiangxuhui, guojian}@idea.edu.cn

Abstract

Knowledge Graphs (KGs) are foundational structures in many AI applications, representing entities and their interrelations through triples. However, triple-based KGs lack the contextual information of relational knowledge, like temporal dynamics and provenance details, which are crucial for comprehensive knowledge representation and effective reasoning. Instead, Context Graphs (CGs) expand upon the conventional structure by incorporating additional information such as time validity, geographic location, and source provenance. This integration provides a more nuanced and accurate understanding of knowledge, enabling KGs to offer richer insights and support more sophisticated reasoning processes. In this work, we first discuss the inherent limitations of triple-based KGs and introduce the concept of CGs, highlighting their advantages in knowledge representation and reasoning. We then present a context graph reasoning (CGR) paradigm that leverages large language models (LLMs) to retrieve candidate entities and related contexts, rank them based on the retrieved information, and reason whether sufficient information has been obtained to answer a query. Our experimental results demonstrate that CGR significantly improves performance on KG completion (KGQ) and KG question answering (KGQA) tasks, validating the effectiveness of incorporating contextual information on KG representation and reasoning.

understand and reason about knowledge, thereby supporting various intelligent applications such as question answering (Sun et al., 2024), semantic analysis (Wang and Shi, 2023), recommendation systems (Wang et al., 2019), and more. While this triple-based structure offers clear semantics and precision through the use of schemas and ontologies, it loses the contextual information of knowledge and falls short in capturing the complexity and richness of real-world knowledge (Dong, 2023). Since we cannot clearly model the knowledge in a domain only with entities and relations, many recent KGs (Pellissier-Tanon et al., 2020; Thirani, 2021) are designed to be semi-structured: they leverage the clear semantics of structured data provided by the rigidity of schemas (i.e., ontologies) while also embracing the flexibility of unstructured data. Such KGs integrate multi-modal knowledge, including entity descriptions, images, timestamps and other metadata, all of which can be regarded as the contexts of triple knowledge. In this paper, we refer to this type of KGs as contextual graphs (CGs). By incorporating these semantic contexts, CGs provide a more comprehensive and nuanced representation of knowledge, extending beyond the traditional triple-based approach. This enables KGs to possess more advanced capabilities in knowledge representation and reasoning.

Moreover, large language models (LLMs), pretrained on vast text corpora, have exhibited strong semantic understanding capability (Brown et al., 2020a). And the use of LLMs for KG reasoning has become a research hotspot (Wei et al., 2023; Liao et al., 2024; Sun et al., 2024). However, KGs

1 Introduction

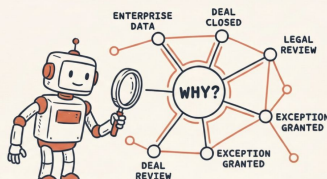
Knowledge Graphs (KGs) are structured knowledge bases (KBs) that organize factual knowledge



Graphs, Graphs Everywhere—with Sudhir Hasbe (Ep 43)

<https://www.youtube.com/watch?v=iO4rWqj1Grs>

CONTEXT GRAPHS: CAPTURING THE 'WHY'



@dharmesh

Context Graphs: Capturing The "Why" In The Age of AI



Dharmesh Shah

Founder and CTO at HubSpot. Helping millions grow better.



<https://www.linkedin.com/pulse/context-graphs-capturing-why-age-ai-dharmesh-shah-oyvze/>

Hands On With Context Graphs And Neo4j

Building AI systems that don't just make decisions, but remember and use the "why".



William Lyon 9 min read · Jan 14, 2026

<https://neo4j.com/blog/agent-ai/hands-on-with-context-graphs-and-neo4j/>

The Missing "Why"

When an AI agent recommends approving a \$100K credit line increase, the question isn't just "is this the right answer?"



No Memory

Ask an agent about last week's conversation — blank stare



No Audit Trail

When something goes wrong, no one knows why the agent decided that



No Shared Learning

Deploy multiple agents and they can't share what they've learned

95% of AI pilots fail to deliver returns — MIT

What is a Context Graph?

A knowledge graph specifically designed to capture decision traces — the full context, reasoning, and causal relationships behind every significant decision.

Traditional Audit Log

Records actions only

"Transaction rejected at 14:32"

No relationships between events

No causal chain or reasoning

Flat, disconnected records

Context Graph

Captures the full "why"

Decision traces & causal chains

Entities, relationships, events

Tribal knowledge made queryable

Connected, traversable structure

Jaya Gupta & Ashu Garg, *"AI's Trillion-Dollar Opportunity: Context Graphs"* — Foundation Capital, Dec 2025

Context Graph: Financial Services Data Model

ENTITIES (What exists)

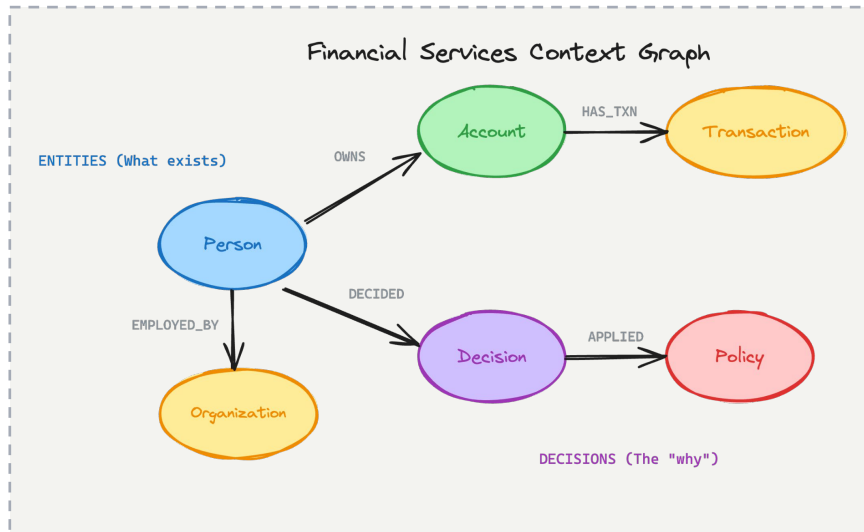
Person, Account, Transaction, Organization

EVENTS (What happened)

Decisions, Transactions, Approvals, Rejections

CONTEXT (The why)

Policies applied, Risk factors, Employee reasoning



Blog: neo4j.com/blog/agent-ai/hands-on-with-context-graphs-and-neo4j/



Context Graph Demo

<https://context-graph-demo.vercel.app/>



AI Assistant

Ask questions about customers, decisions, and policies

Recommendation

⚠️ **ESCALATE FOR MANUAL REVIEW** - Do not approve automatically

Reasoning: While Jessica Norris has strong financial indicators (substantial account balance, low personal risk score, clean transaction history), there are significant geopolitical and compliance concerns that require human judgment:

The employer watchlist status requires enhanced due diligence under sanctions compliance protocols

The open fraud investigation ticket needs resolution before extending additional credit

The recent trading escalation suggests ongoing risk assessment

The pending high-value wire transaction should be cleared first

Recommended Next Steps:

Verify Employment Status: Confirm current employment relationship with watchlisted organization

Resolve Fraud Ticket: Close out TKT000017 before proceeding

Clear Pending Transaction: Investigate the \$7,947.97 pending wire transfer

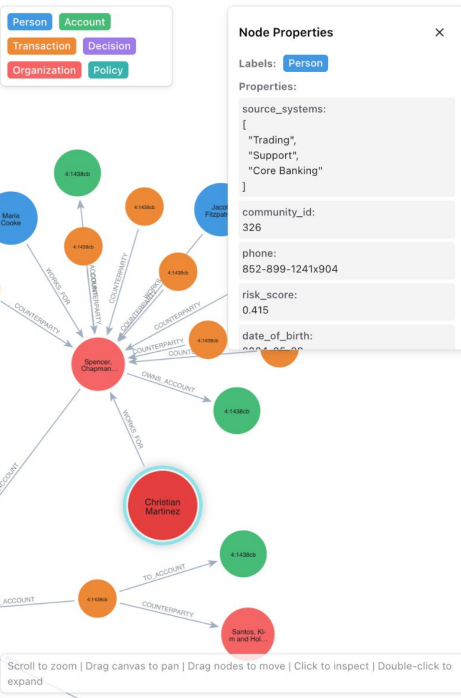
Enhanced Due Diligence: Conduct

Ask about customers, decisions, or policies...



Context Graph

Visualize entities, decisions, and causal relationships



Decision Trace

Inspect reasoning, precedents, and causal chains

← Back to list

escalation trading escalated

1/7/2026, 6:31:48 PM

ID: 52782694...

Reasoning

Decision made for trading case. Type: escalation. Standard review completed with confidence score of 0.94.

Confidence Score

76%

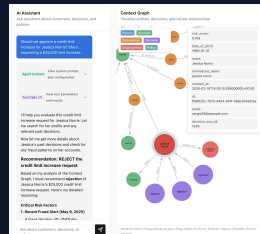
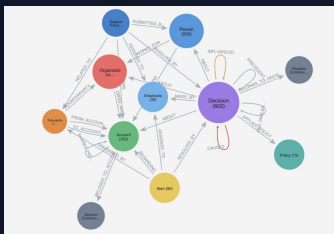
Causal Chain

No causal chain data.

Similar Decisions

No similar decisions found.

AI-Powered Decision Tracing for Financial Institutions



Chakra UI v3
NVL Graphs

Claude Agent
10 MCP Tools

Sample Data: 200 persons · 350 accounts · 2K transactions · 600 decisions · 50 orgs · 15 policies

Hybrid semantic + structural

Structural decision matching

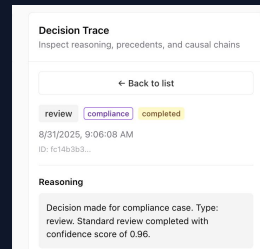
Neighborhood pattern matching

Decision cluster detection

Cause → effect tracing

Index new decisions

KEY INSIGHT: Hybrid search combines semantic similarity (text embeddings) with structural similarity (FastRP graph embeddings) — meaning + context, not just meaning alone.



3 Exception Request + Audit Trail Causal chain → justification & policy links

LATER **Government & Defense** Policy enforcement, grant decisions, audit

How the Agent Traces Decision History

REJECT or ESCALATE — grounded in institutional knowledge, not just a credit score

1 Query Context Graph

Agent retrieves Jessica's full profile: accounts, transactions, employer, risk tier

2 Trace Decision History

Past fraud flag (April 2025), compliance rejections, velocity check failures

3 Hybrid Search

Semantic similarity (text embeddings) + structural similarity (FastRP graph embeddings)

4 Recommendation

31 nodes · 30 relationships · Jessica's complete context in a single graph traversal

View tool parameters and results

Tool Calls (7)

Tool Call

search_customer

Arguments:

```
{
  "query": "Jessica Norris",
  "limit": 5
}
```

Tool Call

find_precedents

Arguments:

```
{
  "scenario": "credit limit increase $25000",
  "category": "credit",
  "limit": 10
}
```

Tool Call

detect_fraud_patterns

Arguments:

```
{
  "account_id": "6a5a92ab-7e1a-4990-9ca0-fe08574b2ddc",
  "neighbor_count": "10"
}
```

Recommendation

⚠️

ESCALATE FOR MANUAL REVIEW

-

Do not approve automatically

Reasoning: While Jessica Norris has strong financial indicators (substantial account balance, low personal risk score, clean transaction history), there are significant geopolitical and compliance concerns that require human judgment:

Hybrid Search: Vector + Graph

Neo4j GDS provides algorithms impossible in relational databases



Text Embeddings

Semantic similarity — what does the text mean?

Finds decisions with similar descriptions, rationale, and risk factors. "Fraud rejection" matches "suspicious activity alert."



Graph Embeddings (FastRP)

Structural similarity — how is this node connected?

Finds decisions involving customers with similar account structures, employer types, and transaction patterns.

Combined: find precedent decisions from structurally AND semantically similar situations



**Where do I learn
more?**



The Knowledge Graph Difference

3x
improvement

A **Knowledge Graph** improved the accuracy of LLM responses by an average of 3x.



Day-to-day Analytics

Analytics Type	Percentage
Day-to-day Analytics	71%
Strategic Analytics	25.5%
Ad-hoc Analytics	3.5%



Operational Analytics



Category	Percentage
Operational Analytics	37.4%
Operational Analytics	66.9%
Operational Analytics	37.4%

● LLM Accuracy Alone ● LLM Accuracy with a Knowledge Graph

PERFORMANCE

Knowledge Graph achieves unmatched accuracy

In a benchmarking study, Writer Knowledge Graph achieved top scores on RobustQA, which measures accuracy in open-domain question-answering, outperforming seven popular RAG approaches that use vector retrieval.

Learn more →

#1 RATED IN ACCURACY
OVER
86%
on
RobustQA

Brian Shi | Alfred Clemedtson | Zach Blumenfeld

Experiments: G-retriever with STaRK-Prime

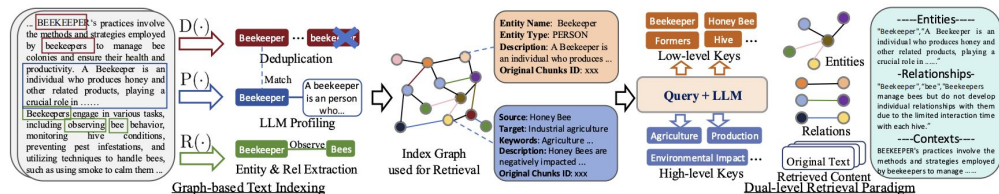
Method	Hits@1	Hits@5	Recall@20	MRR
Pipeline*	32.09	48.34	47.85	38.48
G-Retriever	32.27 $\pm$ 0.3	37.92 $\pm$ 0.2	27.16 $\pm$ 0.1	34.73 $\pm$ 0.3
Subgraph Pruning**	12.60	31.60	35.93	20.96
Baseline	15.57	33.42	39.09	24.11

*Pipeline adds nodes in subgraph context to G-retriever output.

**Frozen LLM. No fine-tuning.

Retrieval-Augmented Generation with Knowledge Graphs for Customer Service Question Answering

in BLEU. Our method has been deployed within LinkedIn’s customer service team for approximately six months and has reduced the median per-issue resolution time by 28.6%.



NODES AI 2026

The largest technical community gathering dedicated to Graph + AI

CONFERENCE DATE

April 15, 2026 - Leonhard Euler's Birthday

PRE-EVENT

Five 2 hour AI workshops

March 5, 11, 17, 26 and 31

EVENT

- One day free online conference
- Live technical sessions from community experts
- From: 10:00 am ET / 4:00 pm CET
To: 5:00 pm ET / 11:00 pm CET

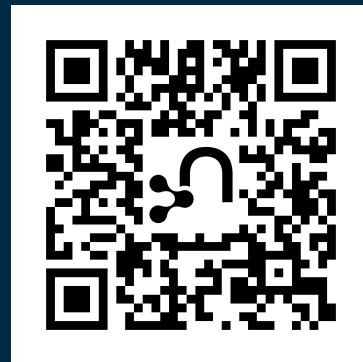
AI-FOCUSED SESSIONS IN THREE TRACKS

- Knowledge Graphs and GraphRAG
- Graph Memory and Agents
- Graph + AI in Production

WHO SHOULD PARTICIPATE

Anyone interested in Graphs + AI

REGISTER: neo4j.com/nodes-ai





Ground Controller Dances To Work





Thank you!

Nyah.Macklin@neo4j.com



Nyah Macklin

[linkedin.com/in/nyahmacklin](https://www.linkedin.com/in/nyahmacklin)

