



Descomplicando grafos

Será o fim da linha dos relacionais?

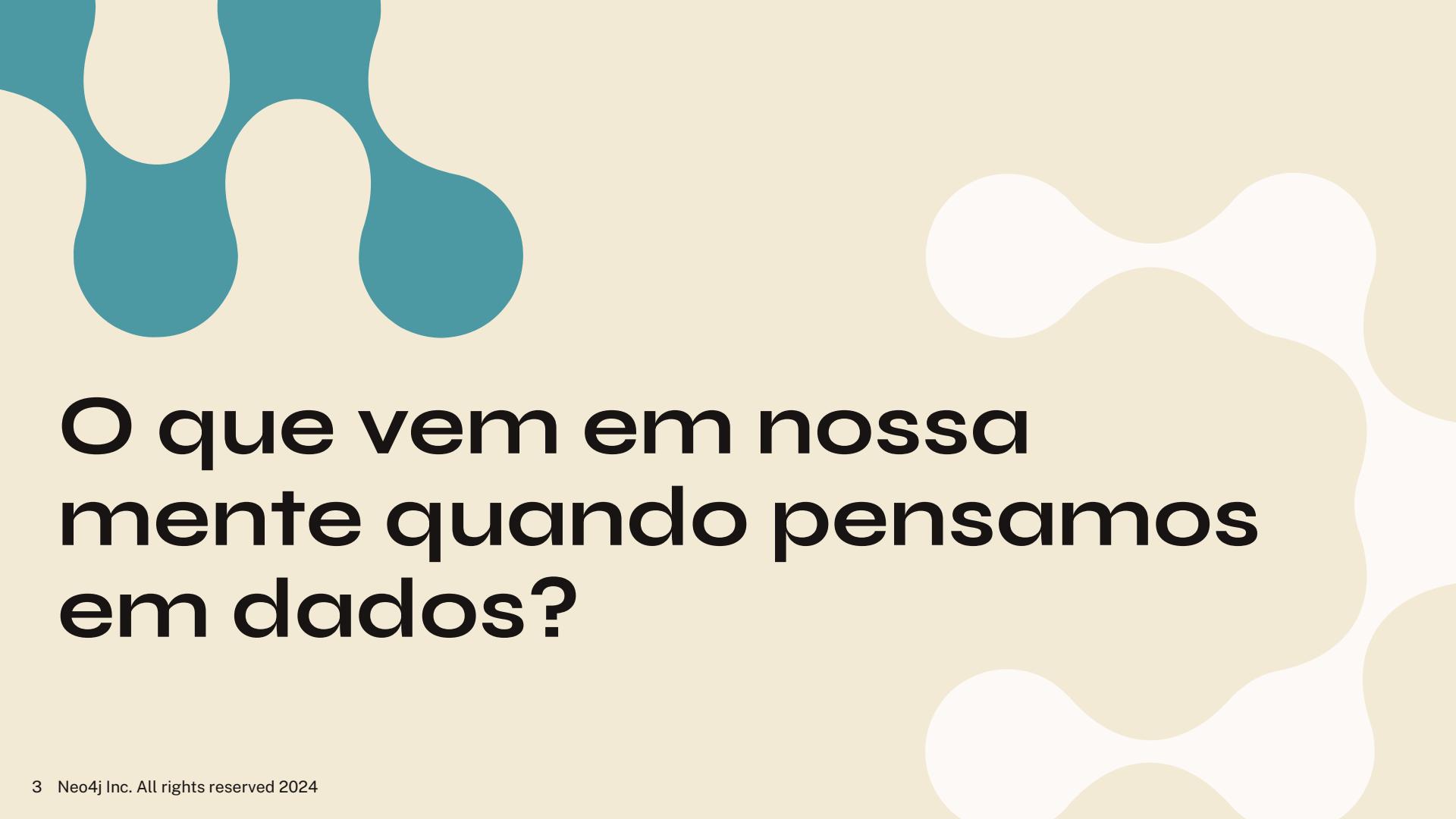
Eliézer Zarpelão - Consulting Engineer - Neo4J

Agenda

1. Grafos
2. Plataforma Neo4j
3. Neo4j + GenAI
4. Demo NodeJS



Eliézer Zarpelão
Consulting Engineer

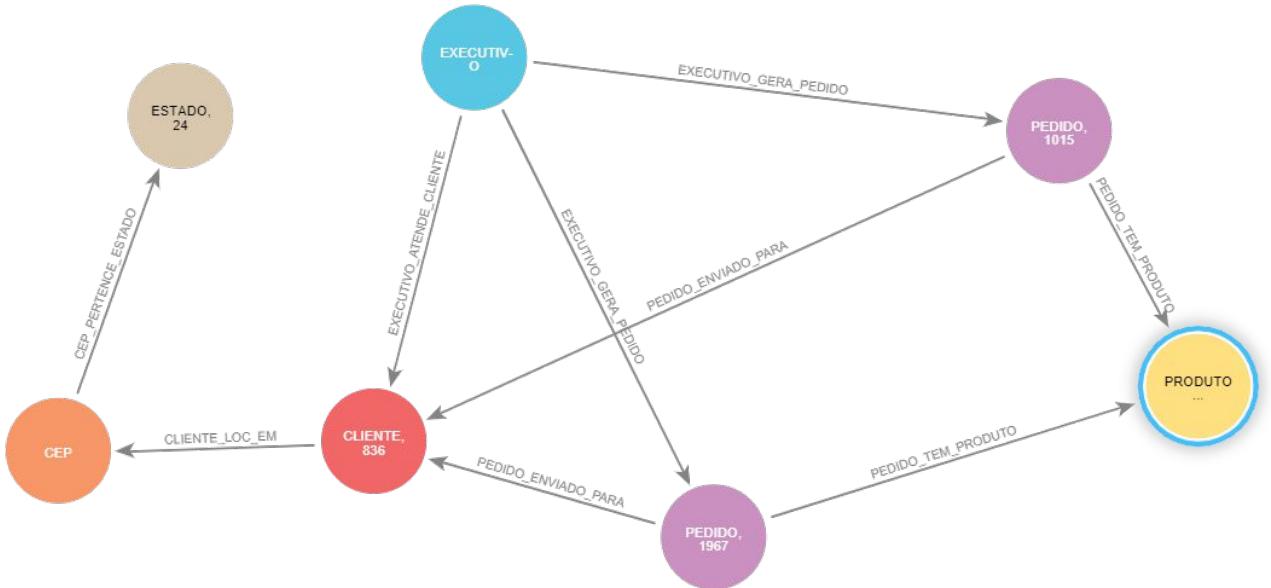


**O que vem em nossa
mente quando pensamos
em dados?**

Sobrenome	Nome	ID da pessoa	Idade	Salário
Ferreira	Maria	1	35	280
Nogueira	Carlos	5	28	325
Souza	Bruna	4	41	265
Oliveira	Jéssica	2	48	359
Silva	João	3	22	250

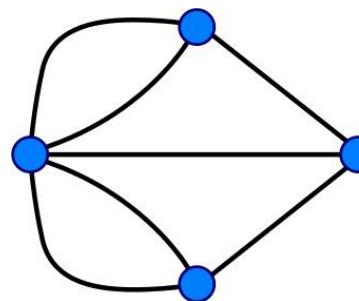
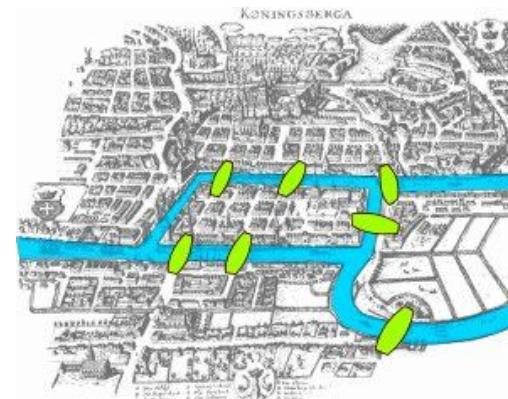
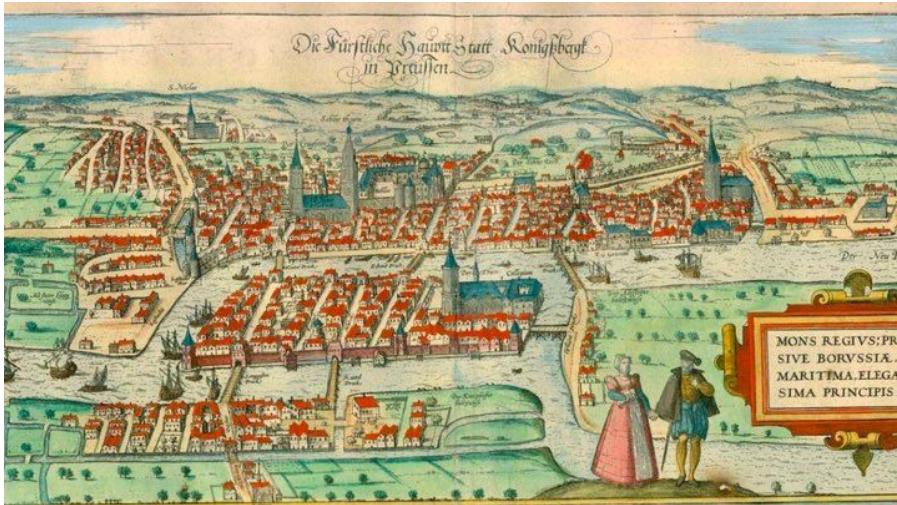
```
Code,Country,Latitude,Longitude,Lat,Lon,Name
AMS,NL,52.18.29N,004.45.51E,0.9129477947605550000000000,0.08315039444709650000,"Amsterdam, Netherlands (AMS)"
ATL,US,33.38.12N,084.25.41W,0.587070582729159000000000,-1.47354755050113000000,"Atlanta, GA (ATL)"
AUS,US,30.11.40N,097.48.12W,0.526992471366065000000000,-1.70466308042286000000,"Austin, TX (AUS)"
BDL,US,41.56.20N,072.41.00W,0.731971695739176000000000,-1.26856347799121000000,"Hartford, CT (BDL)"
DEN,US,39.51.42N,104.40.23W,0.695717328665806000000000,-1.82688945756739000000,"Denver, CO (DEN)"
DFW,US,32.53.49N,097.02.17W,0.574159994491212000000000,-1.69363356917762000000,"Dallas/Fort Worth, TX (DFW)"
EWR,US,40.41.33N,074.10.07W,0.710218105867791000000000,-1.29448646552014000000,"New York/Newark, NJ (EWR - Liberty)"
HND,JP,35.33.08N,139.46.47E,0.620583334178473000000000,2.43961638030086000000,"Tokyo, Japan (HND - Haneda)"
IAD,US,38.56.40N,077.27.21W,0.679708786915568000000000,-1.35185931654264000000,"Washington, DC (IAD - Dulles)"
IAH,US,29.59.04N,095.28.29W,0.523327279936878000000000,-1.66402114953454500000,"Houston, TX (IAH - Intercontinental)"
LAX,US,33.56.33N,118.24.29W,0.592408381358174000000000,-2.06661043032881000000,"Los Angeles, CA (LAX)"
```

```
{
  "name": "create-mixin",
  "author": "Lloyd Brookes <75pound@gmail.com>",
  "version": "3.0.0",
  "description": "Mix the prototype of one class into another",
  "repository": "https://github.com/75lb/create-mixin.git",
  "keywords": [
    "create",
    "mixin",
    "class",
    "extends",
    "es2015",
    "es6"
  ],
  "main": "dist/index.js",
  "module": "index.mjs",
  "files": [
    "index.mjs",
    "dist/index.js"
  ],
  "license": "MIT",
  "engines": {
    "node": ">=10"
  }
},
```



Era uma vez Leonard Euler...

Königsberg, Prussia, 1736



Banco de Dados Grafos

Armazena dados e relacionamentos

Rapidez na navegação dos relacionamentos

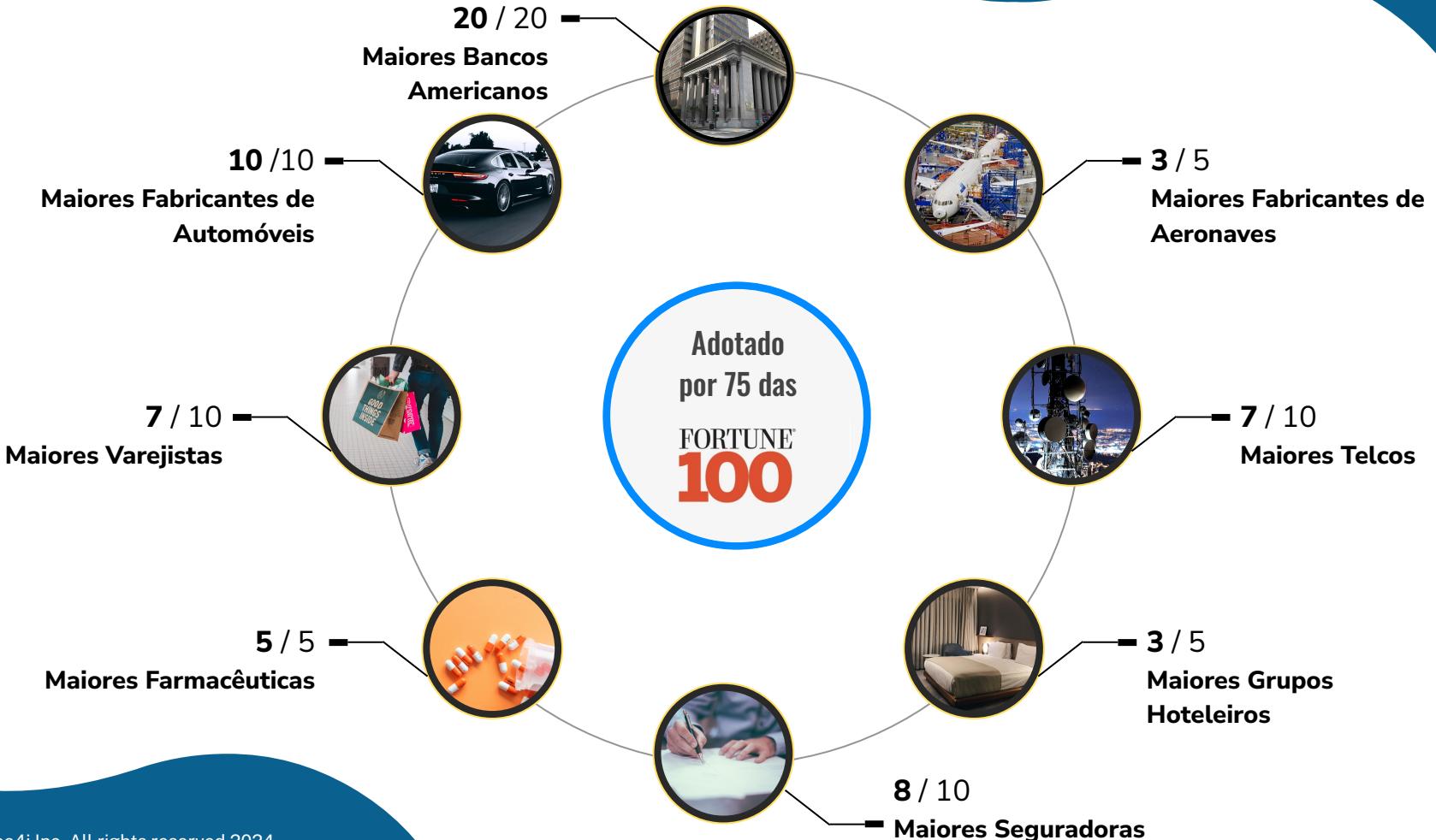
Ótimo para quando relações entre os dados
importam



Você já usou Grafos hoje?

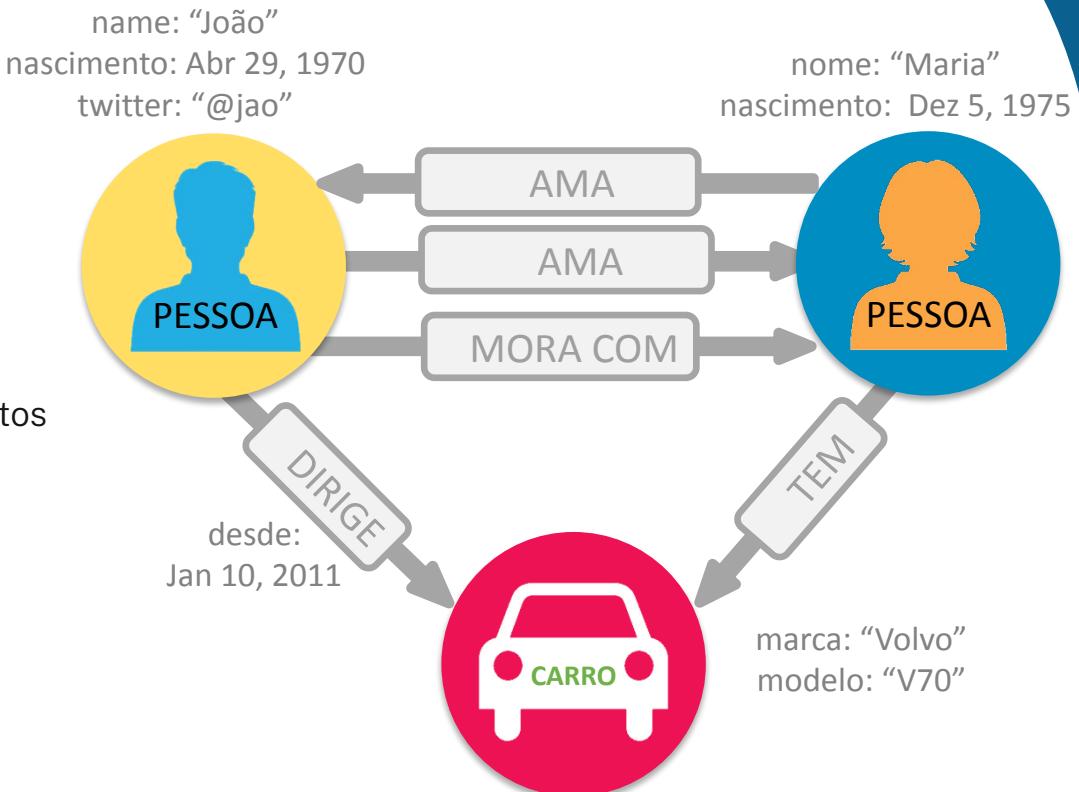




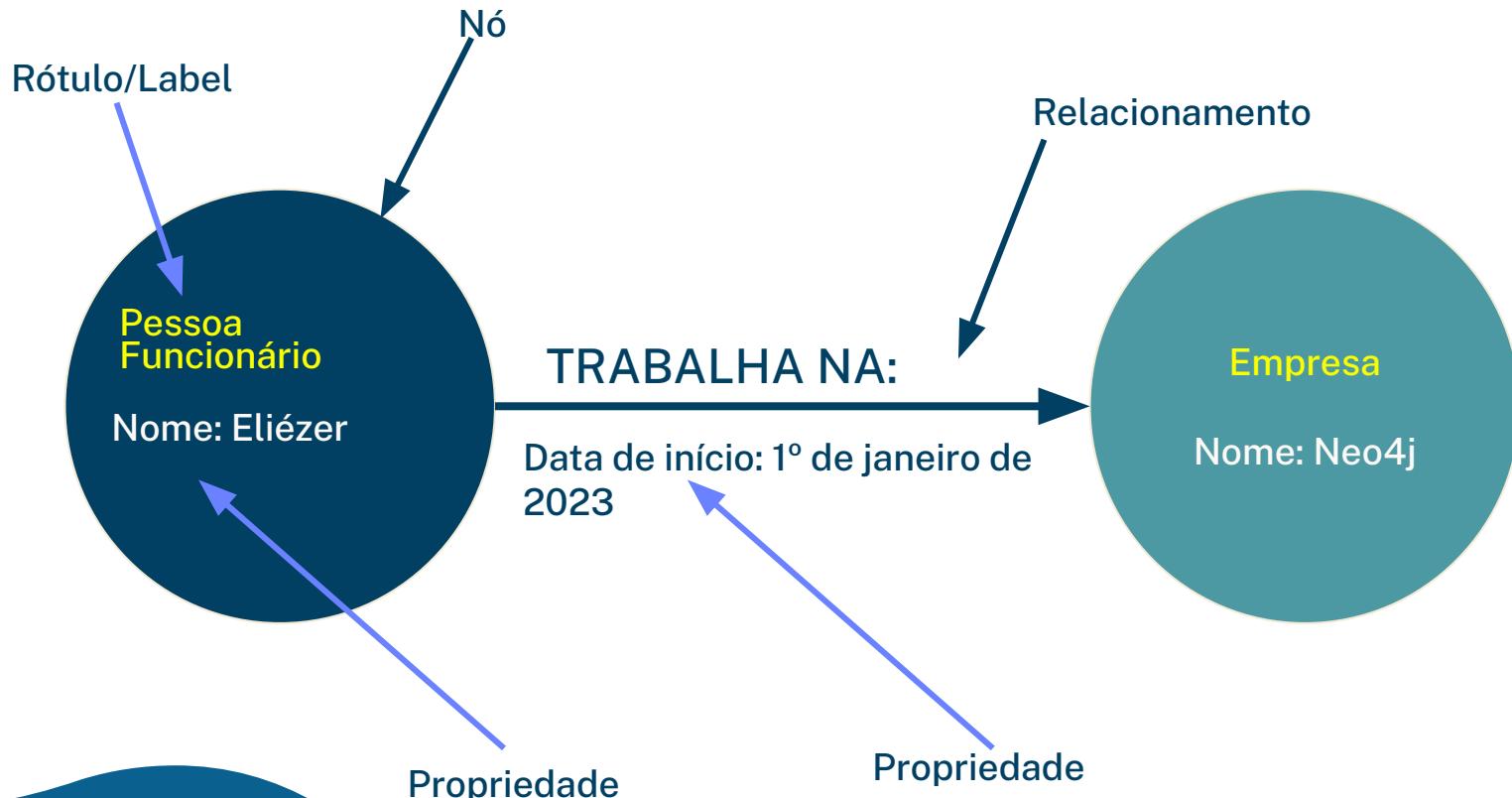


Labeled Property Graph Model

- **Nós / Nodes**
 - Pode ter rótulos/labels para classificar nós
 - Rótulos tem índices nativos
- **Relacionamentos / Relationships**
 - Relaciona nós por tipo e direção
- **Propriedades / Properties**
 - Atributos de nós e relacionamentos
 - Armazenados como chave/valor
 - Podem ter índices simples e compostos



Exemplo:





Recomendações em tempo real



Detecção de Fraude



Rede e operações de TI



Gerenciamento de dados mestres



Risco e conformidade



Gerenciamento de identidade e acesso

ebay

BEST BUY

JCPenney

Kroger

NORDSTROM

Marriott

adeo

浙商银行
CZBANK

DISCOVER
FINANCIAL SERVICES

WESTERN UNION
WU

dun & bradstreet

BMO

AVIVA

at&t

SOCIETE
GENERALE

orange

BT

edf

Alcatel-Lucent

1

verizon

LOCKHEED MARTIN

Adobe

UBS

CITI

VOLVO

RBC

MERCK

Santander

BASF
The Chemical Company

AkerBP

Current

AIRBUS

COMCAST

J.P.Morgan

CISCO

Royal Bank
of Scotland

Microsoft



Varejo



Telecomunicações



Manufatura



Serviços financeiros



Assistência médica



Governo



neo4j.com/case-studies/nasa/

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 2025, grafos serão usados em 80% das inovações em dados e analytics”

ANÁLISE ATRAVÉS DE FILTROS

PRÍNCIPE CHARLES

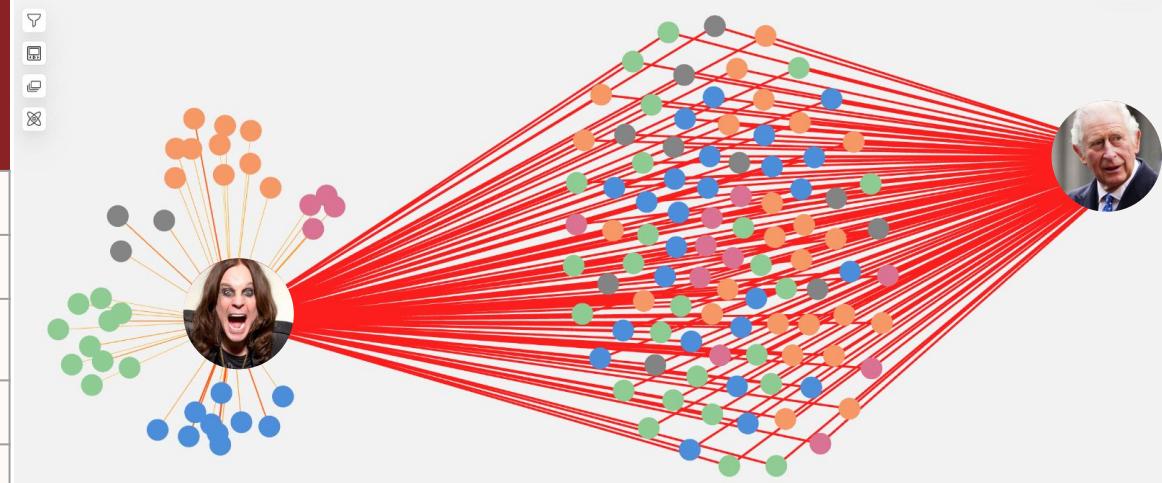


OZZY OSBOURNE

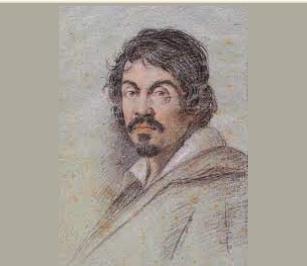


Homem	Homem
Nascido em 1948	Nascido em 1948
Residente no Reino Unido	Residente no Reino Unido
Casado 2 vezes	Casado 2 vezes
Mora em um castelo	Mora em um castelo
Classe A	Classe A
Grande poder aquisitivo	Grande poder aquisitivo
Conservador	Agressivo

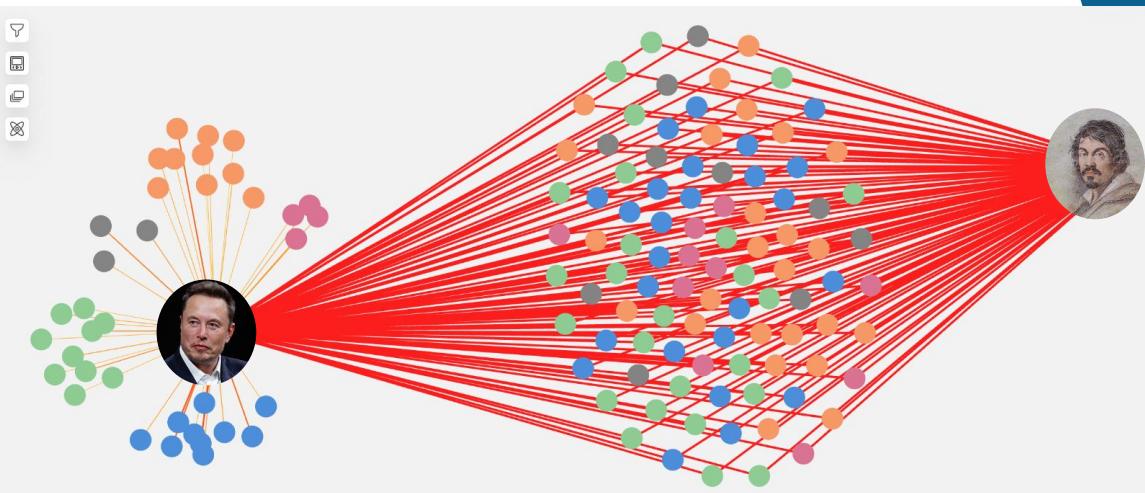
ANÁLISE ATRAVÉS DE CONTEXTO



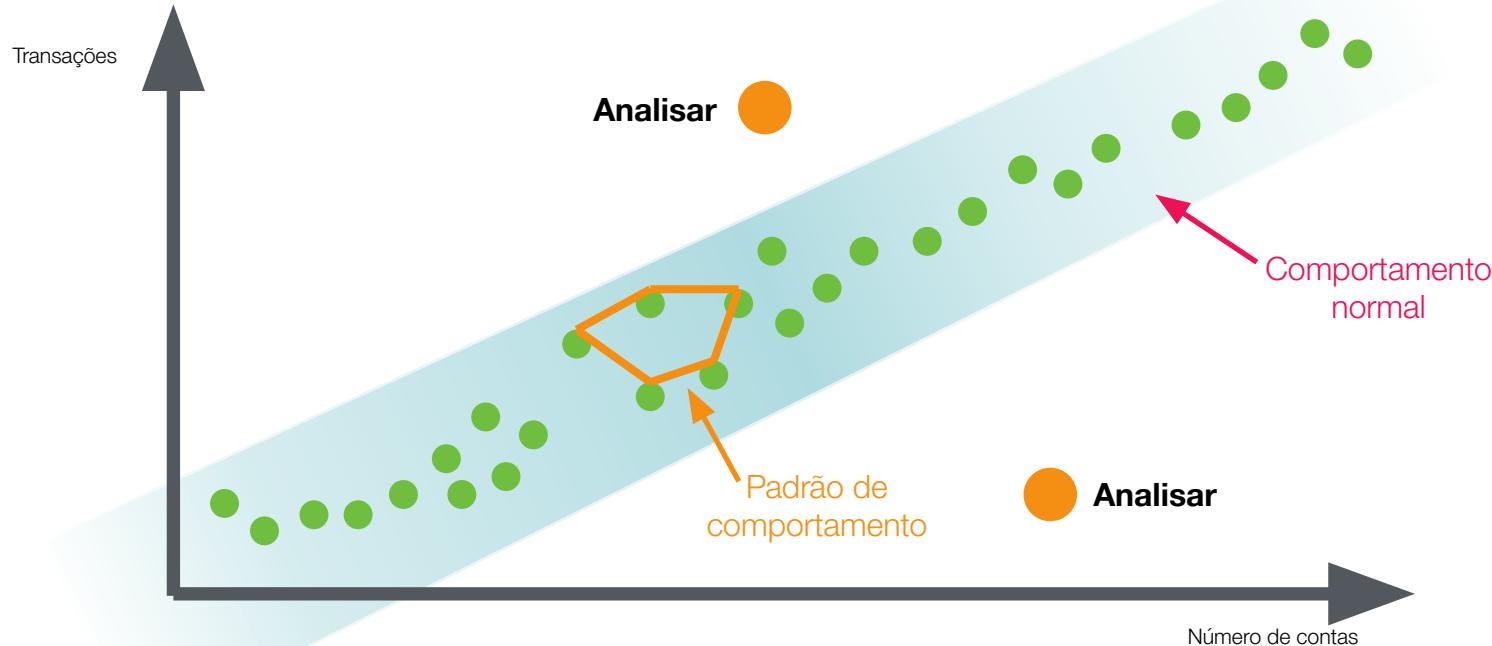
ANÁLISE ATRAVÉS DE FILTROS

	
Elon Musk	Caravaggio
Inovador	Inovador
Provocador	Provocador
Desafiou convenções	Desafiou convenções
Impactante	Impactante
Controverso	Controverso
Genial	Genial
Problemas pessoais	Problemas pessoais
Empresário	Pintor barroco italiano

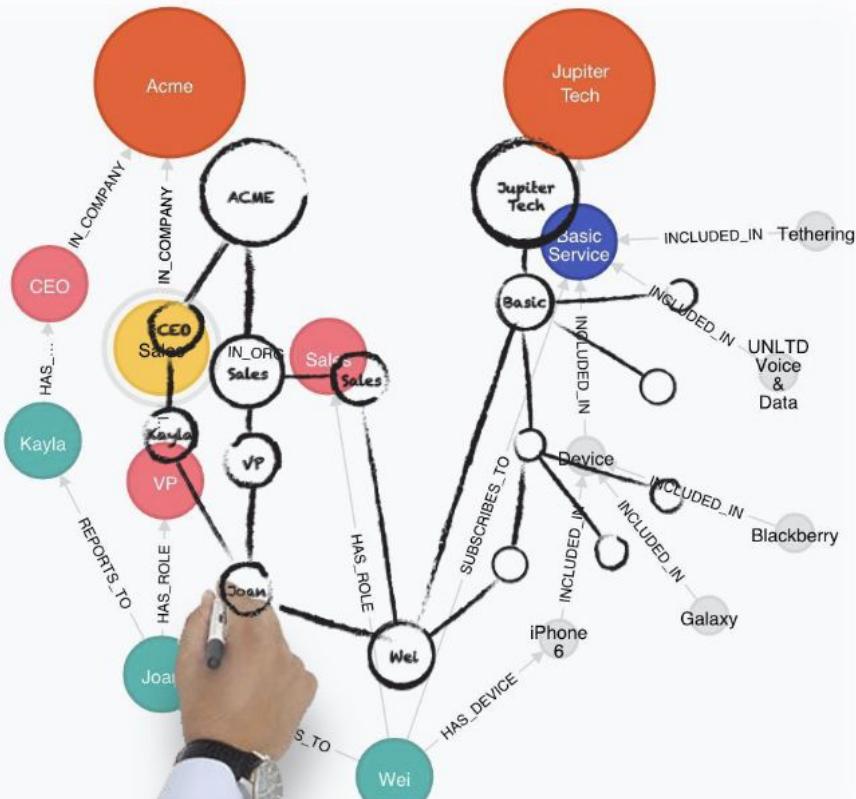
ANÁLISE ATRAVÉS DE CONTEXTO



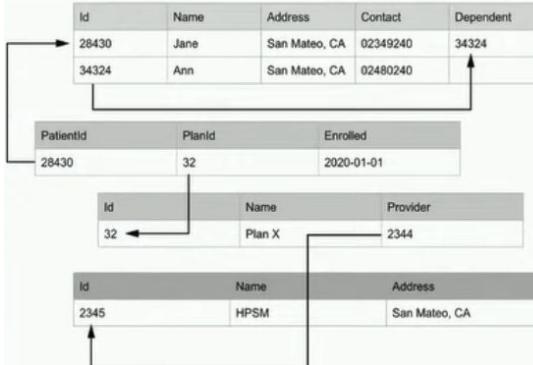
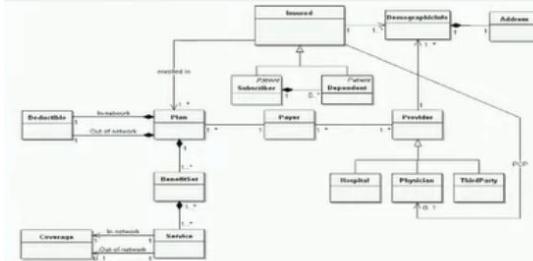
ANÁLISE DE ANOMALIAS



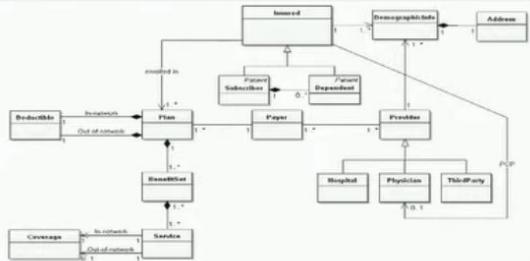
GRAFOS REPRESENTAM A FORMA NATURAL COMO PENSAMOS AO RESOLVER PROBLEMAS



PRECISAMOS MESMO TRADUZIR COMO PENSAMOS?



Relational

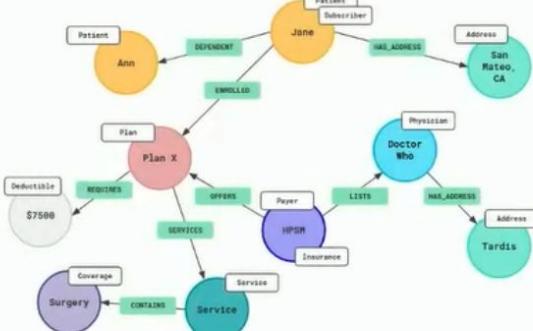
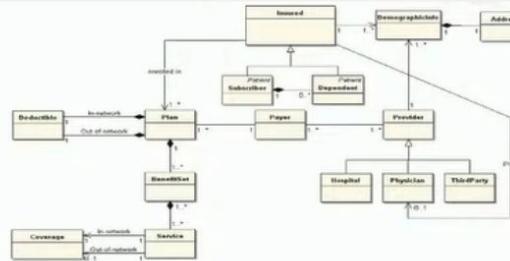


```

// 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

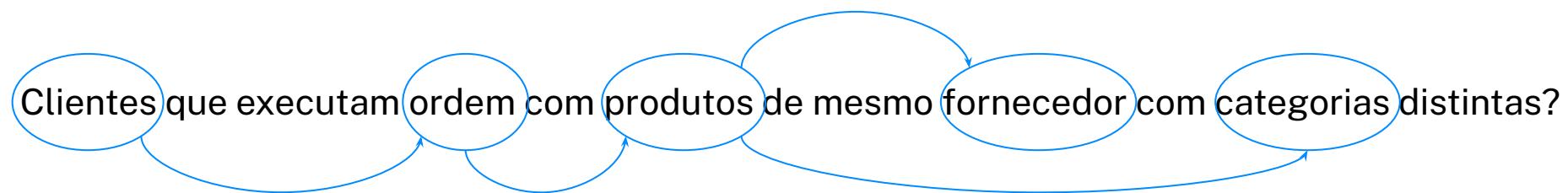


Graph



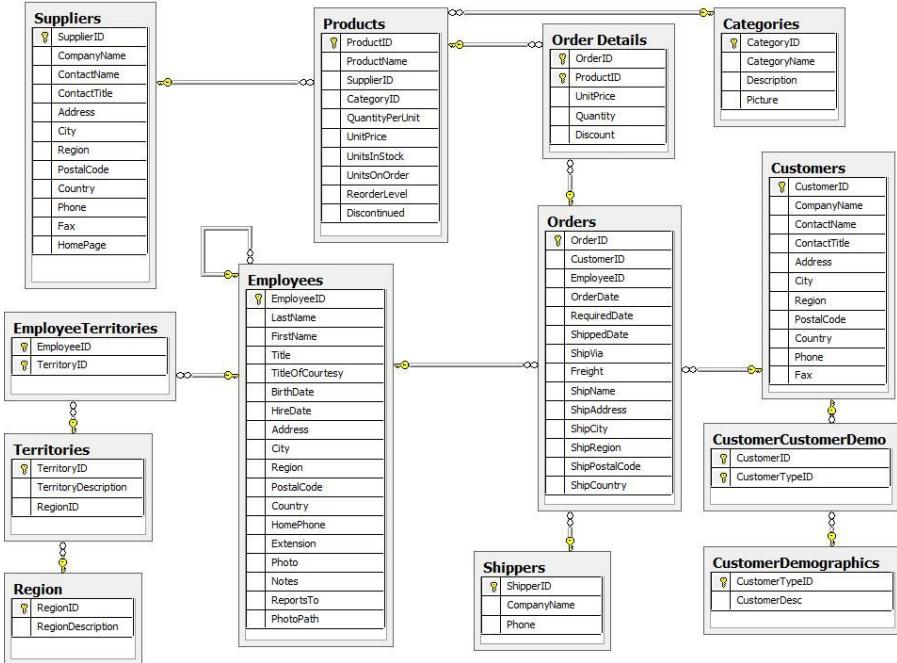
VAMOS UTILIZAR UM EXEMPLO DE VAREJO

Começamos a levantar as perguntas que queremos responder, como:

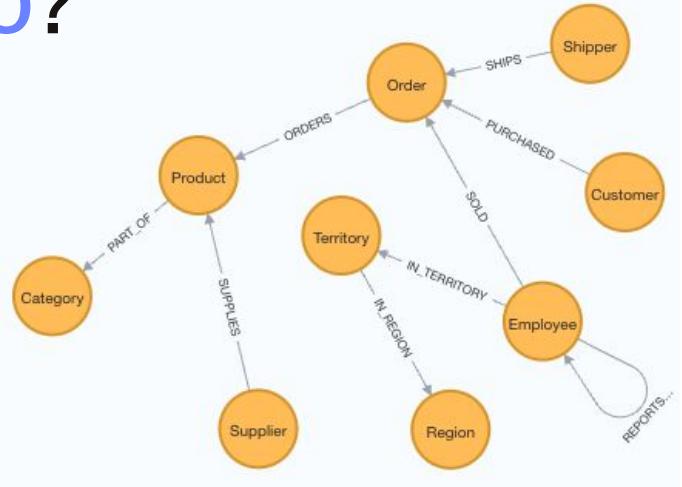


E assim começa a ser modelado nosso grafo...

COMO SERIA ESSE MODELO?



13 tabelas e 92 campos



Simplicidade de modelo
Eliminação de dados duplicados
Redução de volume

ENTENDENDO 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")

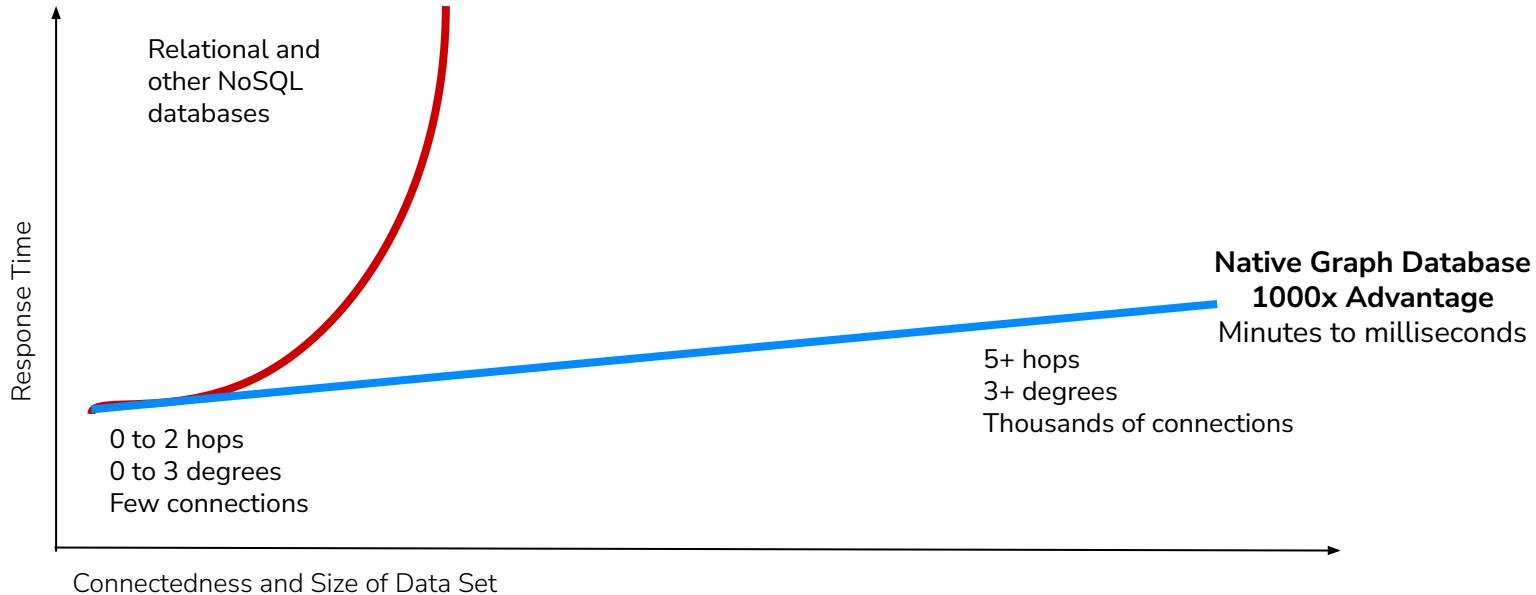
JOIN ProductCategory pc3 ON (p.CategoryID = pc3.CategoryID
JOIN ProductCategory pc4 ON (pc3.ParentID = pc4.CategoryID)
JOIN ProductCategory pc5 ON (pc4.ParentID = pc5.CategoryID AND pc5.CategoryName = "Dairy Products")
;
```

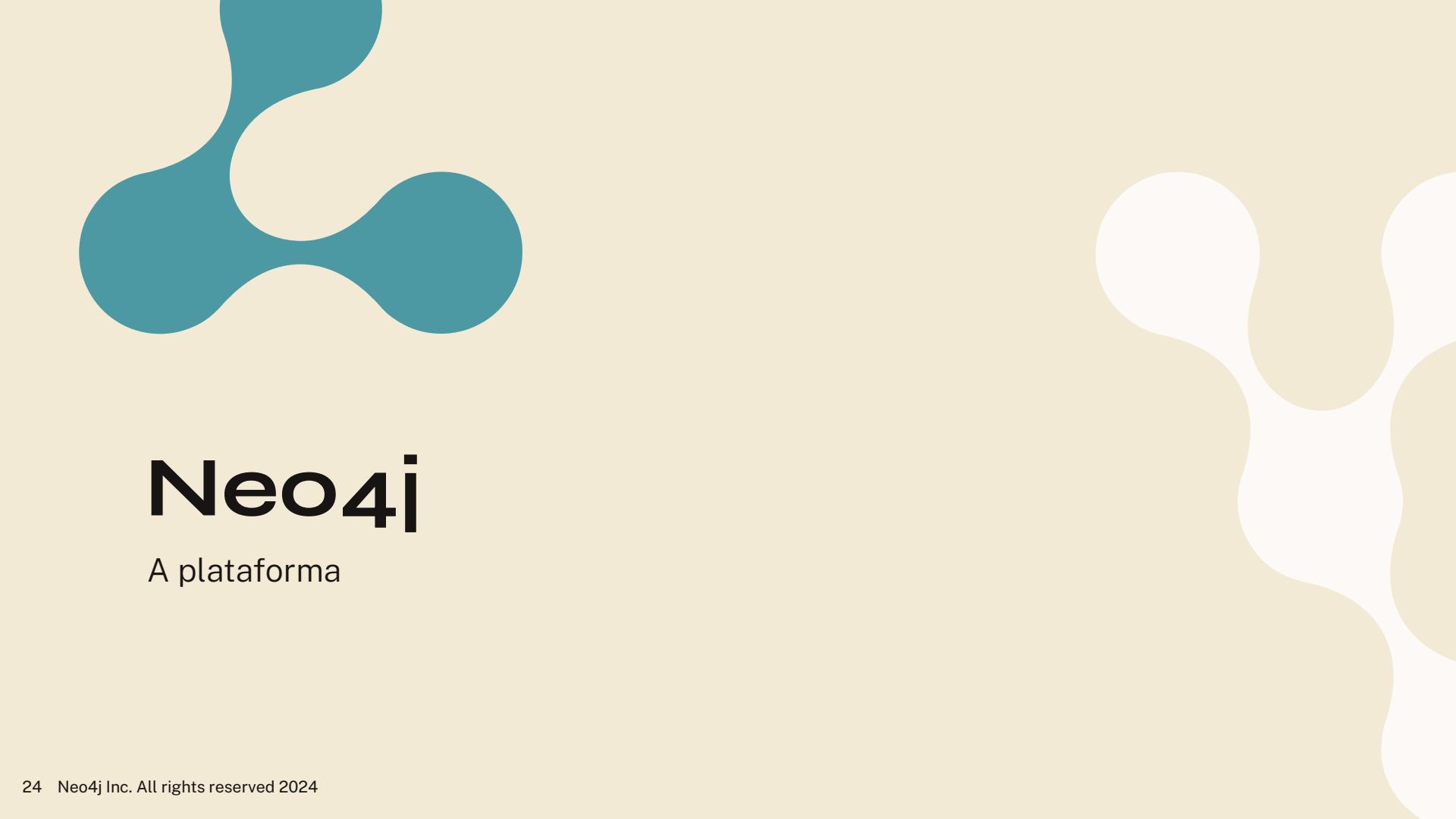
// Como consultar produtos e categorias de forma recursiva usando Cypher (sem limite de níveis)?

```
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**

PERFORMANCE: GRAFOS x BANCOS TRADICIONAIS

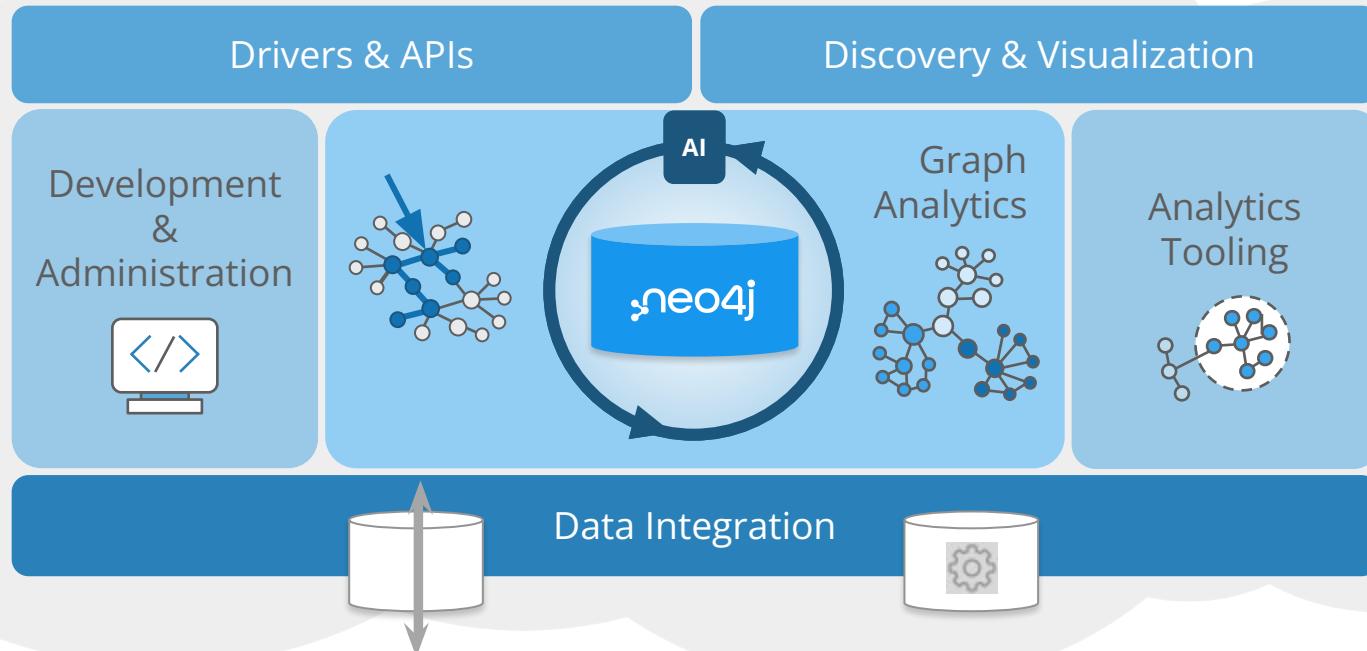




Neo4j

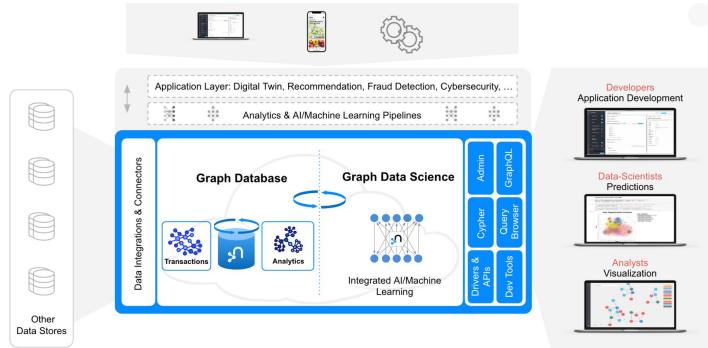
A plataforma

Plataforma Neo4j

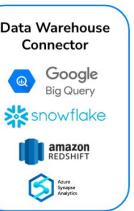


Plataforma Neo4j

+65 Algoritmos prontos



Integrado ao seu ecossistema de Dados



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

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

Exploração

Actor View

Woody Allen

Movies for the selected person

The highest voted movies for this person

Dashboard

Benedicta Lemos

Afonso de Matos

Luciene Schreibl

Luciene Schreibl

Luciene Schreibl

Metrics of 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

Prevê-se que a IA generativa possa desbloquear

\$6.6 Trilhões Em Valor Econômico

Além de 3,3% de melhoria de produtividade anual *

LLMs trazem uma oportunidade incrível para:

1

Automatizar tarefas de
recuperação de dados

2

Melhorar experiências de
atendimento ao cliente

3

Agilizar a leitura, compreensão e resumo

4

Gerar conteúdo e código

75% do valor da IA generativa virá dessas áreas

Mas há desafios...

O que as AIs Generativas podem fazer!

	Papagaio	Chat GPT
Aprende frases aleatórias e de pessoas aleatórias.	✓	✓
Fala como uma pessoa, mas não entende o que está falando.	✓	✓
Ocasionalmente fala coisas que não fazem o menor sentido.	✓	✓
É um pássaro bonitinho!	✓	✗

- 1 Alucinações
- 2 Respostas razoáveis, mas nem sempre precisas
- 3 Pode herdar o viés por meio de dados de treinamento
- 4 Falta de conhecimento do domínio empresarial
- 5 Incapacidade de verificar ou atribuir fontes

Mas o que são alucinações?

Case 4 - ChatGPT cited non-existent medical references

[Professor Robin Emsley](#) detailed his experience experimenting with ChatGPT for writing scientific papers. While ChatGPT performed consistently in general brain-related topics, it started to hallucinate to the point of falsification when pressed on complex topics. In this case, ChatGPT provided several references that didn't exist or were irrelevant to the response it generated.

BREAKING

Lawyer Used ChatGPT In Court —And Cited Fake Cases. A Judge Is Considering Sanctions

Molly Bohannon Forbes Staff
I cover breaking news.

Follow

**Como você pode aproveitar essa
enorme oportunidade e superar
esses desafios?**

Modelos de LLMs com Grafo de conhecimento do Neo4J

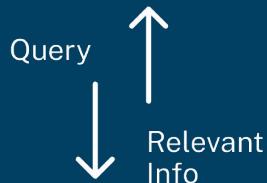


User Asks
a Question

LLM
Smart Search

LLM Question +
Relevant Info

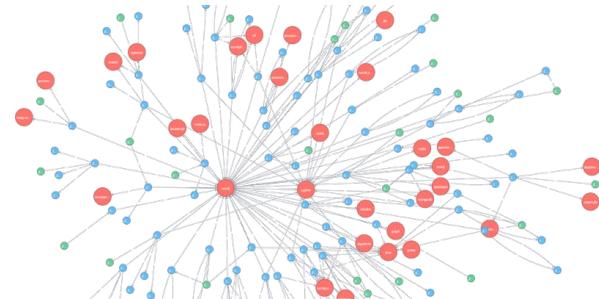
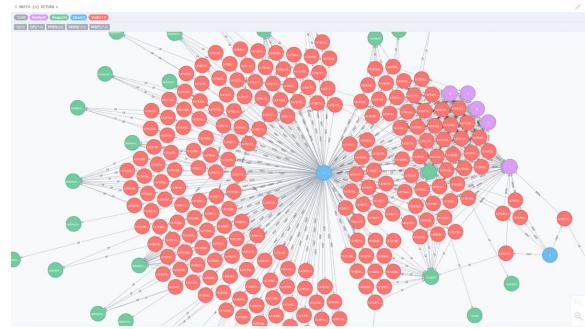
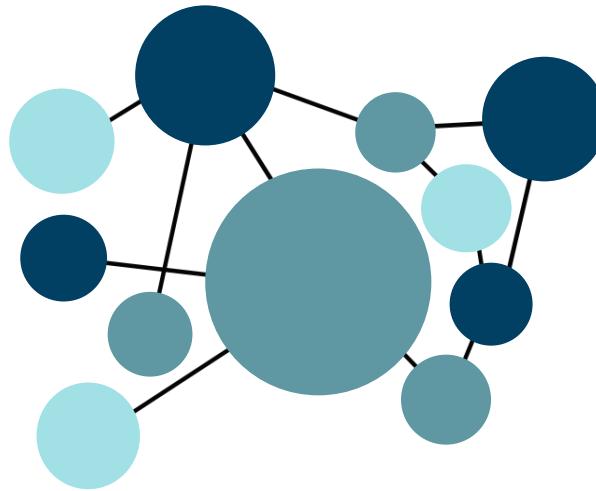
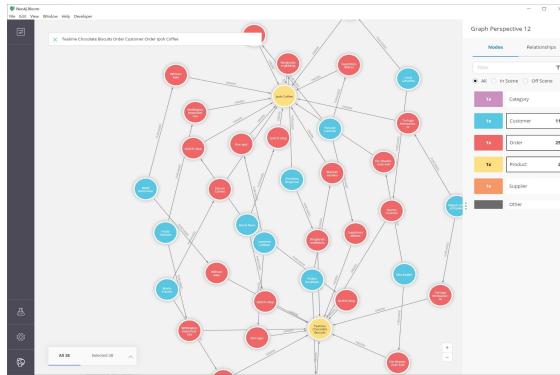
Enriched
Answer



Neo4j's Knowledge Graph

Retrieval Augmented Generation (RAG) combina modelos de LLMs com a recuperação de conhecimento externo, fundamentando as respostas em fatos e prevenindo melhor as respostas com alucinação.

Grafos de conhecimento estruturam dados de forma conectada



Grafo de Conhecimento

Estrutura

Fatos sobre pessoas, lugares ou coisas **interligadas** por
seus **relacionamentos**

Formato entendido por **máquinas e seres humanos**

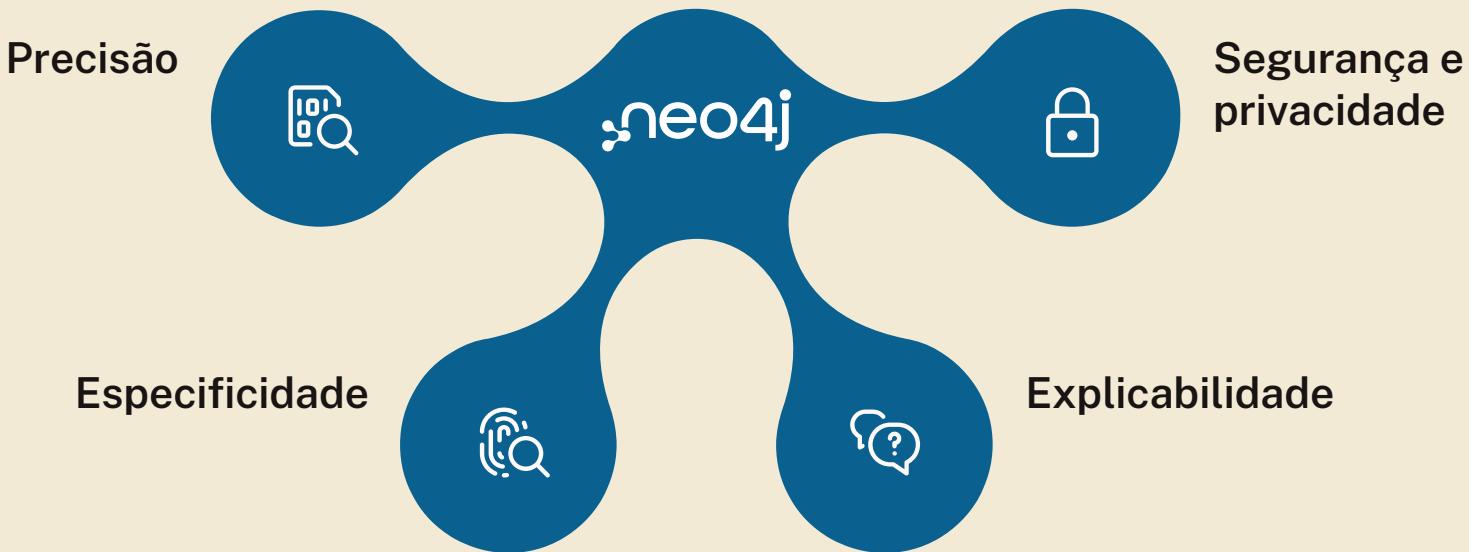
Organizado fornece contexto para o raciocínio sobre os
dados



Vantagens da utilização de Grafos

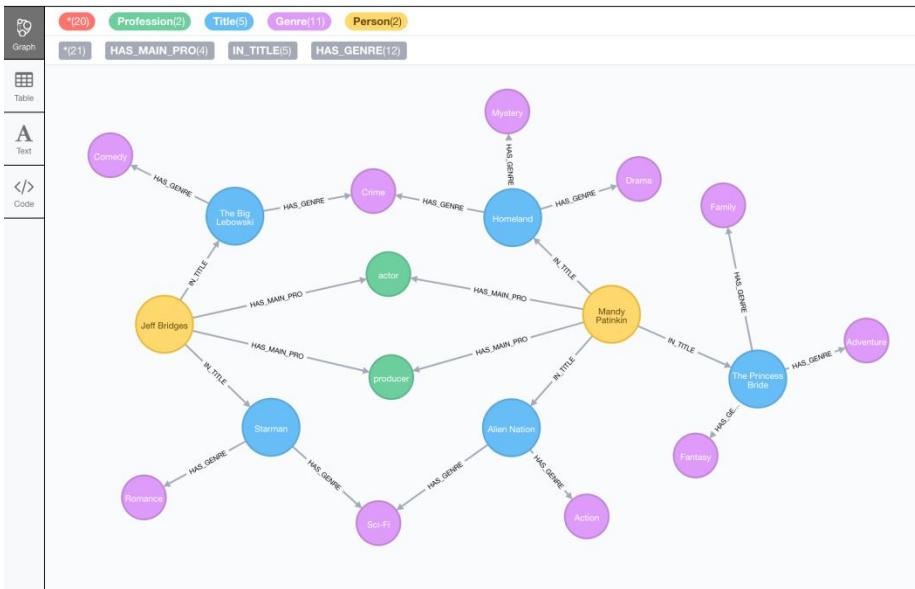
MATCH(:Grafos)-[:MENOS]->(:Alucinacoes)

Grafos + GenAI



Precisão

Os LLMs podem ajudar a gerar **respostas mais precisas** considerando as conexões e dependências dentro do grafo de conhecimento e **mapeando novos links à medida que novos dados são identificados.**



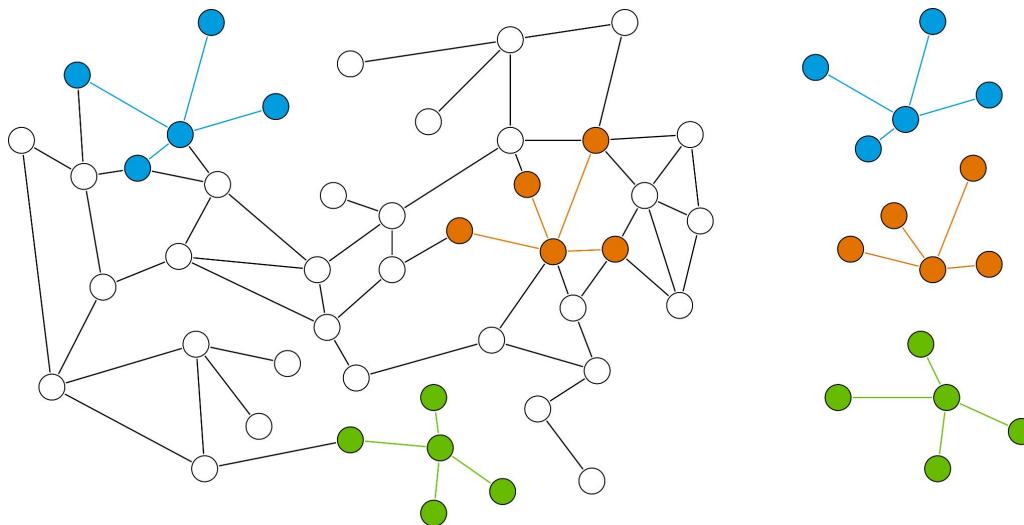
Relacionamentos são dados usados para retornar resultados explícitos

Esquema flexível significa que é fácil aumentar sua base de conhecimento sempre que novas informações estiverem disponíveis

Pesquisa Vetorial adiciona resultados implícitos mostrando respostas semelhantes

Especificidade

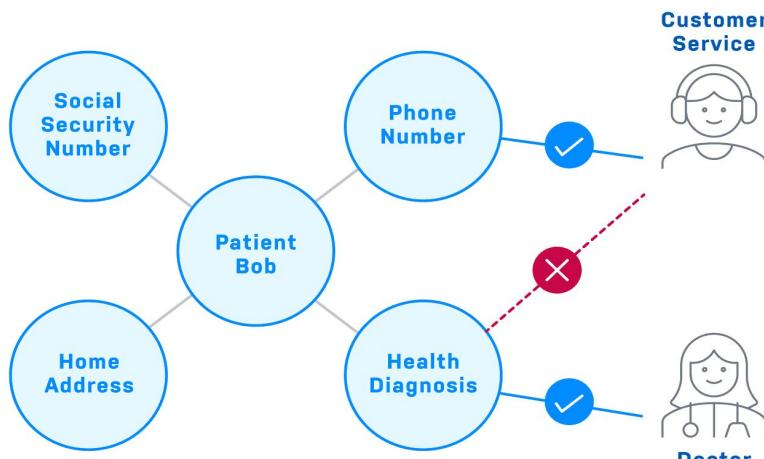
Os gráficos de conhecimento do Neo4j fornecem ao LLM informações sobre sua empresa para que as **respostas sejam específicas para o seu negócio**, dando mais contexto para respostas mais precisas.



Adicione uma camada de contexto sobre seu LLM para precisão e especificidade

Segurança e Privacidade

O LLM recupera e **retorna informações regidas por suas políticas corporativas de segurança e controle de acesso** - até o nível do nó.



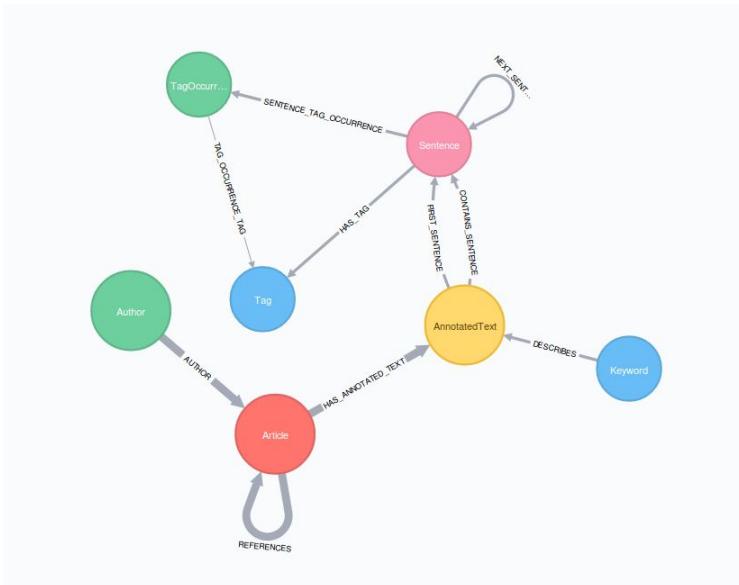
Integração ao provedor de gerenciamento de identidade e acesso com SSO

Define políticas por função ou identidade

Crie restrições em nós, rótulos, relacionamentos, propriedades, partes específicas do gráfico e até mesmo profundidade de travessias

Explicabilidade

Verifique as respostas enriquecidas do seu LLM, pois cada informação está vinculada às suas fontes e origens.



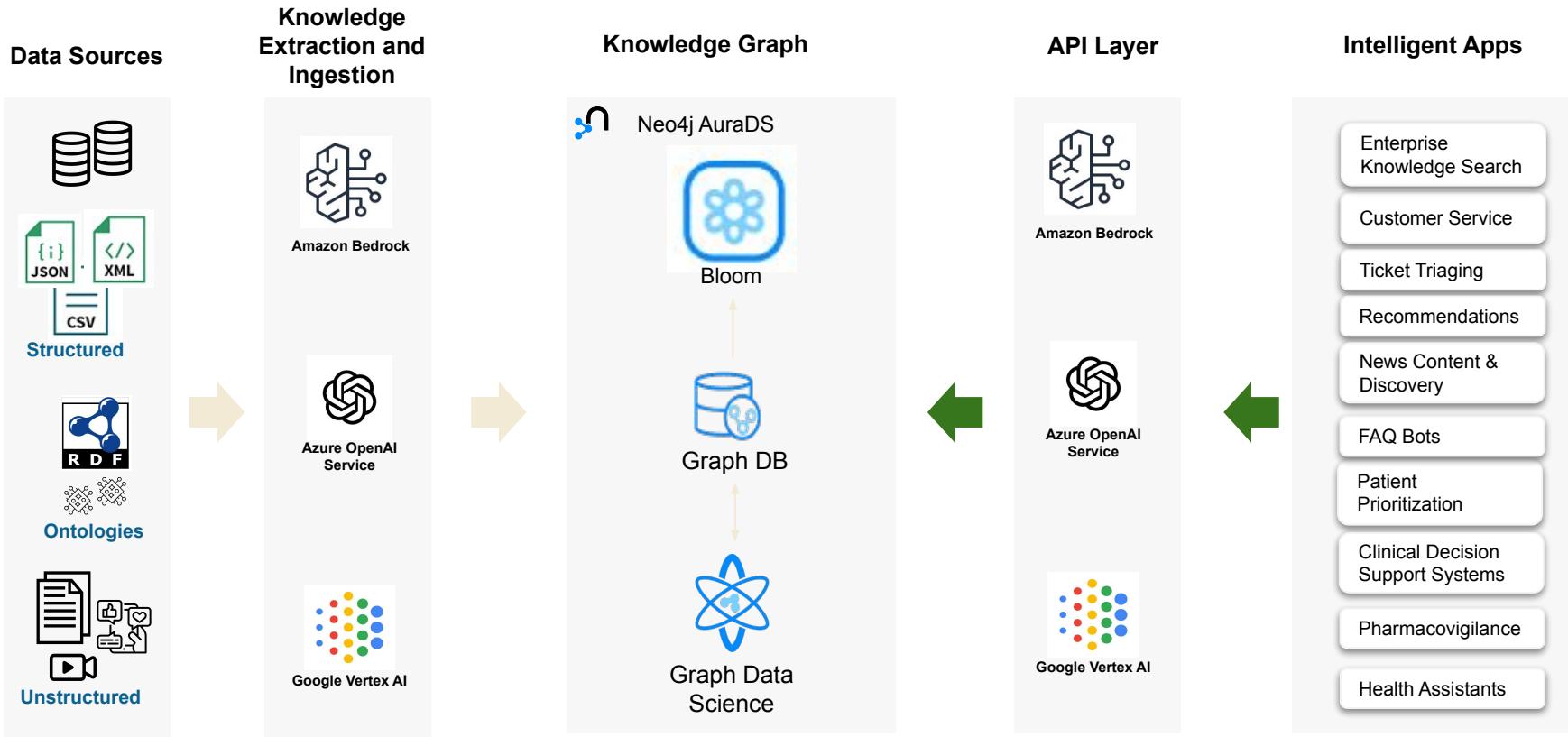
Representar fontes de dados como nós

Mapear relações entre resultados de pesquisa e nós de fonte de dados

Adicionar metadados ou anotações



Arquitetura de Referência de IA Generativa + Neo4j





KEEP
CALM
AND
SHOW ME
THE CODE

Building Neo4j Applications with Node.js

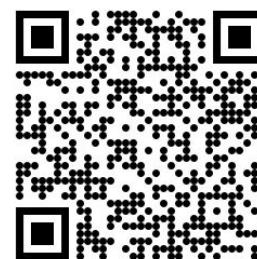
<https://graphacademy.neo4j.com/courses/app-nodejs/>



Comece agora mesmo GRATUITAMENTE!



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



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

Será o fim da linha dos relacionais?

Obrigado!

eliezer.zarpelao@neo4j.com

<https://www.youtube.com/@eliezerzarpelao>

<https://github.com/elizarp>