



THE
DEVELOPER'S
CONFERENCE

TDC SP - Workshop

Combatendo corrupção com grafos

Eliézer Zarpelão - Sr Solutions Engineer - Neo4J

Agenda

1. Grafos
2. Neo4j
3. Hands-on
4. GenAI



Eliézer Zarpelão
Sr. Solutions Engineer

Astrazeneca

Identificar novas **relações moleculares em grandes conjuntos de dados** para acelerar a descoberta de medicamentos era uma barreira para a AstraZeneca.



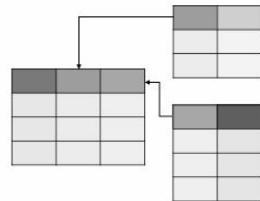


Caterpillar

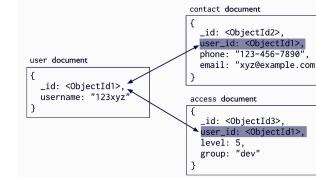
Gerar **insights** para reparo e manutenção eficientes de equipamentos em **dezenas de milhões de documentos técnicos** foi um enorme desafio para a Caterpillar.

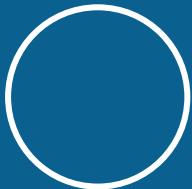
Sistemas legados não conseguem acompanhar...

BDs “relacionais” não lidam bem com relacionamentos.



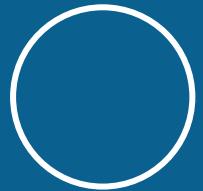
BDs NoSQL não lidam com relacionamentos de forma alguma.





É preciso uma nova abordagem...

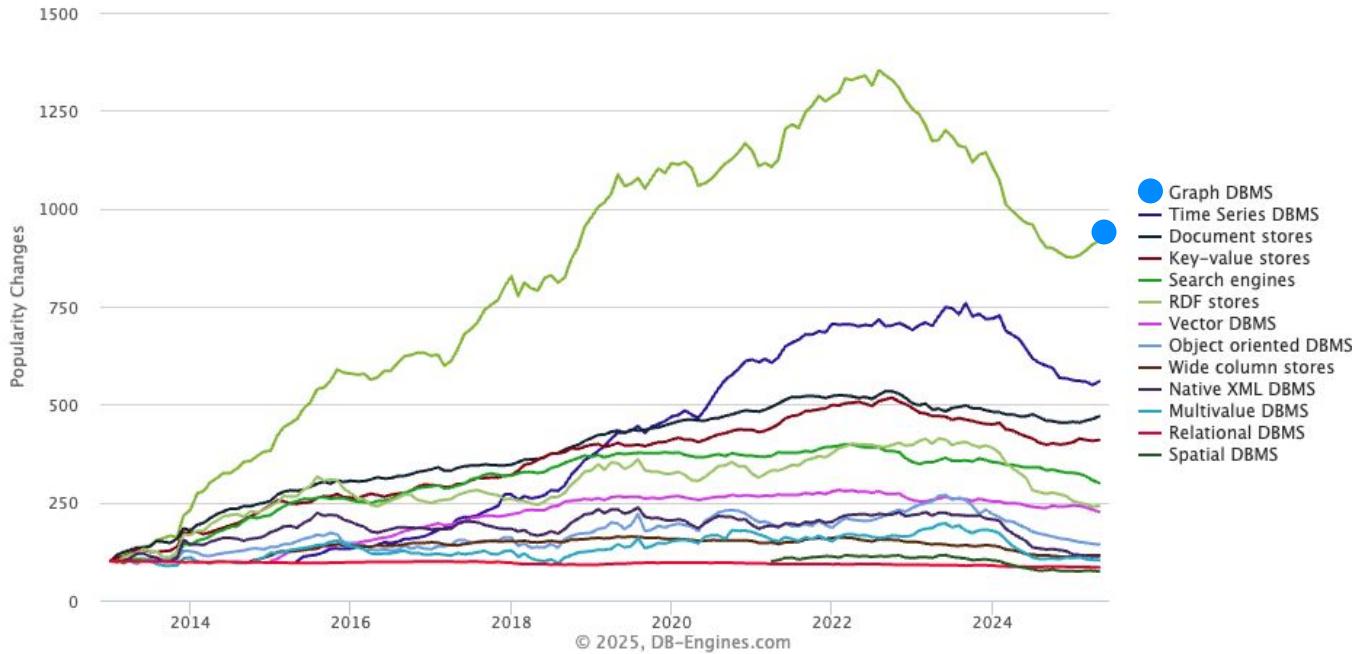




O Grafo cria uma visão mais intuitiva e conectada dos relacionamentos de dados, desbloqueando uma compreensão e um contexto mais profundos.



A categoria de banco de dados de crescimento mais rápido nos últimos anos



AI & ML impactando o interesse em grafos



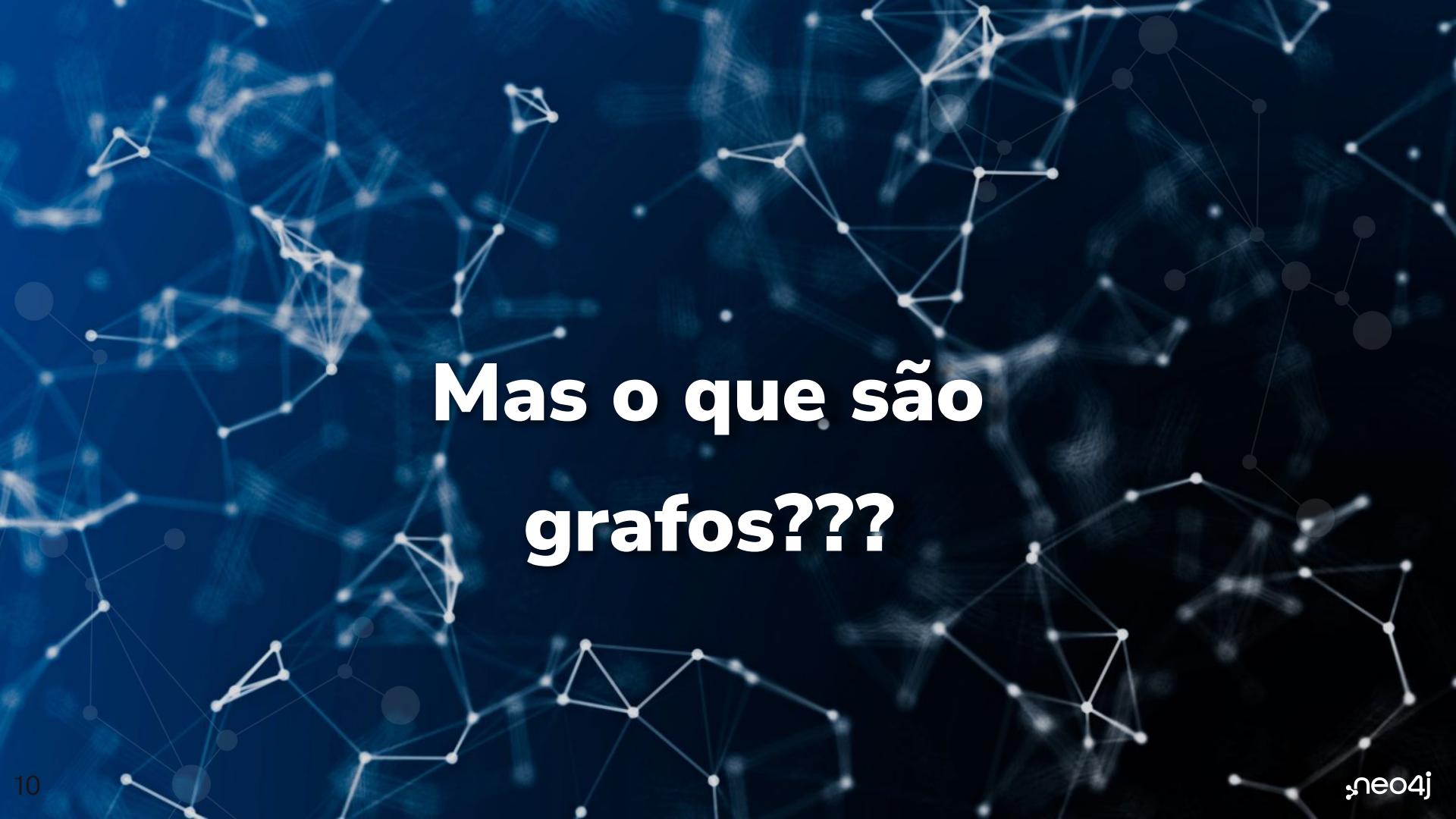
Source: Dimensions Knowledge System

“50% das pesquisas do Gartner sobre o tópico de AI envolvem a discussão do uso da tecnologia de grafos.”

Gartner®

"Até 2026, grafos serão usados em 80% das inovações em dados e analytics"

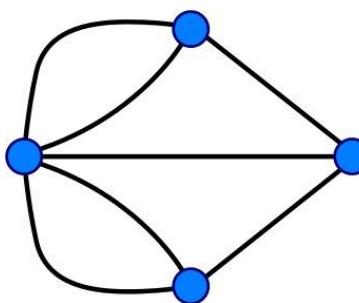
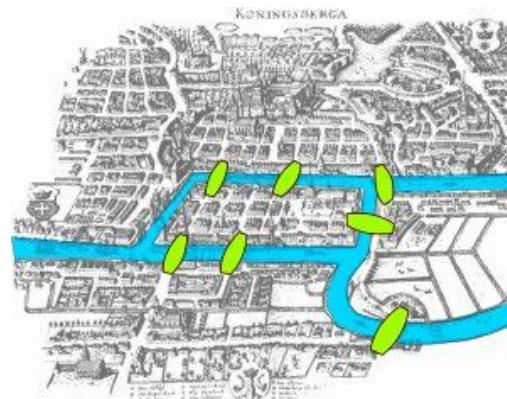
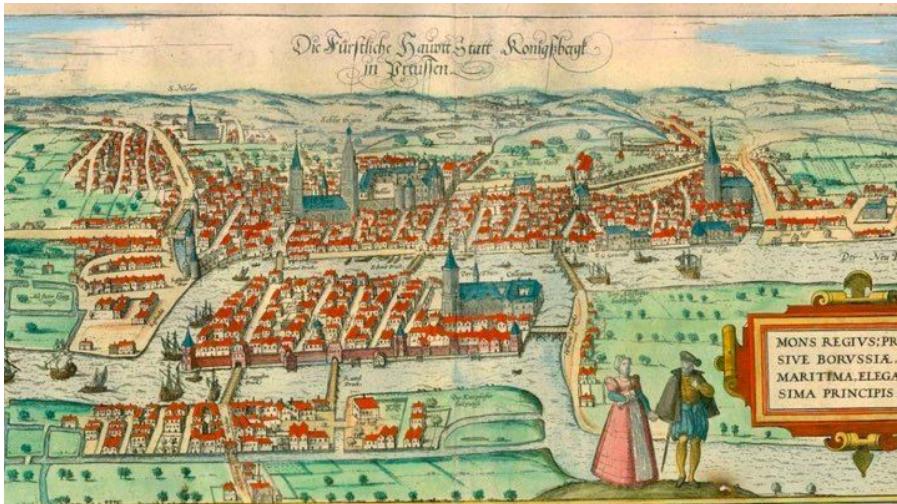
<https://www.gartner.com/en/newsroom/press-releases/2025-03-05-gartner-data-and-analytics-summit-2025-orlando-day-3-highlights>



Mas o que são
grafos???

Era uma vez Leonard Euler...

Königsberg, Prussia, 1736





**Você já usou Grafos
hoje?**



Labeled property graph

Nós podem ter zero ou mais rótulos/labels

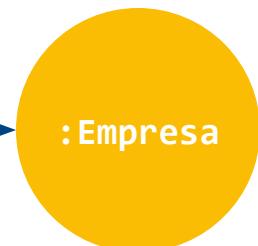


nome: Eliézer Zarpelão
nascimento: 19XX-03-01
matrícula: 14213

Nós e relacionamentos podem ter qualquer quantidade de propriedades

Relacionamentos tem um tipo

:TRABALHA_EM
inicio: 2023-09-18



nome: Neo4j

Relacionamento tem direção

:LOCALIZADA_EM



Relacionamentos representam como as coisas estão relacionadas, conectadas, montadas....

Nós representam entidades

ANÁLISE ATRAVÉS DE FILTROS

REI CHARLES

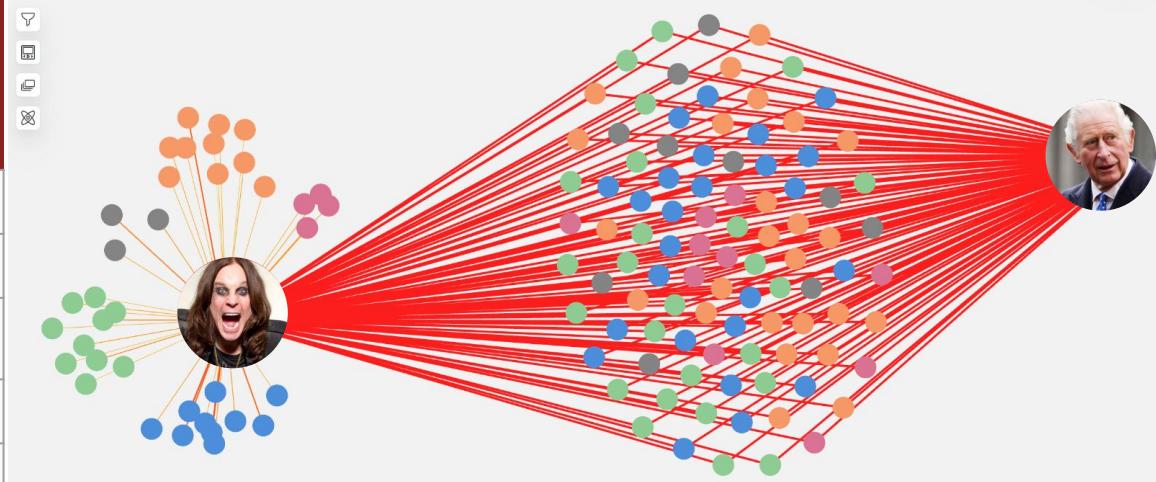


OZZY OSBOURNE

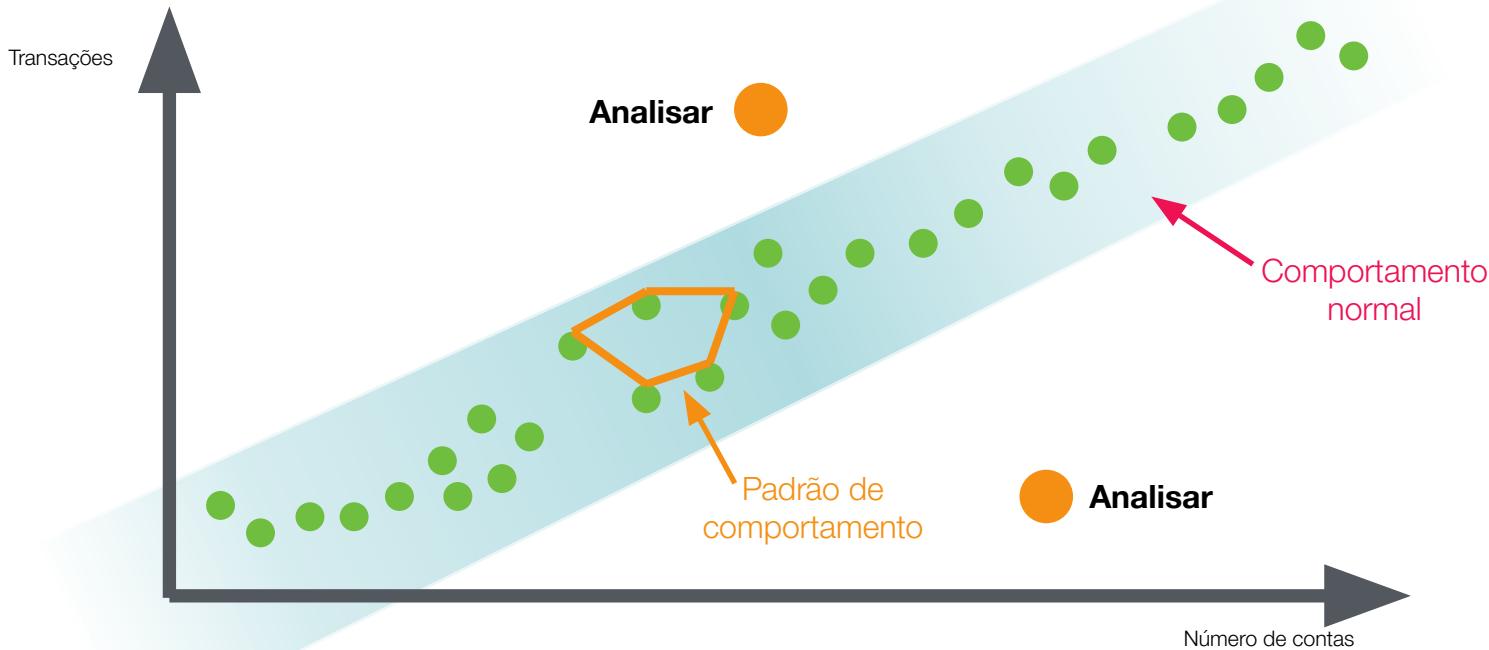


Homem	Homem
Nascido em 1948	Nascido em 1948
Residente no Reino Unido	Residente no Reino Unido
Casado 2 vezes	Casado 2 vezes
Mora em um castelo	Mora em um castelo
Classe A	Classe A
Grande poder aquisitivo	Grande poder aquisitivo
Conservador	Agressivo

ANÁLISE ATRAVÉS DE CONTEXTO



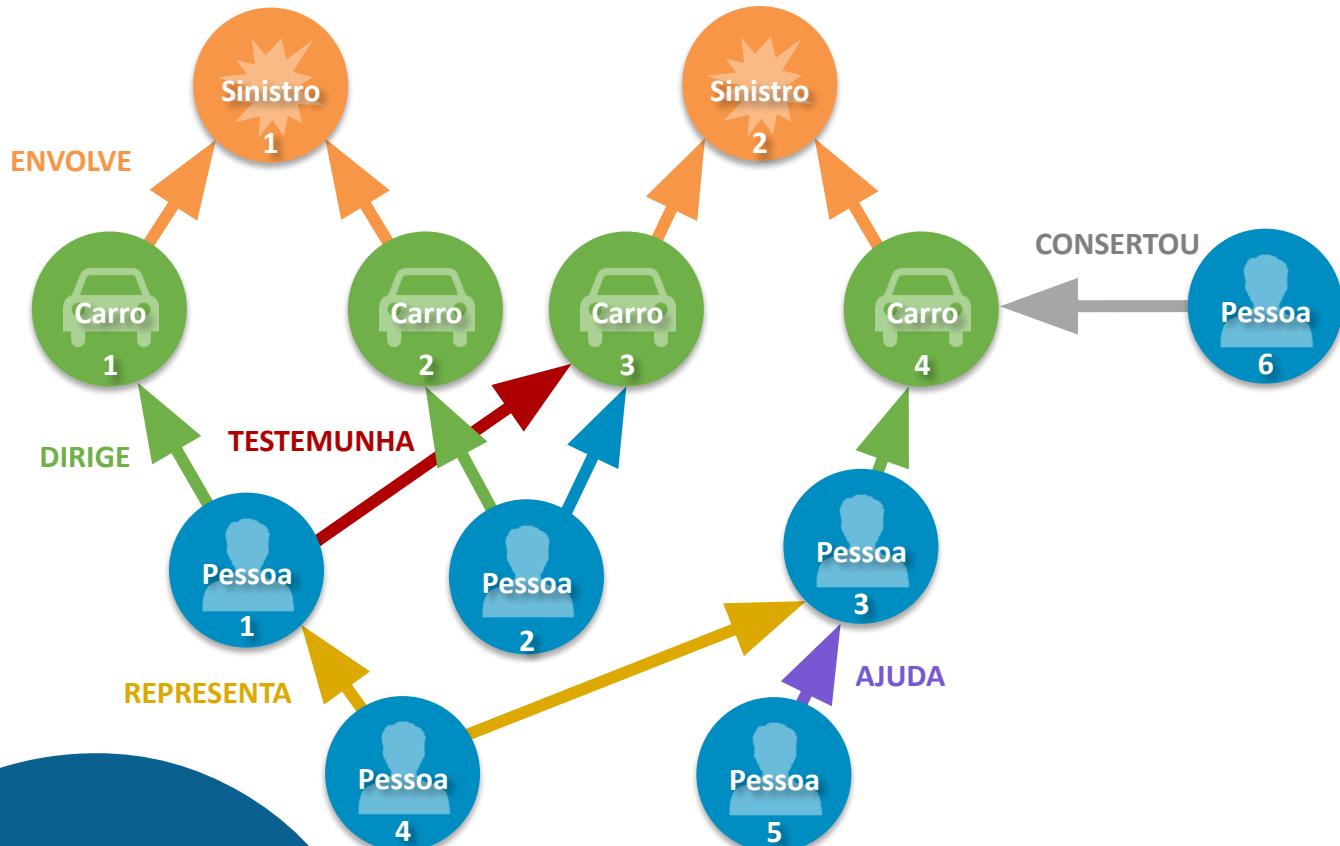
Análise de anomalias



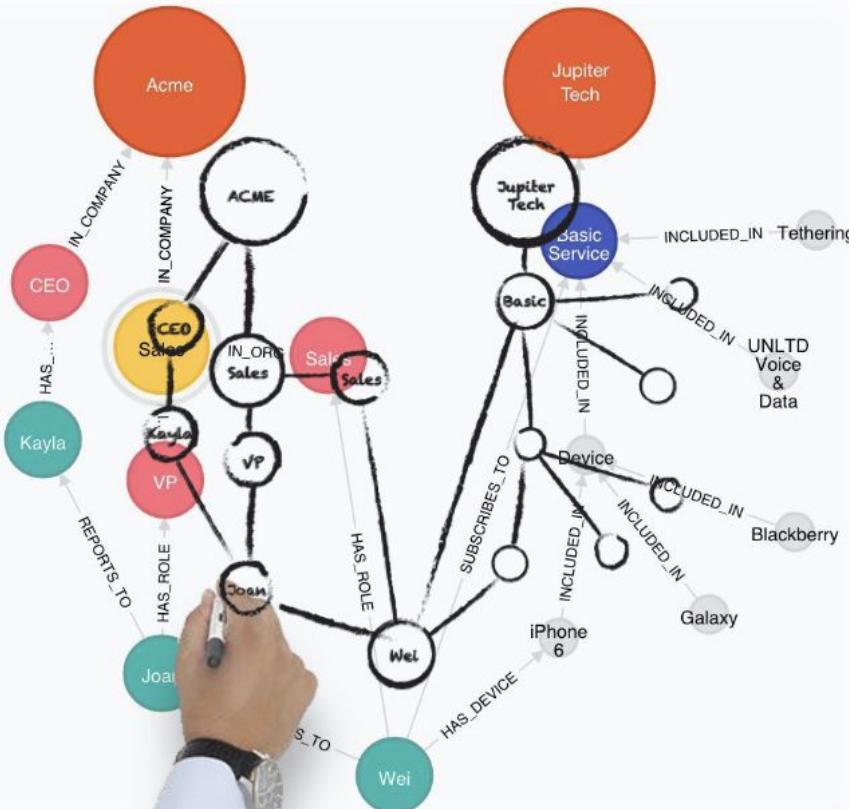
Modelando uma fraude ring em grafos



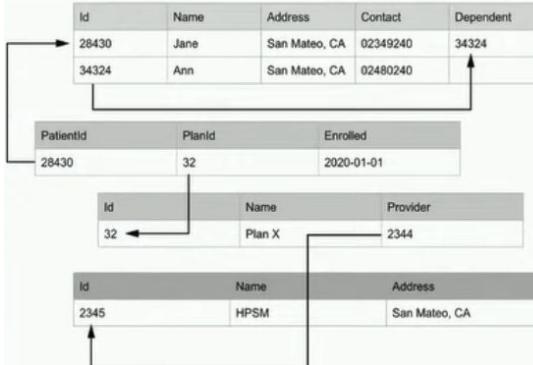
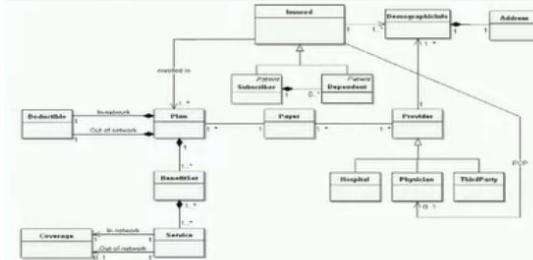
Modelando uma fraude de seguros em grafos



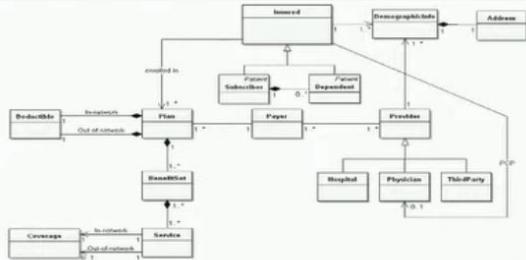
Grafos representam a forma **natural** como pensamos ao resolver problemas



Precisamos mesmo traduzir como pensamos?



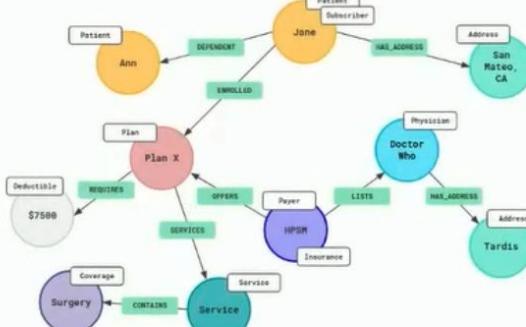
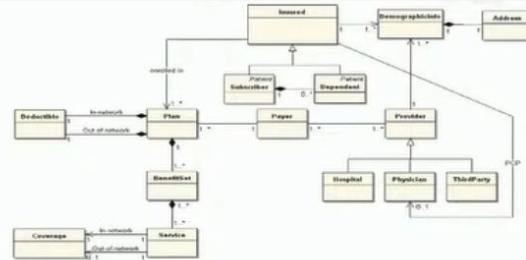
Relational



```
// Patient
{
  "_id": 593340651,
  "first": "Gregorio",
  "last": "Lang",
  "addr": {
    "street": "623 Flowers Rd",
    "city": "Groton",
    "state": "NH",
    "zip": 3266
  },
  "physicians": [
    10387 33456
  ],
  "procedures": [
    {"id": "551ac", "type": "Chest X-ray"},
    {"id": "843fs", "type": "Blood Test"}
  ]
}

// Procedure
{
  "_id": "551ac",
  "date": "2008-04-26",
  "hospital": 161,
  "patient": 593340651,
  "physician": 10387,
  "type": "Chest X-ray",
  "records": ["67bc6"]
}
```

Document

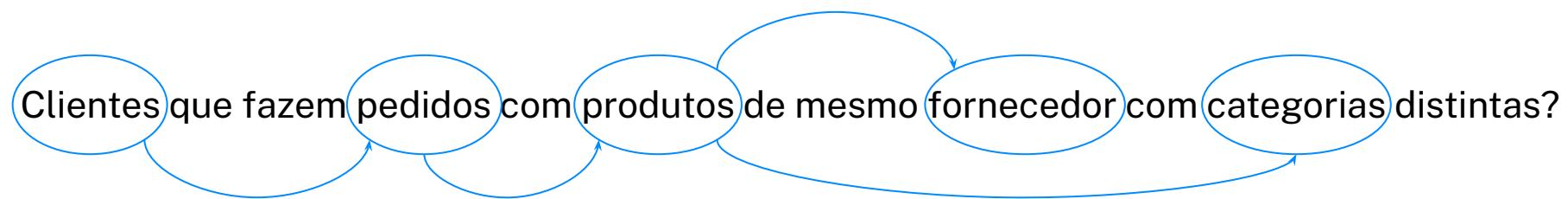


Graph



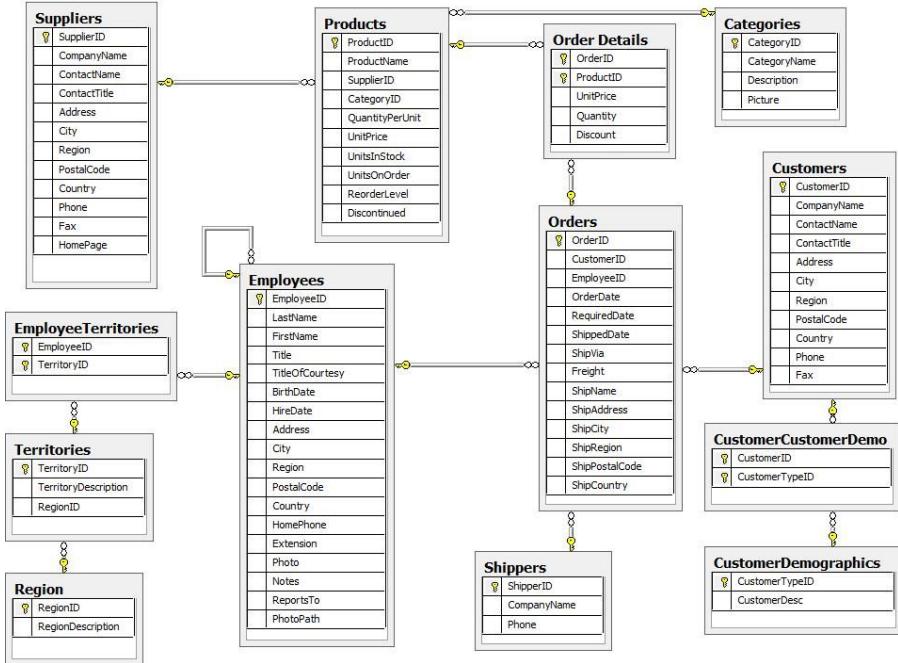
Vamos usar um exemplo de varejo

Começamos a levantar as perguntas que queremos responder, como:

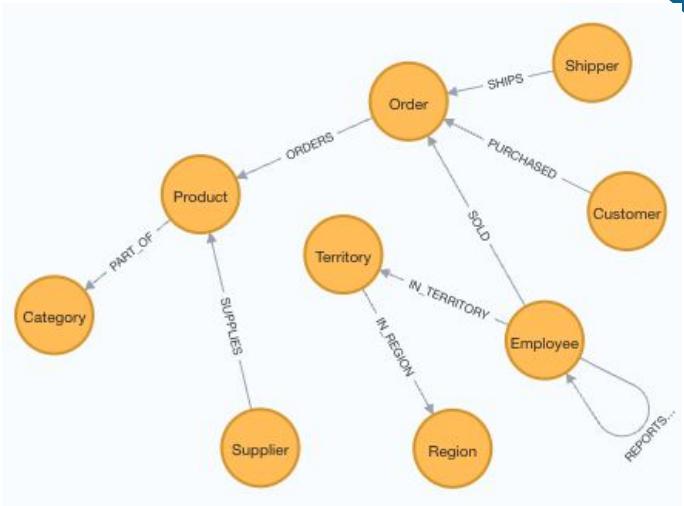


E assim começa a ser modelado nosso grafo...

Como seria esse modelo?



13 tabelas e 92 campos



Simplicidade de modelo
Eliminação de dados duplicados
Redução de volume

O poder do CYPHER, uma linguagem para dados conectados

// Como consultar produtos e categorias de forma recursiva usando SQL (3 níveis)?

```
SELECT p.ProductName  
FROM Product AS p  
JOIN ProductCategory pc ON (p.CategoryID = pc.CategoryID AND pc.CategoryName = "Dairy Products")  
  
JOIN ProductCategory pc1 ON (p.CategoryID = pc1.CategoryID)  
JOIN ProductCategory pc2 ON (pc2.ParentID = pc2.CategoryID AND pc2.CategoryName = "Dairy Products")  
  
JOIN ProductCategory pc3 ON (p.CategoryID = pc3.CategoryID)  
JOIN ProductCategory pc4 ON (pc3.ParentID = pc4.CategoryID)  
JOIN ProductCategory pc5 ON (pc4.ParentID = pc5.CategoryID AND pc5.CategoryName = "Dairy Products")  
;
```

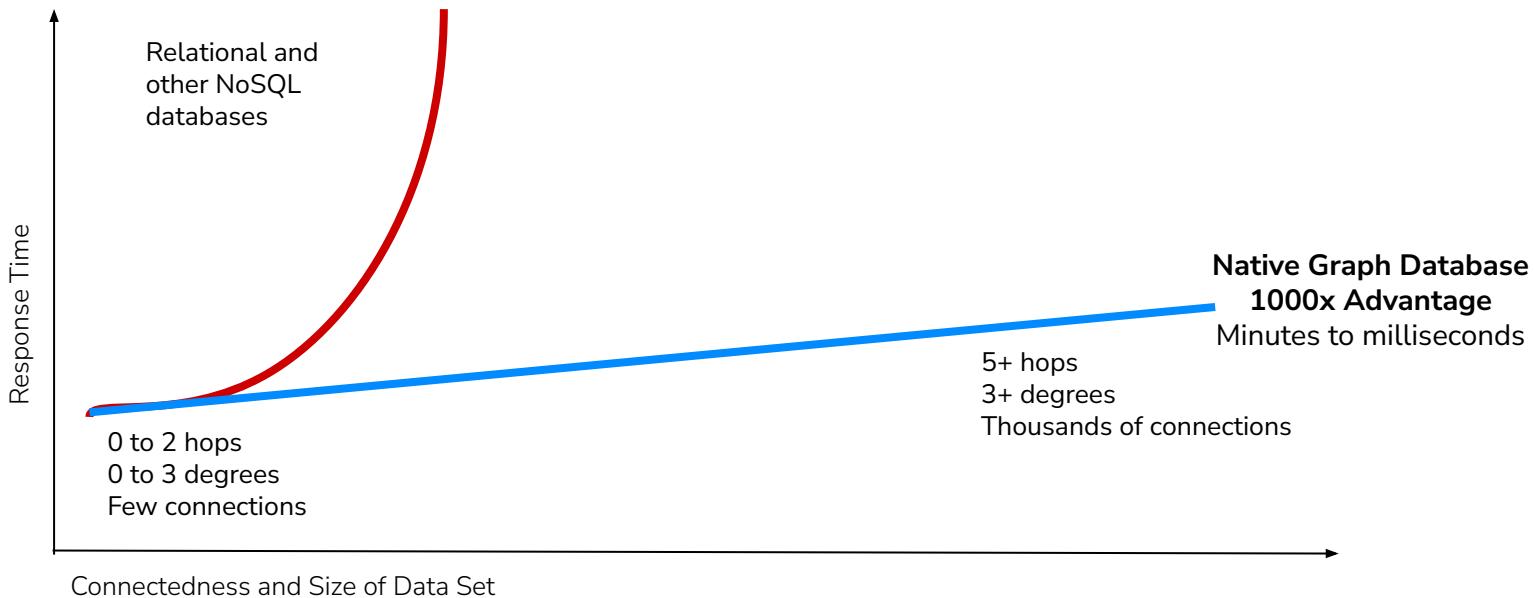
// Como consultar produtos e categorias de forma recursiva usando Cypher (sem limite de níveis)?

```
MATCH (p:Product)-[:CATEGORY]->(l:ProductCategory)-[:PARENT*0..]-(:ProductCategory {name:"Dairy Products"})  
RETURN p.name
```

75% menos código e sem limite de níveis

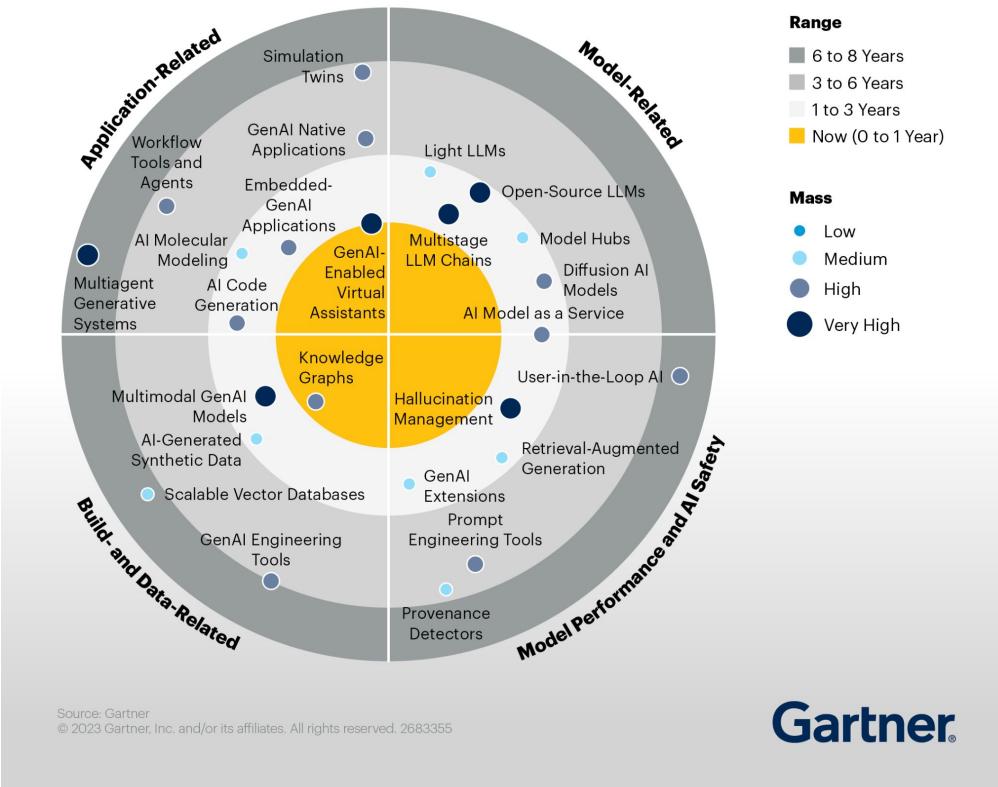
Redução do custo de manutenção

How does it affect performance?



Impact Radar Gartner

Tecnologias chave e imprescindíveis a curto prazo



Neo4j

Quem somos e o que oferecemos





11 / 11
Serviços Financeiros



3 / 3
Montadoras



3 / 3
Aeroespacial e Defesa



6 / 7
Varejistas



4 / 4
Telecomunicações



5 / 5
Farmacêuticas



12 / 12
Tecnologia & Software



13 / 16
Companhias de Seguro

Potenciais casos de uso





AIRBUS



AVIVA



Boston
Scientific



CATERPILLAR®



citibank



DAIMLER

DB BAHN

Deloitte.

Discovery

El Corte Inglés

FUJITSU

1,700+ clientes no
mundo todo

intuit

LEVI STRAUSS & CO.



Microsoft

NASA

NORDSTROM

NOVARTIS



pitney bowes

SAMSUNG

Santander

Staples



UBS

U.S. ARMY

verizon^v

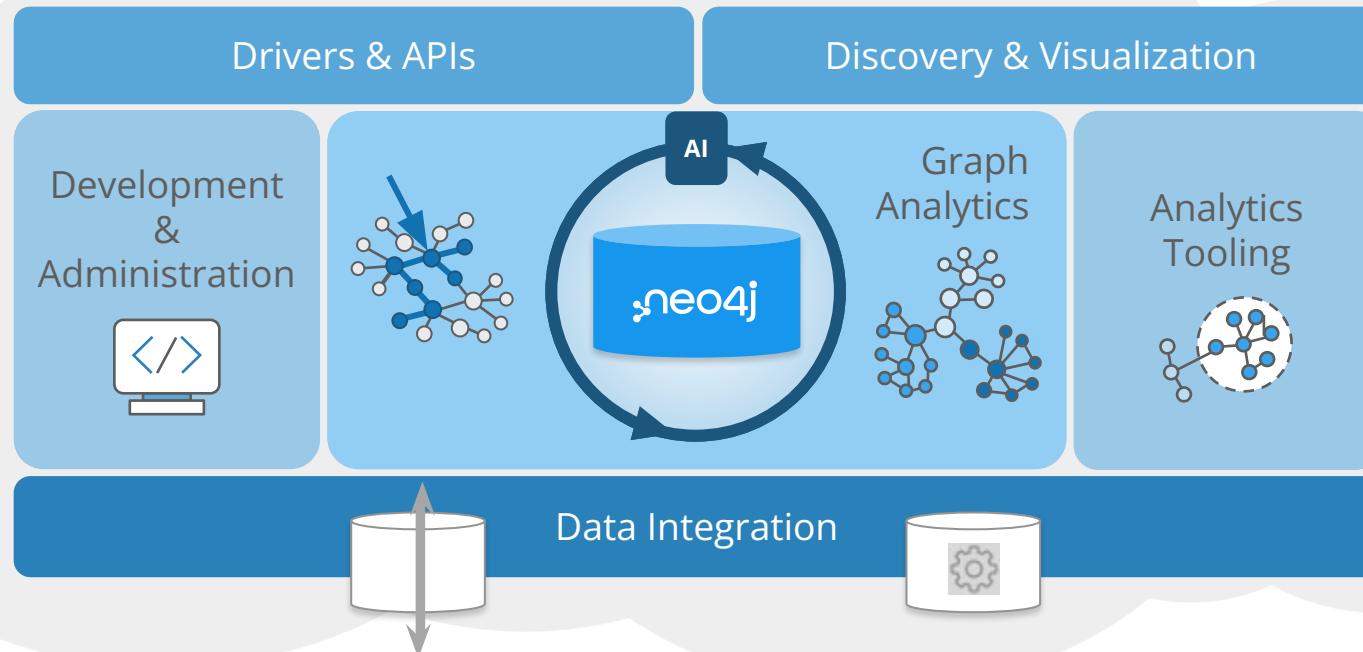


Walmart

Western Union

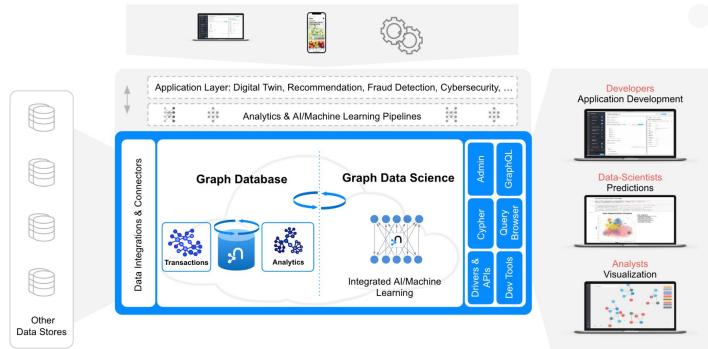
ZURICH[®]

Plataforma Neo4j

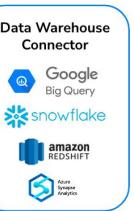


Plataforma Neo4j

+65 Algoritmos prontos



Integrado ao seu ecossistema de Dados



neo4j\$ match (c1:CLIENTE)-[l:EMBEDDEDITEM_CONTA]->(conta:CONTAS)-[l:ESPECIFICOITEM_POSICAO]-(posicao:ITEM_POSICAO) where c1.clienteNome = "José" return c1.clienteNome, posicao.posicaoQtd, posicao.posicaoValor, dia.diaNumero

c1.clienteNome	posicao.posicaoQtd	posicao.posicaoValor	dia.diaNumero
"José"	100.0	1000.0	14
"José"	100.0	1900.0	14
"José"	100.0	1000.0	14
"José"	100.0	1000.0	14

Exploração

Actor View

Woody Allen

Movies for the selected person

The highest voted movies for this person

Dashboard

Benedicto Lemos

Afonso de Mattos

Luciene Schreibl

Luciene Schreibl

Luciene Schreibl

Metrics de Sales

Aplicação

65+ Algoritmos de grafos na Neo4j



Pathfinding & Search

- A* Shortest Path
- All Pairs Shortest Path
- Breadth & Depth First Search
- Delta-Stepping Single-Source
- Dijkstra Single-Source
- Dijkstra Source-Target
- K-Spanning Tree (MST)
- Minimum Weight Spanning Tree
- Random Walk
- Yen's K Shortest Path



Centrality & Importance

- ArticleRank
- Betweenness Centrality & Approx.
- Closeness Centrality
- Degree Centrality
- Eigenvector Centrality
- Harmonic Centrality
- Hyperlink Induced Topic Search (HITS)
- Influence Maximization (Greedy, CELF)
- PageRank
- Personalized PageRank



Community Detection

- Conductance Metric
- K-1 Coloring
- K-Means Clustering
- Label Propagation
- Leiden Algorithm
- Local Clustering Coefficient
- Louvain Algorithm
- Max K-Cut
- Modularity Optimization
- Speaker Listener Label Propagation
- Strongly Connected Components
- Triangle Count
- Weakly Connected Components



Supervised Machine Learning

- Link Prediction Pipelines
- Node Classification Pipelines
- Node Regression Pipelines



Heuristic Link Prediction

- Adamic Adar
- Common Neighbors
- Preferential Attachment
- Resource Allocations
- Same Community
- Total Neighbors



Similarity

- K-Nearest Neighbors (KNN)
- Node Similarity
- Filtered KNN & Node Similarity
- Cosine & Pearson Similarity Functions
- Euclidean Distance Similarity Function
- Euclidean Similarity Function
- Jaccard & Overlap Similarity Functions



Graph Embeddings

- Fast Random Projection (FastRP)
- FastRP with Property Weights
- GraphSAGE
- Node2Vec



... and more!

- Collapse Paths
- Graph Sampling
- Graph Stratified Sampling
- One Hot Encoding
- Pregel API (write your own algos)
- Property Scaling
- Split Relationships
- Synthetic Graph Generation



KEEP
CALM
AND
SHOW ME
THE GRAPH

IA Generativa & Neo4J

**Grafos são a
camada de dados
perfeita para IA
Generativa**



94% dos líderes
empresariais concordam
que a IA é crítica para o
sucesso nos próximos
cinco anos

Gerenciar riscos é a maior barreira para escalar iniciativas de IA

Receio generalizado: Mais da metade dos líderes empresariais atualmente desencorajam a adoção de GenAI.²

Falta de Explicabilidade: Mais de 80% dos executivos se preocupam que a natureza não transparente da GenAI possa resultar em decisões ruins ou ilegais.²

Risco de Imprecisão: Imprecisão e alucinação são dois dos riscos mais citados da adoção da tecnologia GenAI em todos os níveis de uma organização.³

Grafos de conhecimento (KG) desbloqueiam GenAI

KGs

Conhecimento
Fatos
Contexto



GenAI

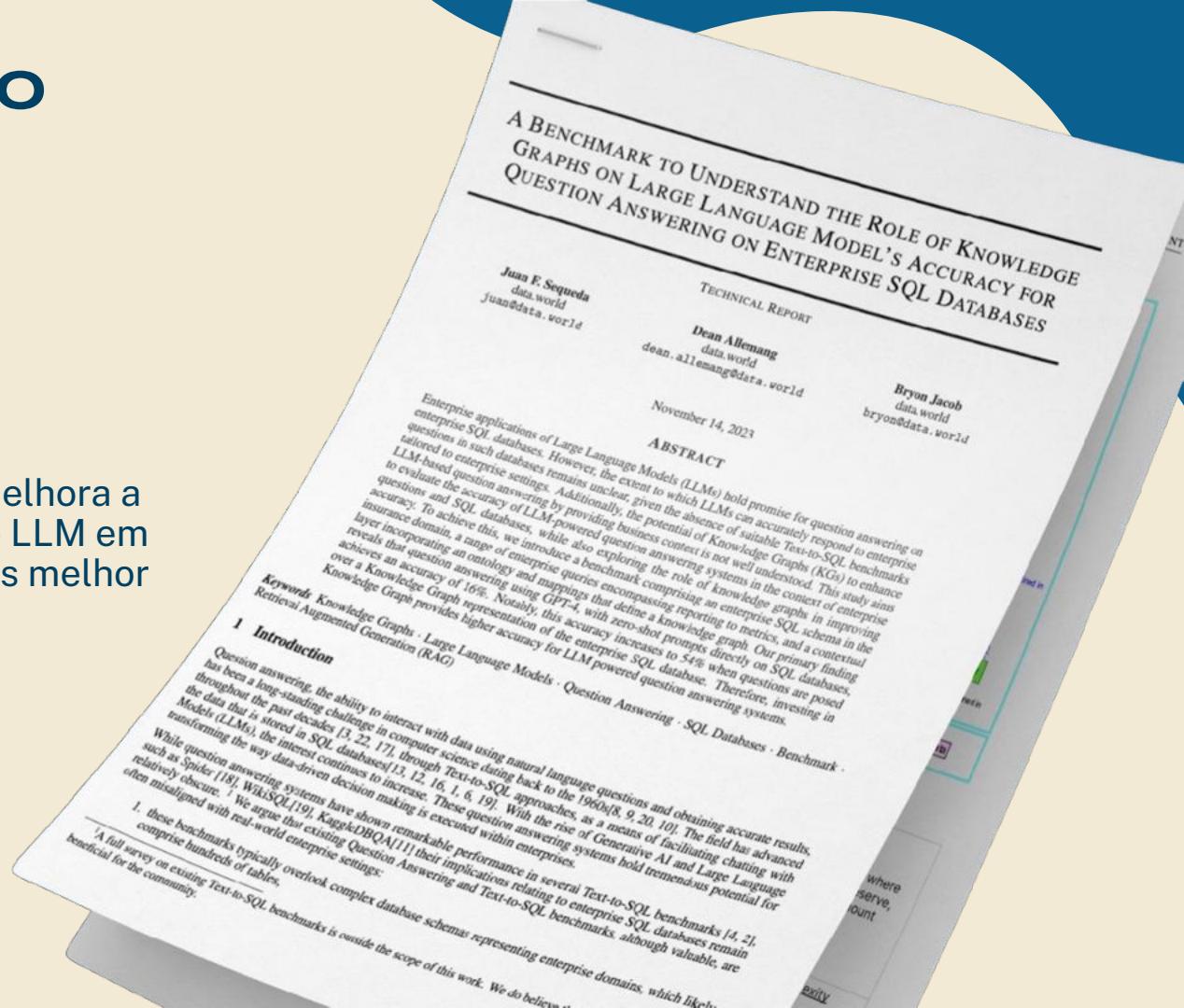
Linguagem
Estatísticas
Criatividade

Preciso
Contextual
Explicável

Maior precisão

3X

Grafo de Conhecimento melhora a precisão das respostas do LLM em **54,2%**, uma média 3 vezes melhor do que apenas SQL



Benefícios do GraphRAG

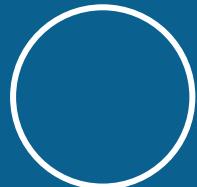
1. Maior
precisão

2. Desenvolvimento
facilitado

3. Explicabilidade
& Governança



KEEP
CALM
AND
SHOW ME
SOME MAGIC



<https://github.com/elizarp/tdc-sp-2025-neo4j>
Baixar arquivo .backup

Abrir conta no AuraDB Free
<https://console.neo4j.io/>
Criar instância

...

Backup and Restore
Restore from backup file
Restaurar o .backup

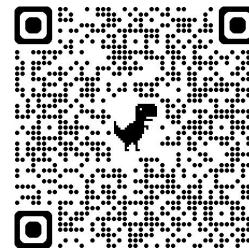
Limite: 200k nós / 400k relacionamentos



Comece agora mesmo GRATUITAMENTE!



<https://graphacademy.neo4j.com/>



<https://neo4j.com/cloud/platform/aura-graph-database/>

neo4j



YouTube



LinkedIn



Comunidade Neo4j

Obrigado!!!