

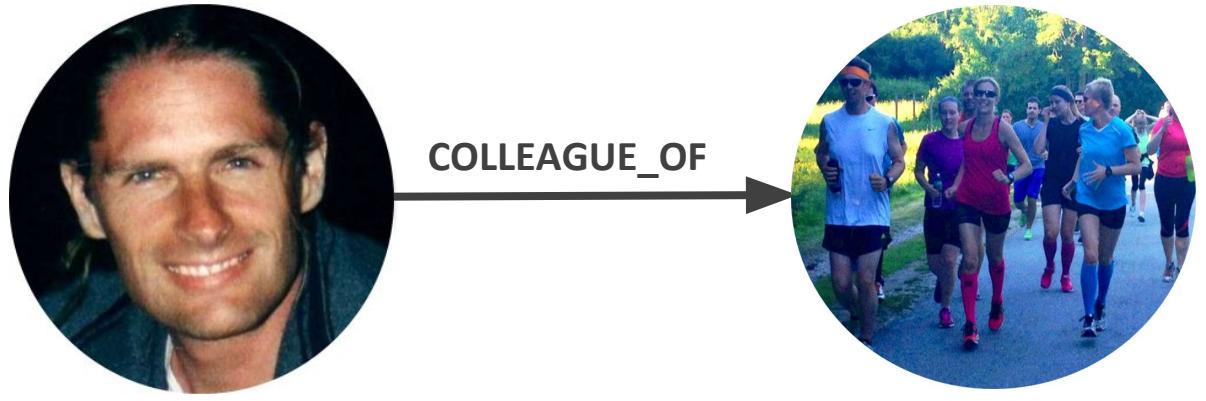
Neo4j Zoom

Cphbusiness 18th April, 2017

Maria Scharin



Today's first graph



```
(:Person { name:"Craig Taverner"} ) - [:COLLEAGUE_OF] -> (:Person { name:"Maria Scharin"} )
```

- Creators of Neo4j, the world's leading graph database
- Swedish founders
- 150 employees; main offices in
 - Malmö
 - San Mateo (HQ)
 - London
- Hiring engineers



Agenda

- Zoom Out - the Business Perspective
 - Why Graph Databases (and why now)?
 - Use Cases
- Zoom in - Performance and the Techie Details
 - Why is Neo4j so Fast?
- Demo

The Business Perspective

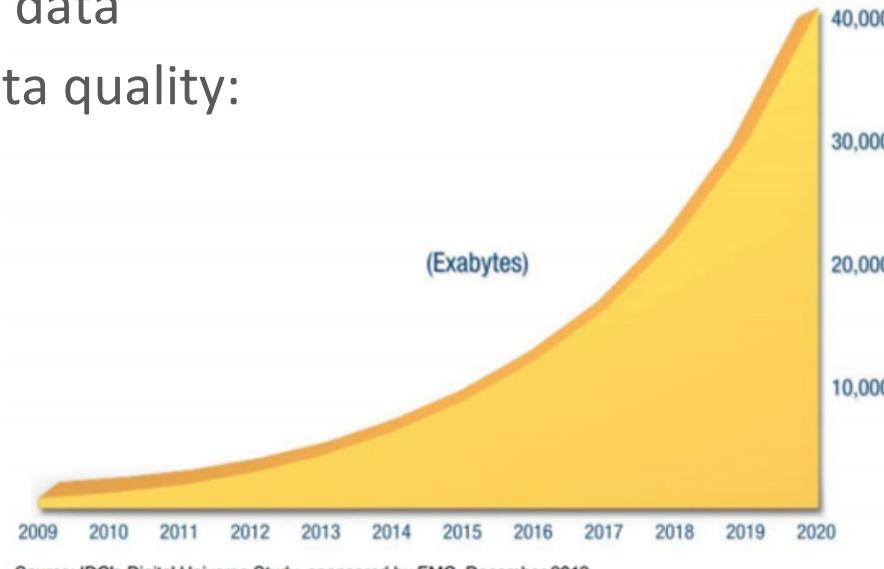
- Why Graph Databases...
... and why now?
- Use Cases
- Resources

Why Graph Databases?

Data Challenges



- Exponentially growing data
- Variety of data and data quality:
 - Structured
 - Semi-structured
 - Unstructured



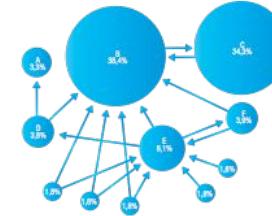
Source: IDC's Digital Universe Study, sponsored by EMC, December 2012

Why Graph Databases?

Graph technology has been transformational



Google



monster
Find better.™



LinkedIn



PayPal

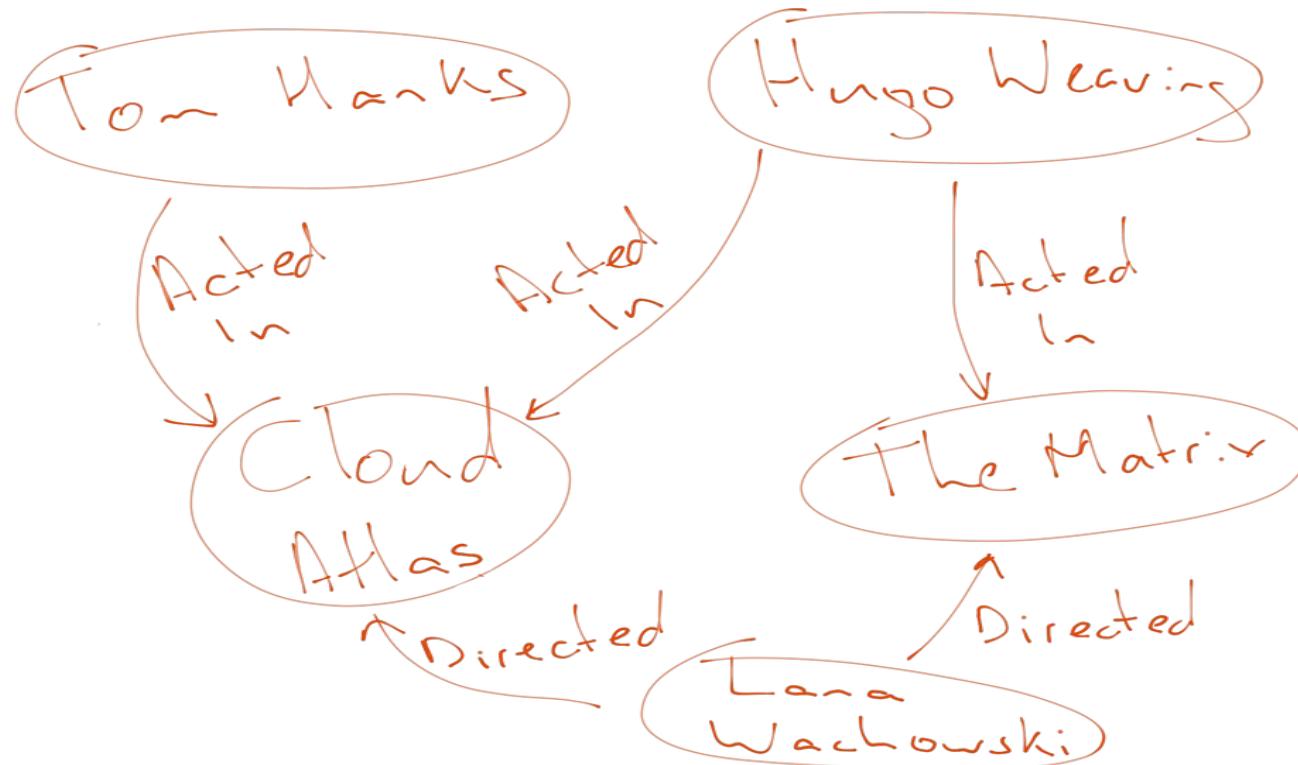


Keyword oriented

Pattern oriented

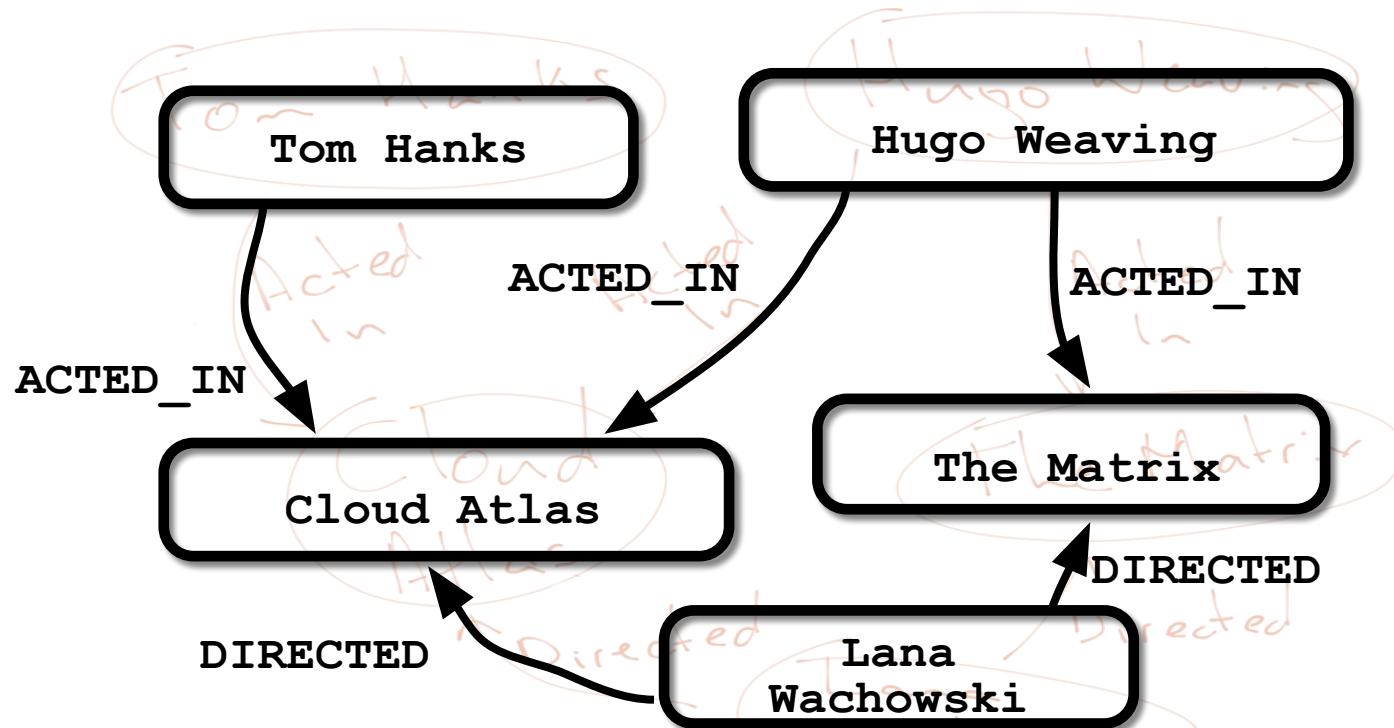
Why Graph Databases?

Whiteboard Friendliness



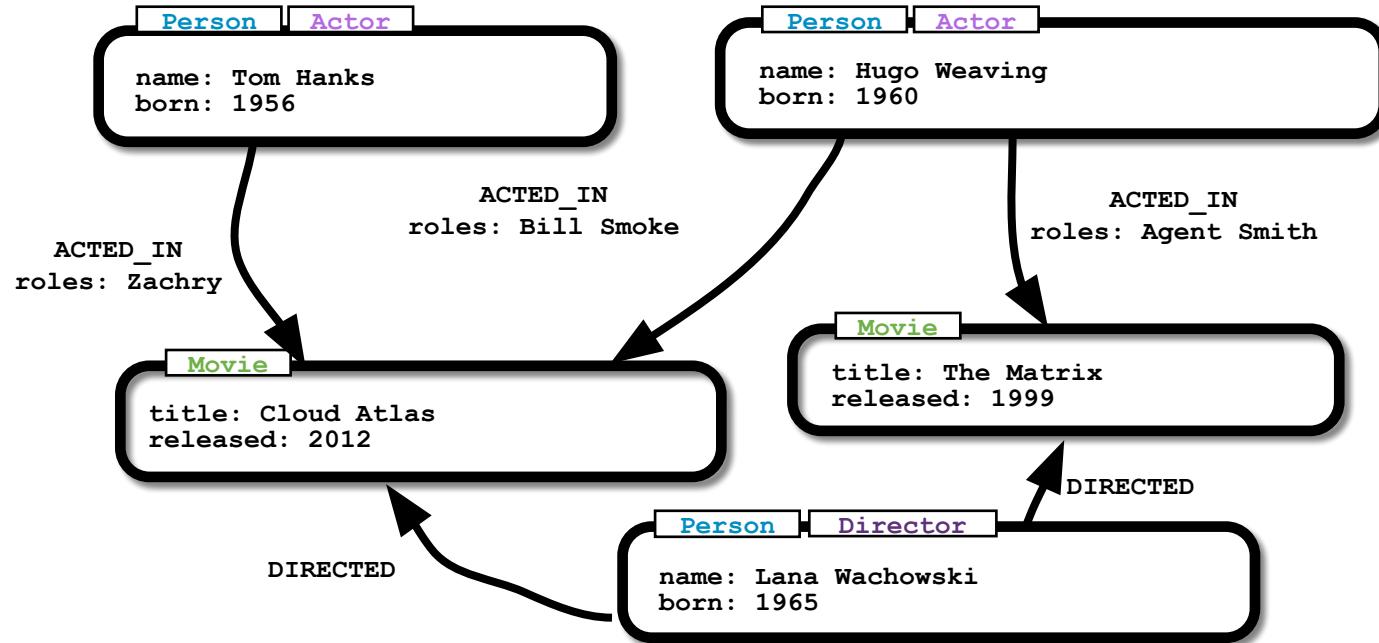
Why Graph Databases?

Whiteboard Friendliness



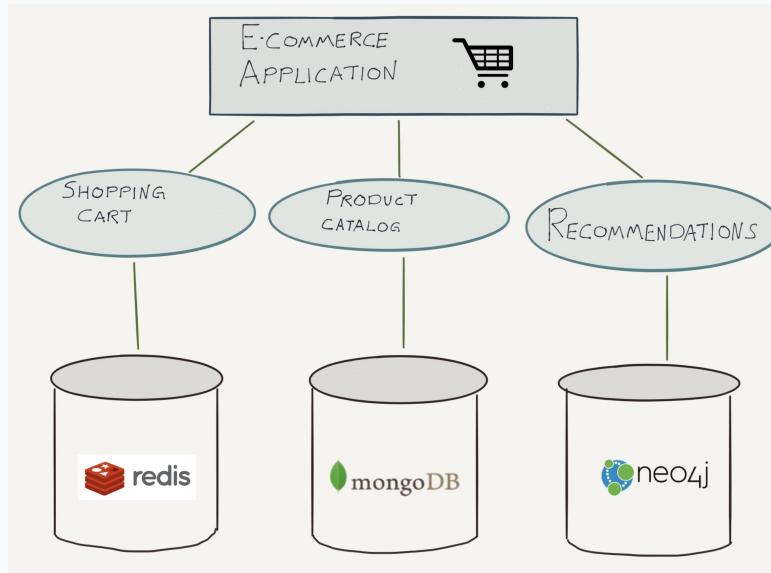
Why Graph Databases?

Whiteboard Friendliness



Why Graph Databases?

Polyglot Persistence: Use the *right tool for the task!*



Functionality	Database type
Shopping Cart	Rapid session reads / writes Key-value store
Orders / Product Catalog	Frequent reads Document
Customer social graph	Recommendation Graph

What is Polyglot Persistence: <http://www.jamesserra.com/archive/2015/07/what-is-polyglot-persistence/>

Webinar on Polyglot Persistence: <https://www.youtube.com/watch?v=50ifNtFAUXg>

The Business Perspective

- Why Graph Databases...
... and why now?
- Use Cases
- Resources

Use Cases



- Graph-based Search
- Network and IT Operations
- Identity and Access Management
- Master Data Management
- Fraud Detection
- Social Networks
- Real-Time Recommendations

Data Journalism

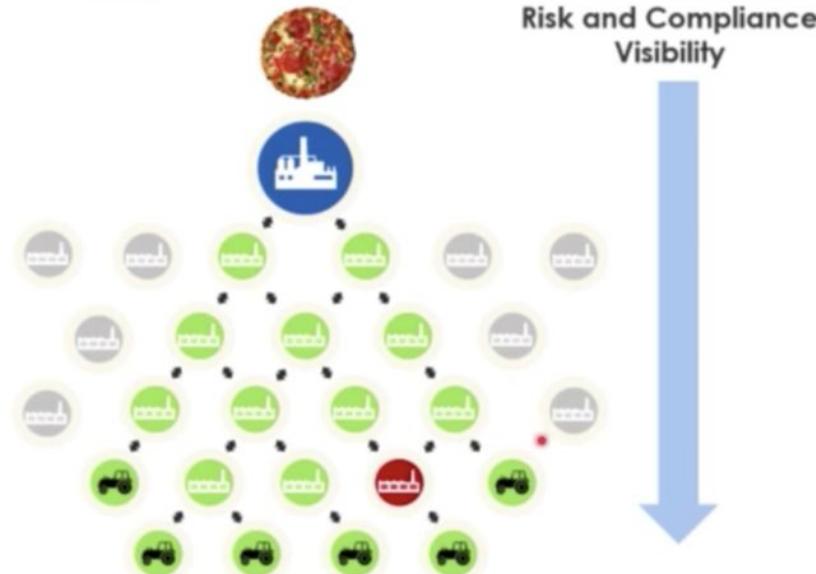
Use Cases

Graph-Based Search



Use Cases

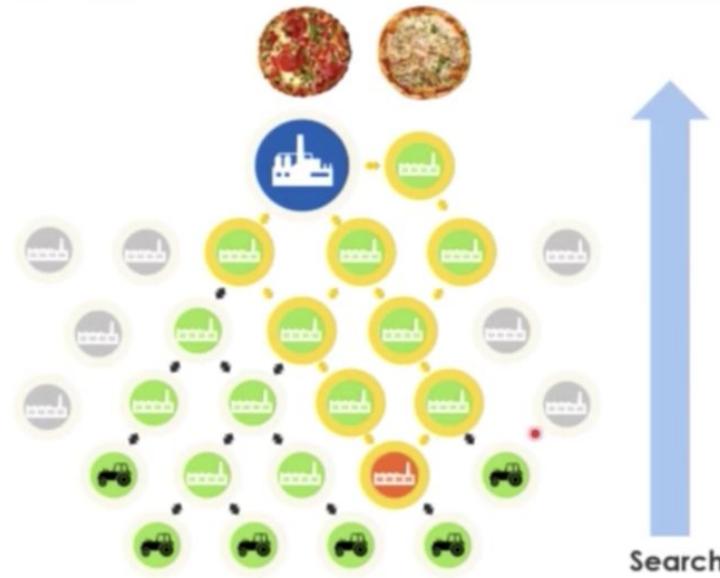
Graph-Based Search: Supply-Chain



Article: <https://neo4j.com/blog/consumer-trust-transparency-compliance-sustainability-food-industry/>
Video - Building trust through transparency: https://www.youtube.com/watch?v=_BzoSOBrBM4

Use Cases

Graph-Based Search: Supply-Chain



Article: <https://neo4j.com/blog/consumer-trust-transparency-compliance-sustainability-food-industry/>
Video - Building trust through transparency: https://www.youtube.com/watch?v=_BzoSOBrBM4

Use Cases

Graph-Based Search: Library Search



Find all publications from 2013 that have the tag "NOSQL"*

MATCH

```
(year)-[:IN_YEAR]-(issue)-[:ISSUE_OF]->(pub:Publication),  
(issue)-[:CONTAINS]->(content:Content)-[:TAGGED]->(nosql:Tag)
```

WHERE

```
year.year = 2013 AND  
nosql.name='NOSQL'
```

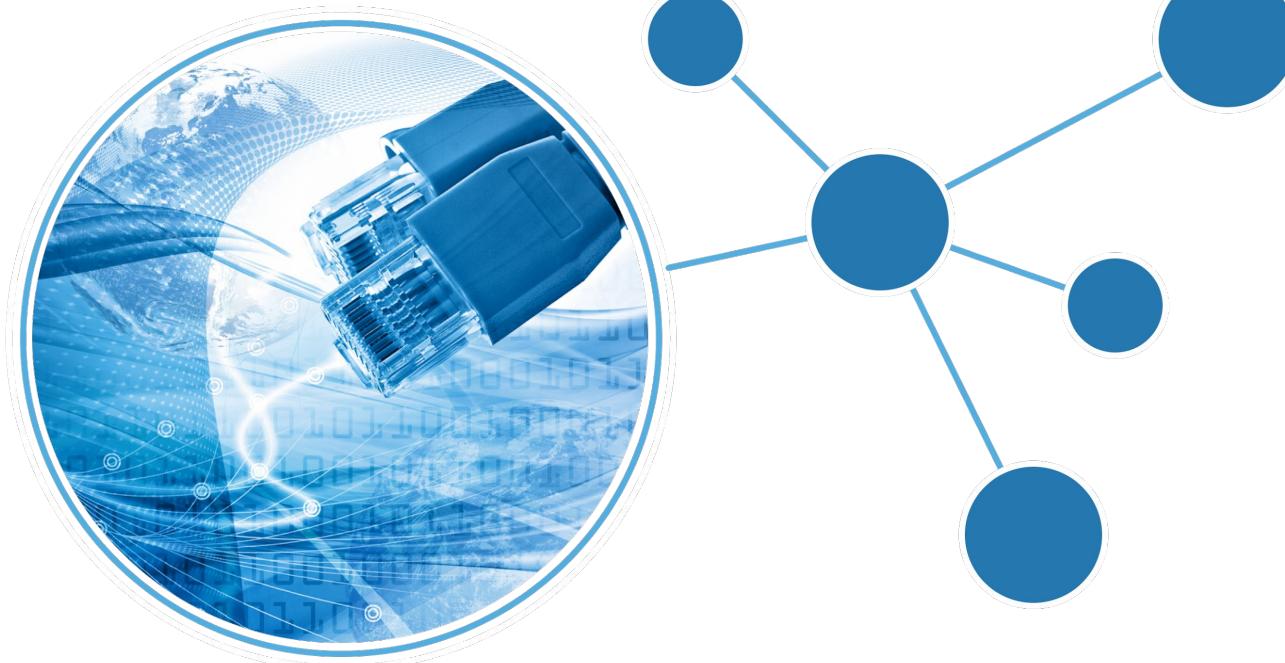
RETURN

```
content.title as Title, issue.title as Issue, pub.name as Publication
```

*GraphGist: <https://neo4j.com/graphgist/9a78922e-df9e-4dd5-a737-074363513b77/?ref=solutions>

Use Cases

Network and IT Operations



Use Cases

Network and IT Operations

Which network routes would be affected by the failure of a specific power supply unit.

```
MATCH (p:PowerSupply)-[:POWERS]->(failed:Router)  
      WHERE p.id = 'JHGUF457D8S'
```

```
WITH collect(failed) as failedRouters
```

```
MATCH (source:Network), (destination:Network) WHERE source <> destination
```

```
MATCH shortRoute=shortestPath((source)-[:ROUTE*..10]-(destination))
```

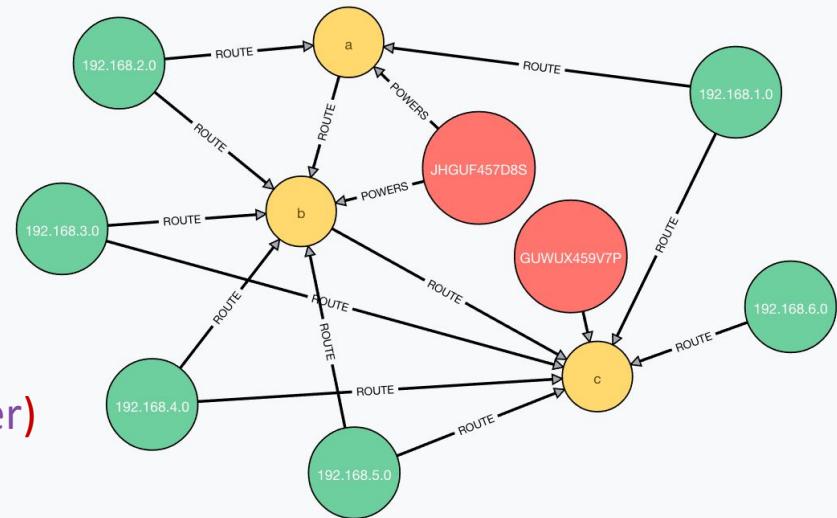
```
OPTIONAL MATCH route=(source)-[:ROUTE*..10]-(destination)
```

```
      WHERE NONE(d in nodes(route) WHERE d IN failedRouters)
```

```
WITH source, destination, shortRoute, collect(route) as routes WHERE length(routes) = 0
```

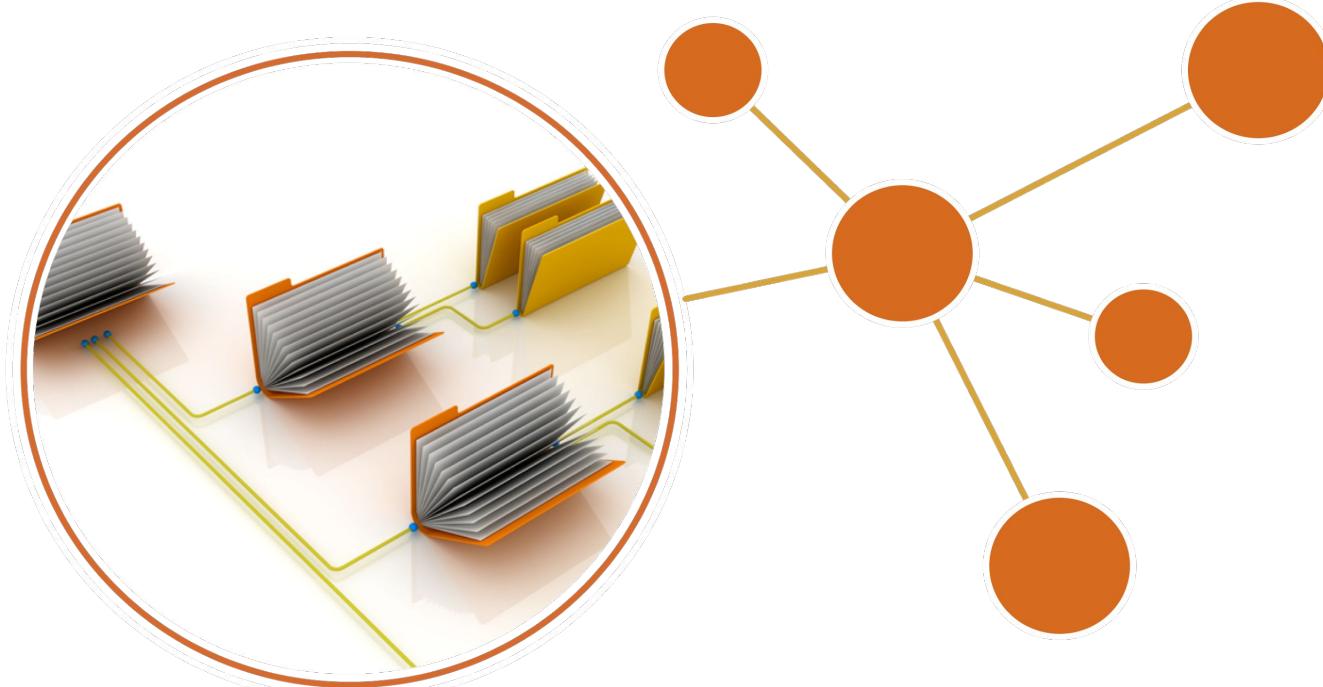
```
RETURN source.name, destination.name,
```

```
      extract(node in nodes(shortRoute) | node.name) as failed
```



Use Cases

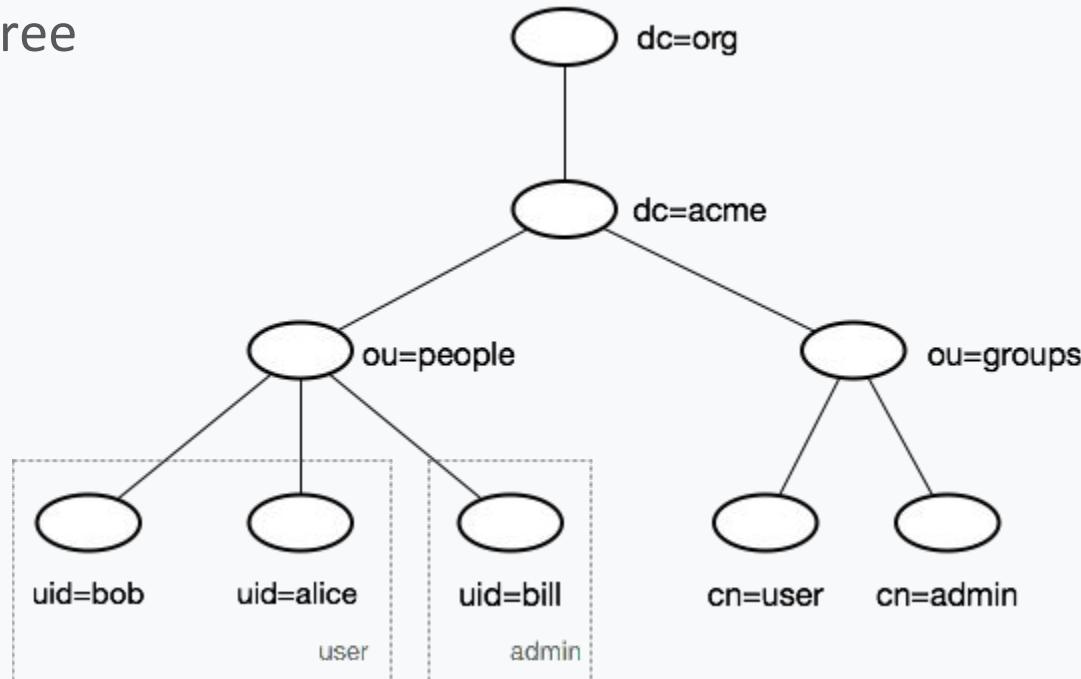
Identity and Access Management



Use Cases

Identity and Access Management

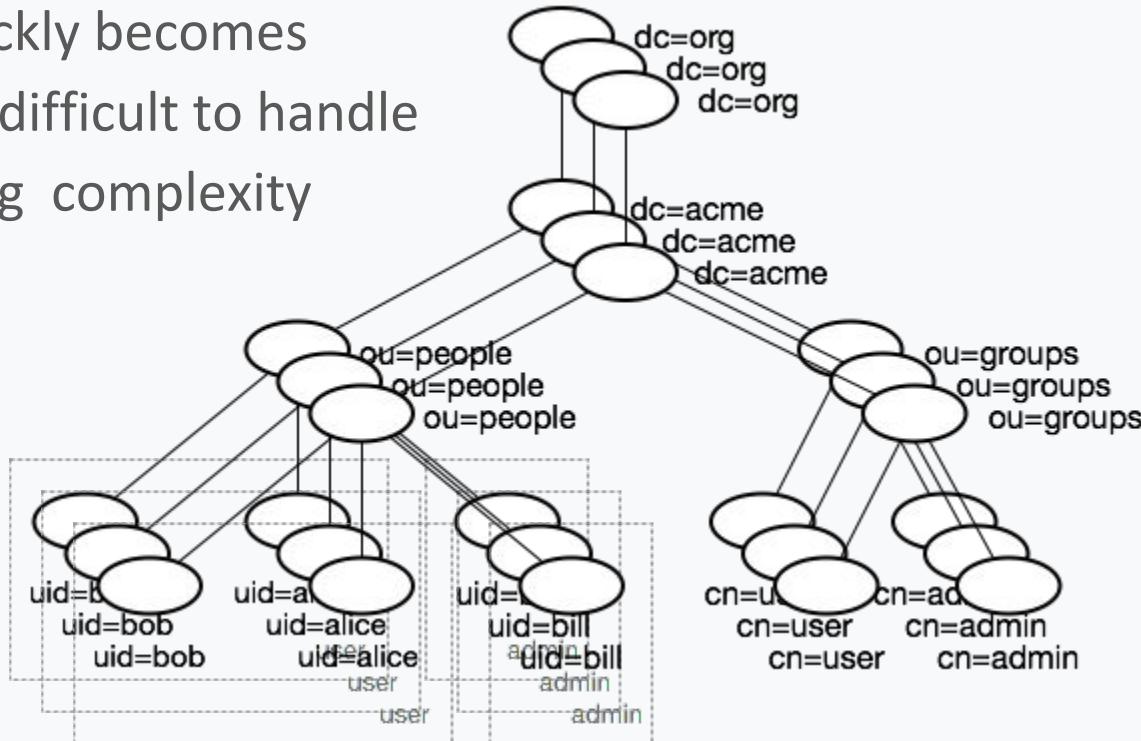
Simple LDAP tree



Use Cases

Identity and Access Management

LDAP tree quickly becomes unwieldy and difficult to handle with increasing complexity



Use Cases

Identity and Access Management

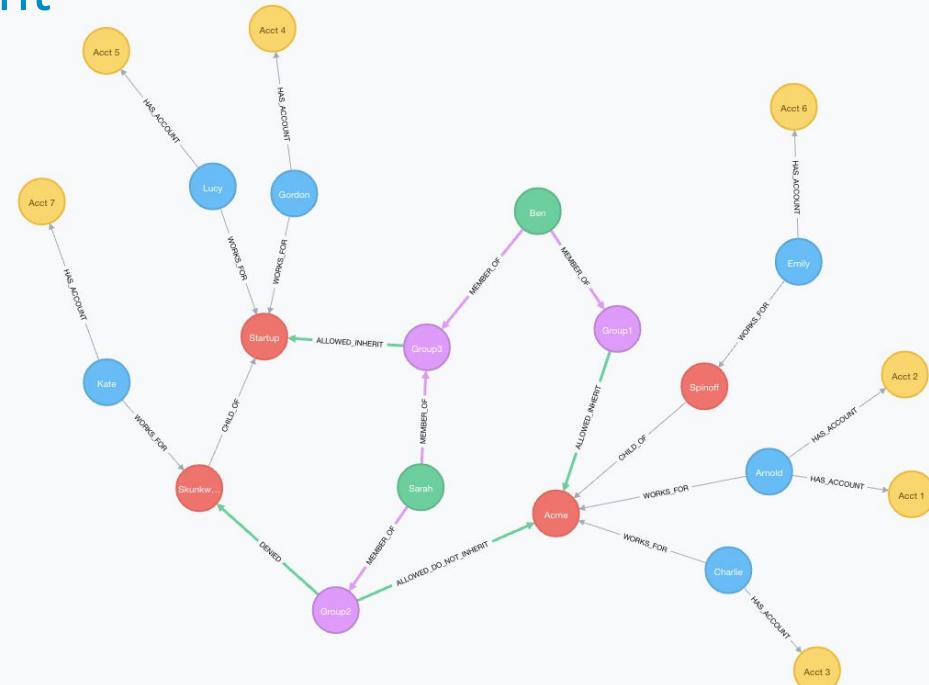
Example graph model

Types of security relationships:

:ALLOWED_INHERIT

:ALLOWED_DO_NOT_INHERIT

:DENIED



account(12)

administrator(4)

company(9)

employee(11)

group(7)

Example from GraphGist: <https://neo4j.com/graphgist/06bbc73b-80e9-45fc-a1d6-c0270745ddfb>

Use Cases

Identity and Access Management

List all the resources any administrator can access:

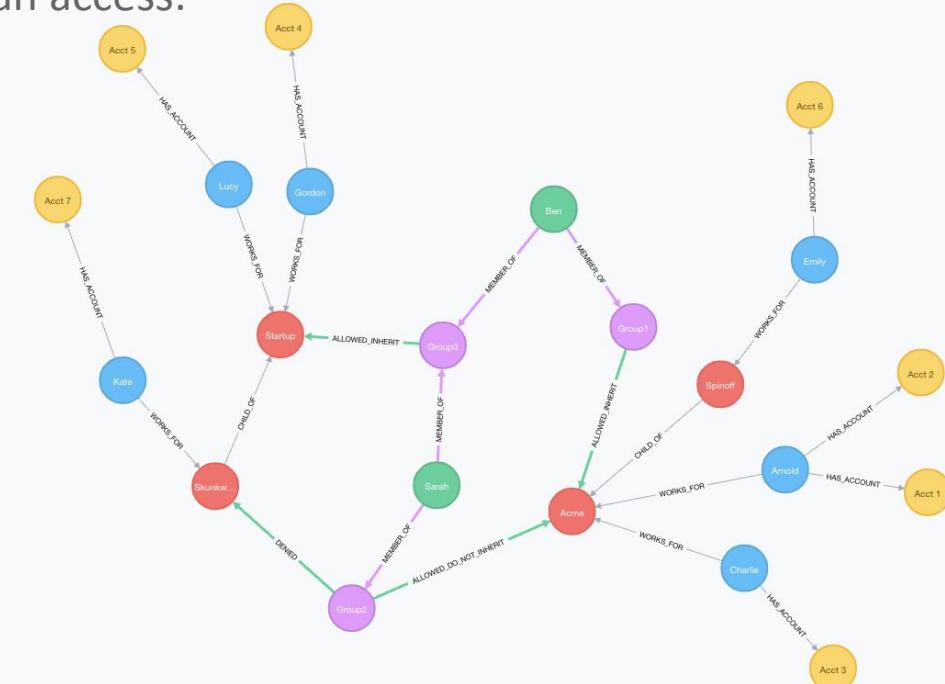
```
MATCH paths=(admin:administrator)-[:MEMBER_OF]->()-[:ALLOWED_INHERIT]->
()<-[:CHILD_OF*0..3]-(company)<-[:WORKS_FOR]-(employee)-[:HAS_ACCOUNT]->(account)
WHERE NOT ((admin)-[:MEMBER_OF]->()-[:DENIED]->())<-[:CHILD_OF*0..3]-(company))
RETURN admin.name
AS Admin, employee.name AS Employee, collect(account.name) AS Accounts
ORDER BY Admin ASC
UNION
MATCH paths=(admin)-[:MEMBER_OF]->()-[:ALLOWED_DO_NOT_INHERIT]->()-[:WORKS_FOR]-(employee)-
[:HAS_ACCOUNT]->(account)
RETURN admin.name
AS Admin, employee.name AS Employee, collect(account.name) AS Accounts
ORDER BY Admin ASC
```

Use Cases

Identity and Access Management

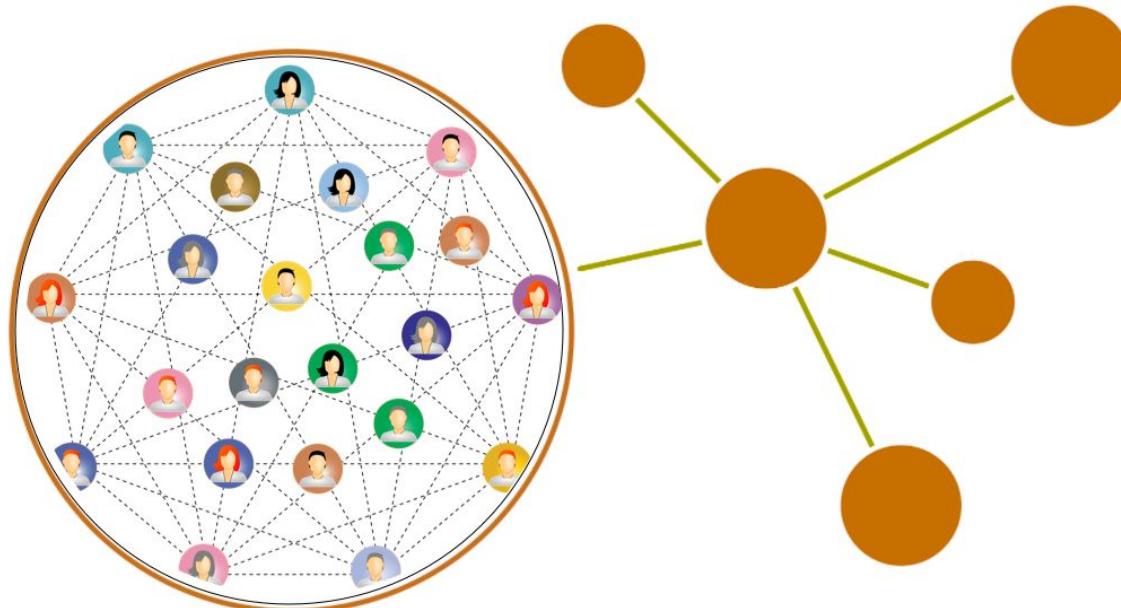
List all the resources any administrator can access:

Admin	Employee	Accounts
Ben	Charlie	[Acct 3]
Ben	Emily	[Acct 6]
Ben	Gordon	[Acct 4]
Ben	Arnold	[Acct 1, Acct 2]
Ben	Kate	[Acct 7]
Ben	Lucy	[Acct 5]
Sarah	Lucy	[Acct 5]
Sarah	Gordon	[Acct 4]
Sarah	Charlie	[Acct 3]
Sarah	Arnold	[Acct 1, Acct 2]



Use Cases

Master Data Management



Use Cases

Master Data Management

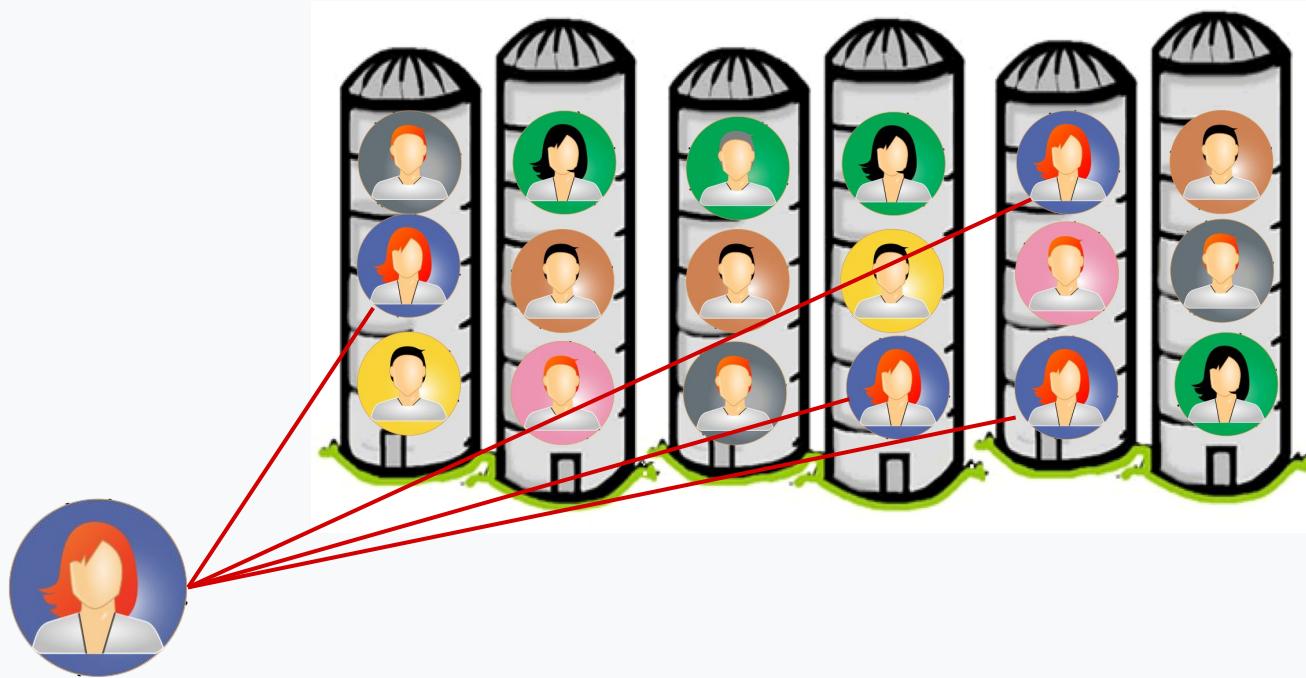
Data in legacy systems or *silos*



Use Cases

Master Data Management

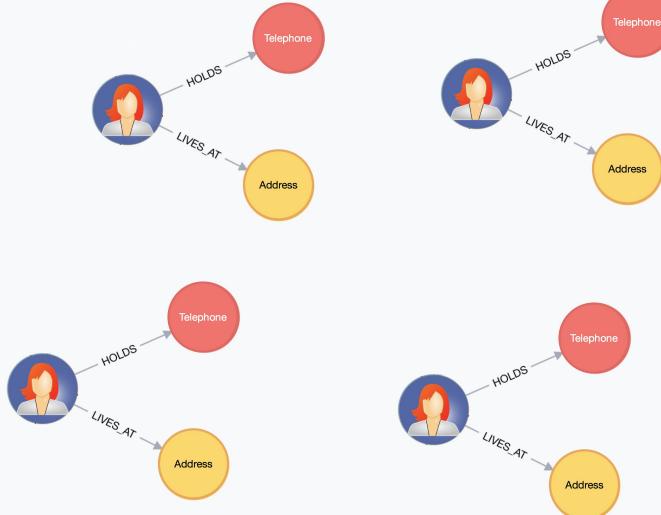
Data in legacy systems or *silos*



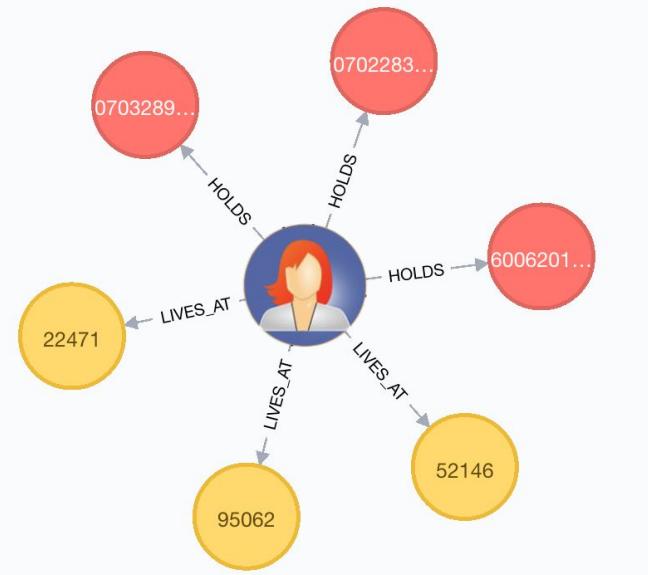
Use Cases

Master Data Management

Data in legacy systems or *silos*



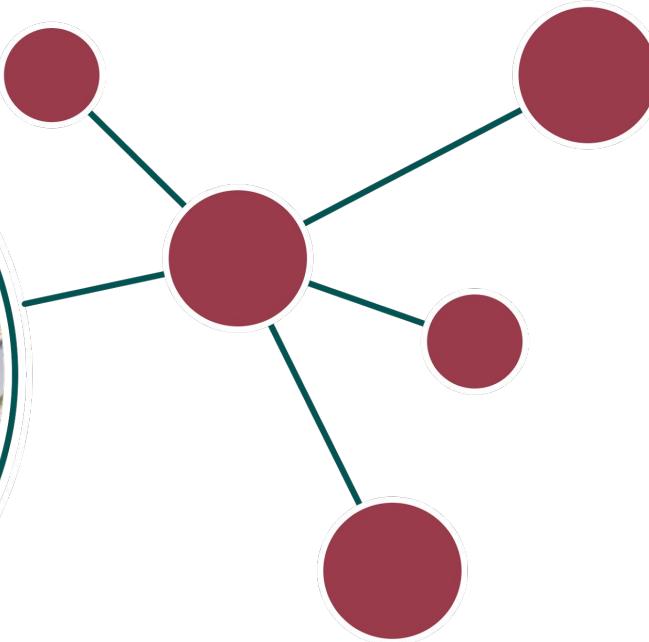
Which one is "the right" one?



The graph model enables
a richer *single version of the
truth*

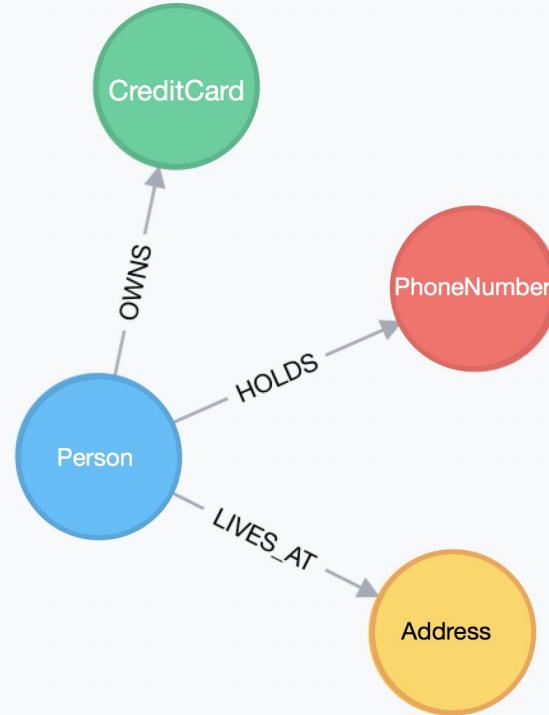
Use Cases

Fraud Detection



Use Cases

Fraud Detection - Credit Card Fraud



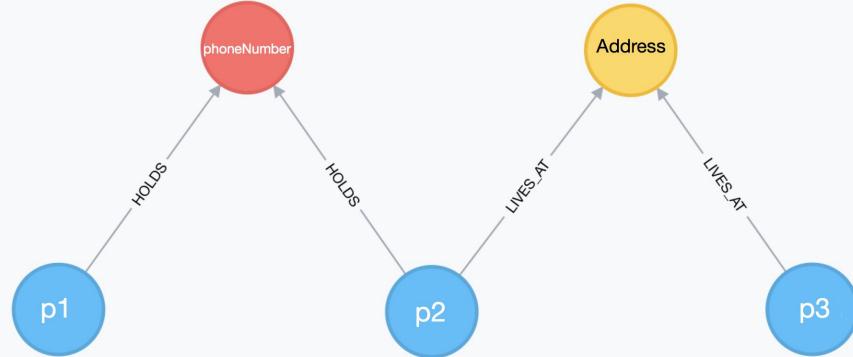
A simple data model:

Use Cases

Fraud Detection - Credit Card Fraud

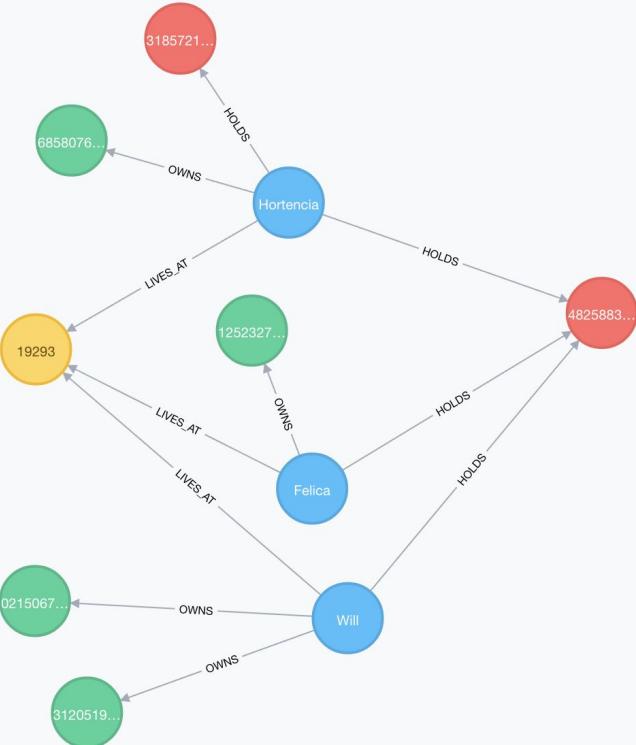
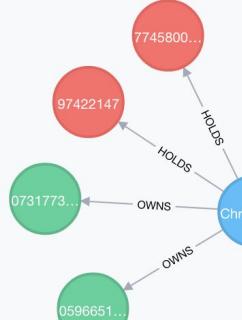
Show persons with suspicious patterns:

```
MATCH (p1:Person)-[:HOLDS|LIVES_AT]->()
<-[:HOLDS|LIVES_AT]-(p2:Person)
  -[:HOLDS|LIVES_AT]->()
<-[:HOLDS|LIVES_AT]-(p3:Person)
WHERE p1 <> p2 AND p2 <> p3 AND p3 <> p1
```



Use Cases

Fraud Detection - Credit Card Fraud



Is this match suspicious?

```
MATCH (p1:Person)-[:HOLDS|LIVES_AT]->()
```

```
<-[:HOLDS|LIVES_AT]-(p2:Person)  
-[:HOLDS|LIVES_AT]->()
```

```
<-[:HOLDS|LIVES_AT]-(p3:Person)
```

```
WHERE p1 <> p2 AND p2 <> p3 AND p3 <> p1
```

```
RETURN p1, p2, p3
```

Use Cases

Fraud Detection - Credit Card Fraud

Find chains that are between 8 and 10 hops long:

```
MATCH p=(n:Person)-[:HOLDS|LIVES_AT*8..10]-(m:Person)
```

```
WHERE m <> n
```

```
WITH nodes(p) as nodes
```

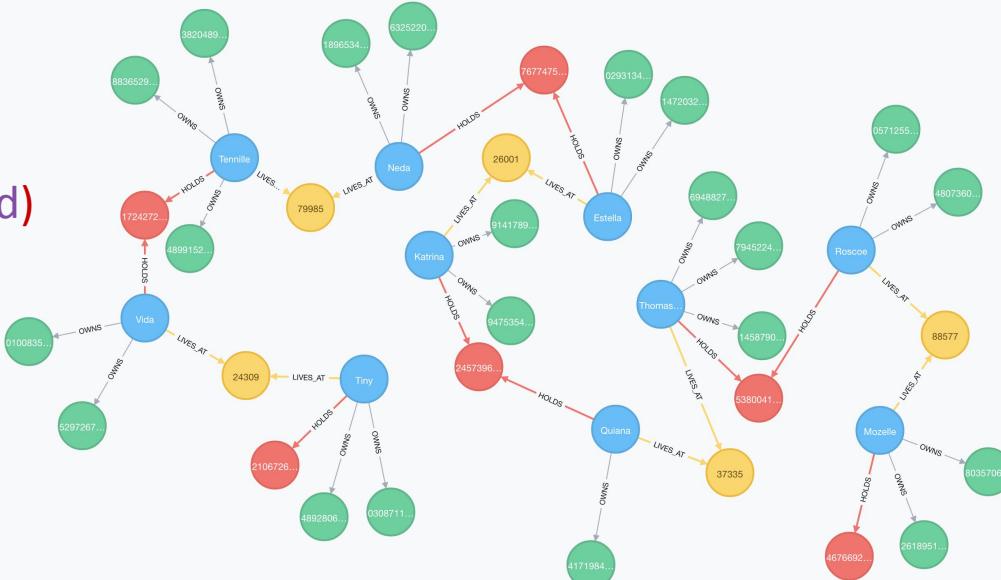
```
UNWIND nodes as n
```

```
MATCH (n)-[:HAS_CARD]->(c:CreditCard)
```

```
WHERE NOT c.suspended = true
```

```
SET c.status = "temporary-suspended"
```

```
RETURN n.name, c.number
```



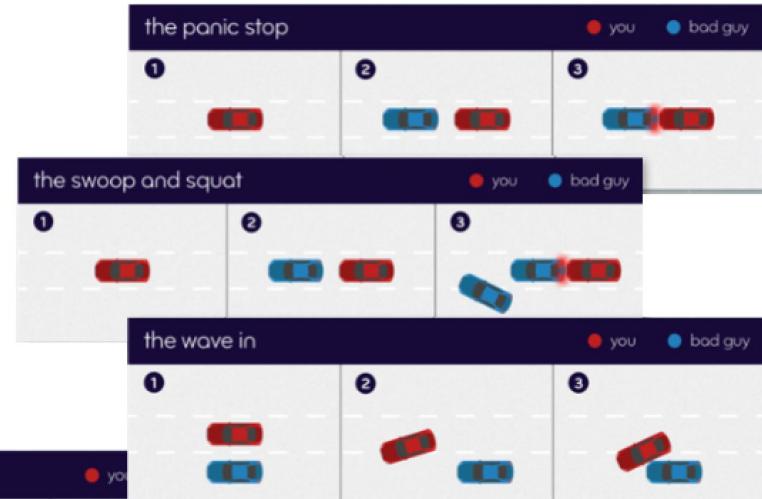
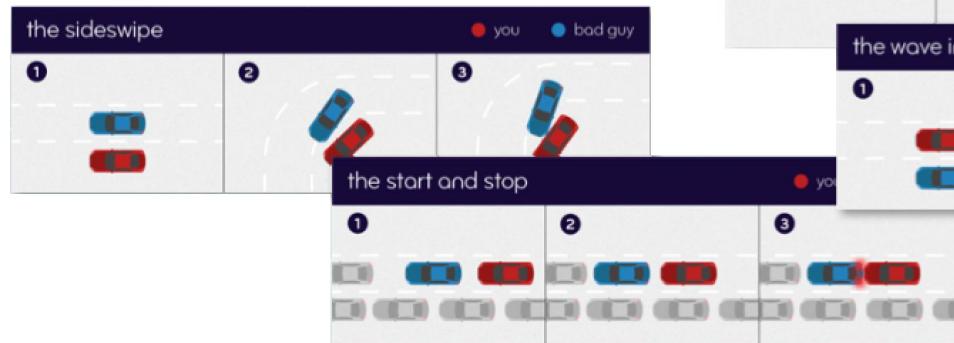
Use Cases

Insurance Fraud "Whiplash for Cash"

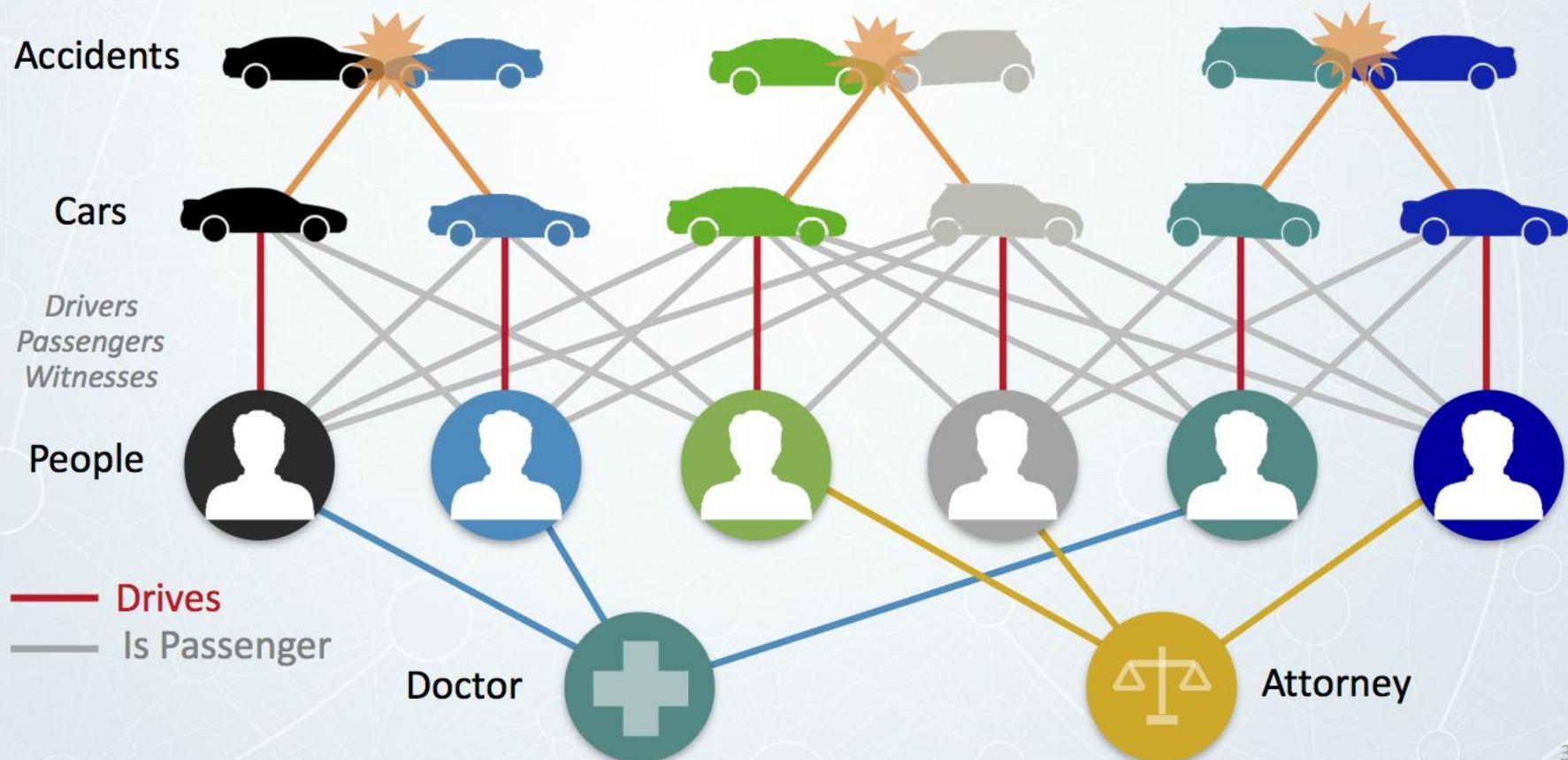


Paper collisions

Insurance scammers invent automobile accidents complete with fake drivers, passengers and witnesses

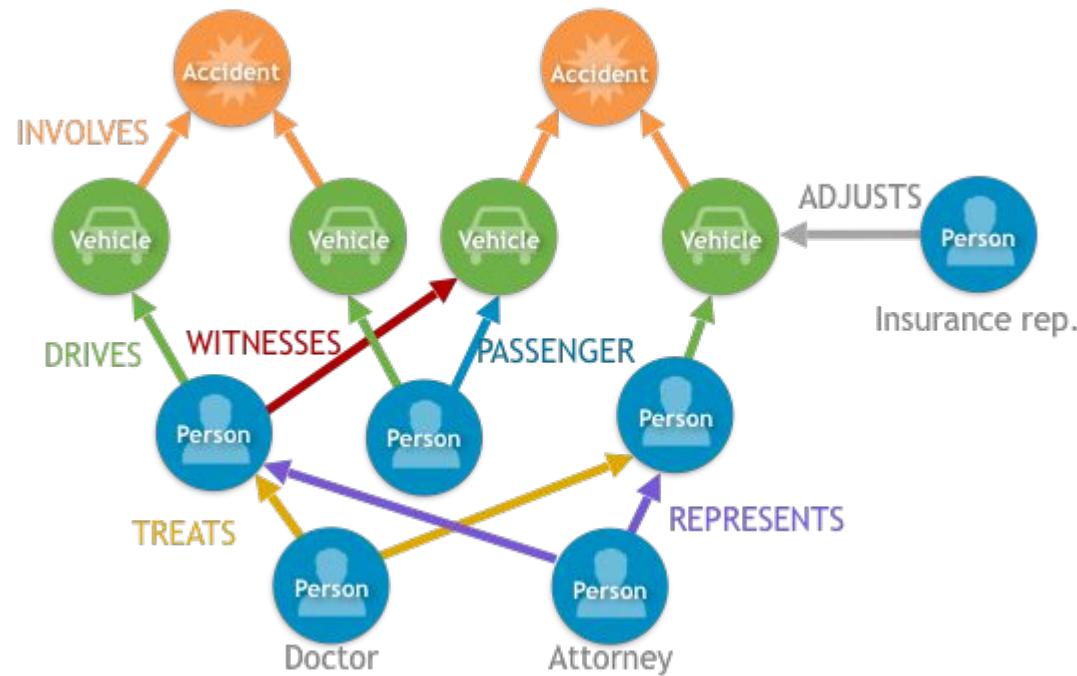


Whiplash for Cash Example



Use Cases

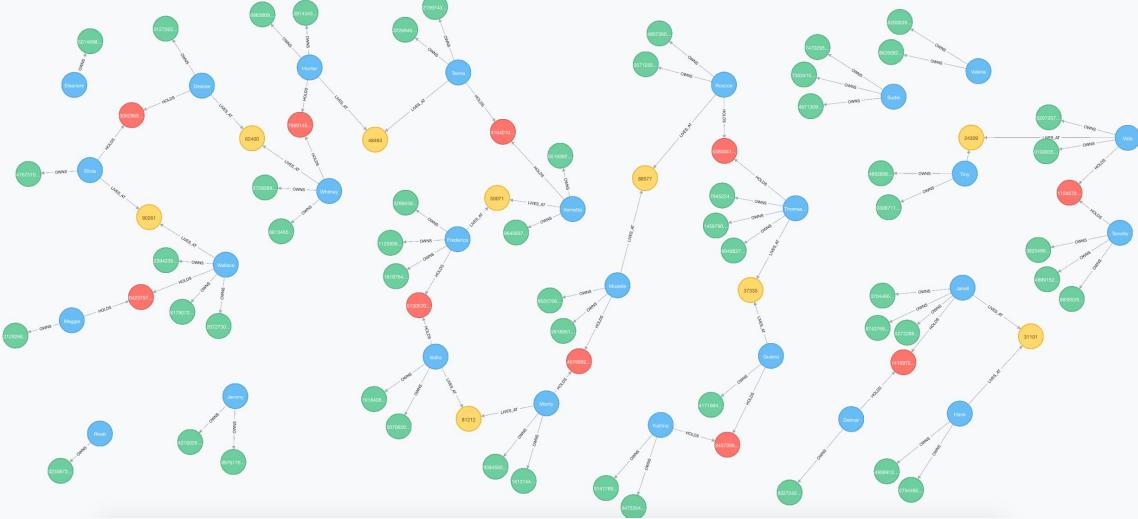
Whiplash for Cash as a Graph



Use Cases

Fraud Detection

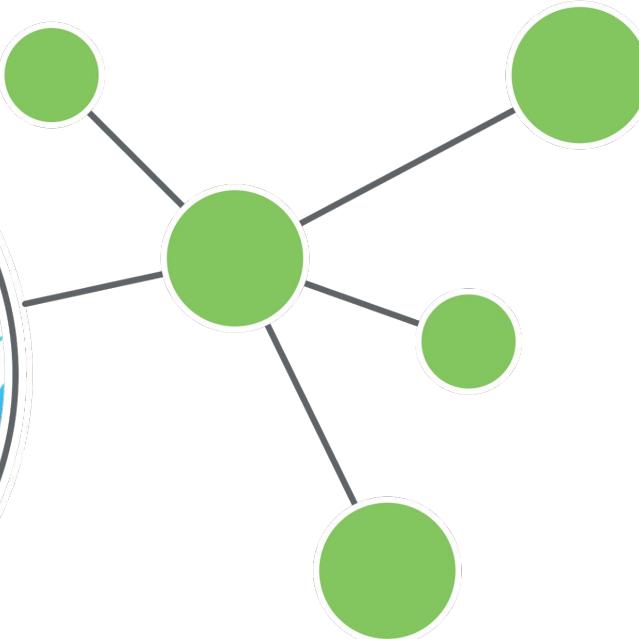
Traditional technologies
don't provide the right tools:



- Large amounts of data and *relationships* must be processed
- New data and relationships are continually being added
- Fraud rings must be uncovered *in real-time* to prevent fraud

Use Cases

Social Network

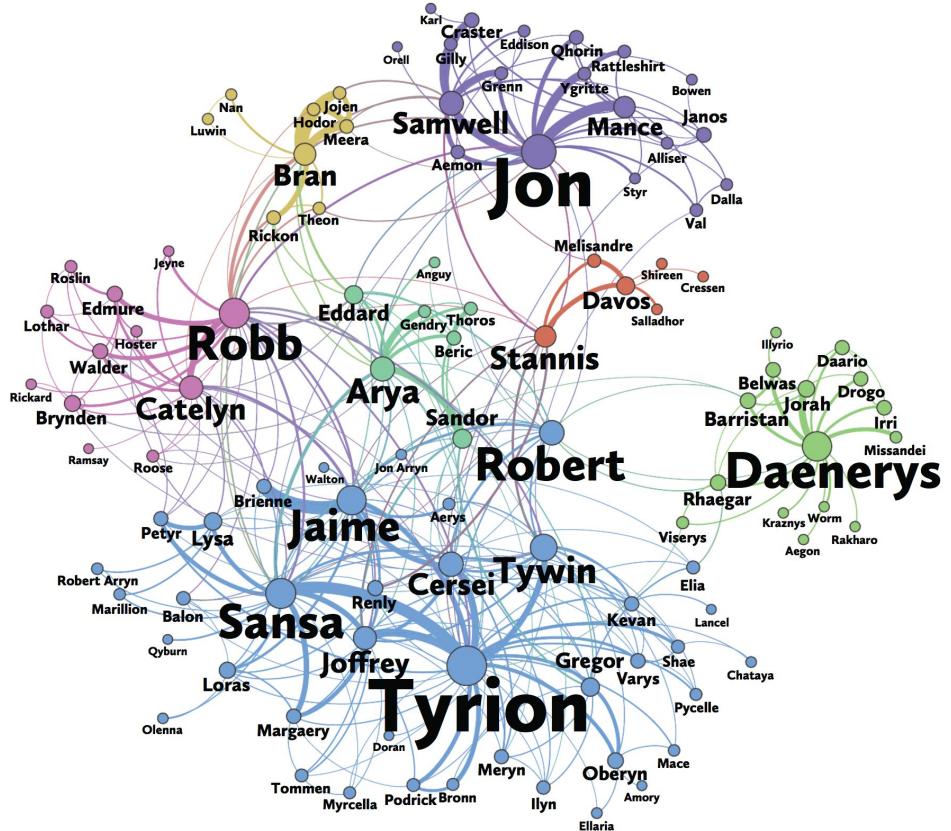


Use Cases

Social Network

Game of Thrones

Book #3: *A Storm of Swords*

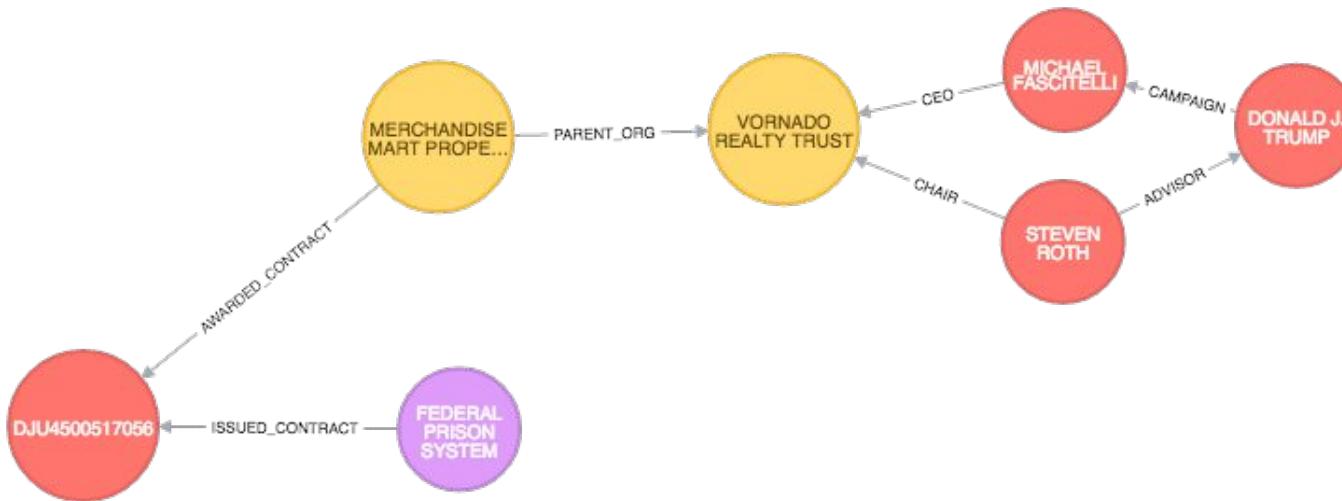
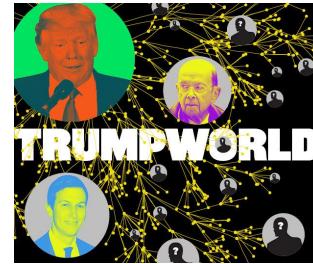


<http://www.maa.org/sites/default/files/pdf/Mathhorizons/NetworkofThrones%20%281%29.pdf>

Additional analysis: <http://www.lyonwj.com/2016/06/26/graph-of-thrones-neo4j-social-network-analysis/>

Use Cases

Social Network: Trump World



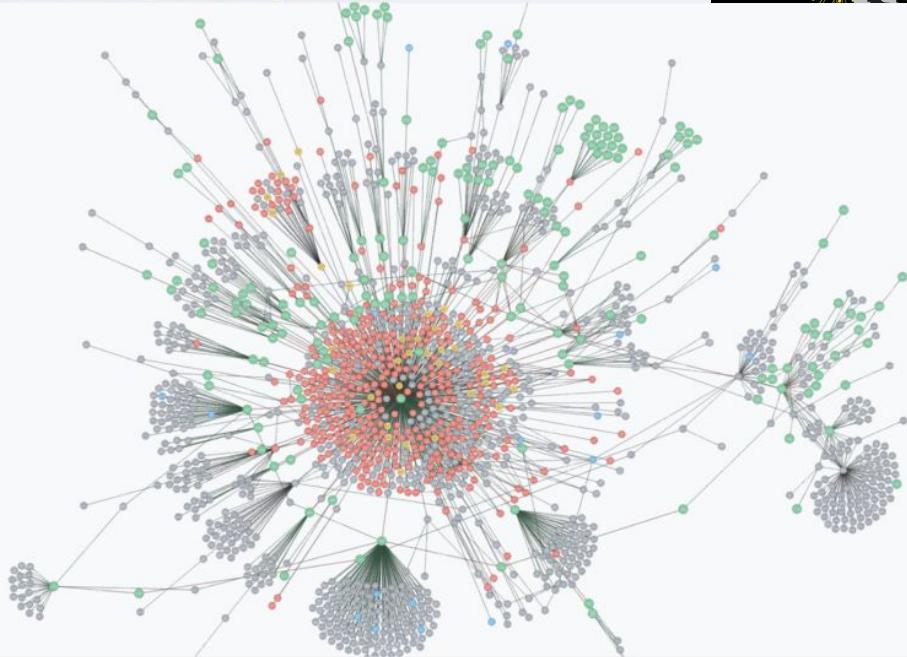
Sandbox containing the data: <https://neo4j.com/sandbox-v2>

Articles: <https://neo4j.com/blog/buzzfeed-trumpworld-dataset-neo4j/>,
<http://www.lyonwj.com/2017/01/30/trumpworld-us-contracting-data-neo4j/>
Video: <https://www.youtube.com/watch?v=pW9boJoUxO0>

Use Cases

Social Network: Trump World

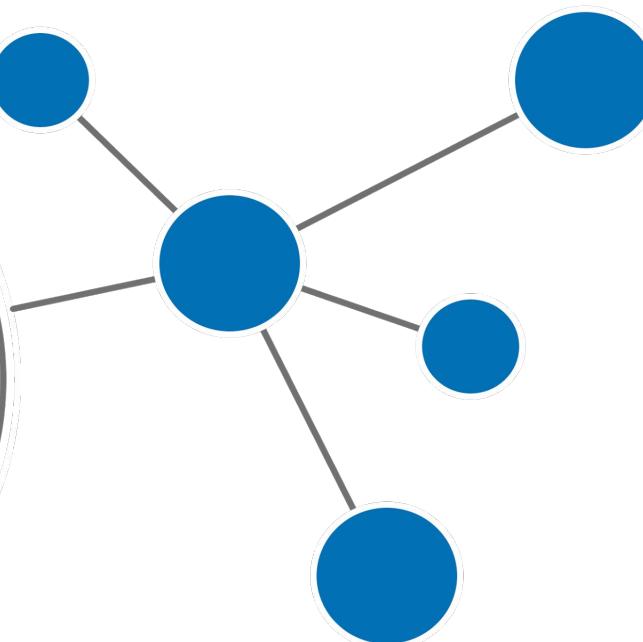
1514 nodes, 1857 relationships.
Bank(17), Hotel(34), Organization(133), Person(164), Trump(405)
CONNECTED_TO(611), INVOLVED_WITH(1139), RELATED_TO(107)



Displaying 1514 nodes, 1857 relationships.

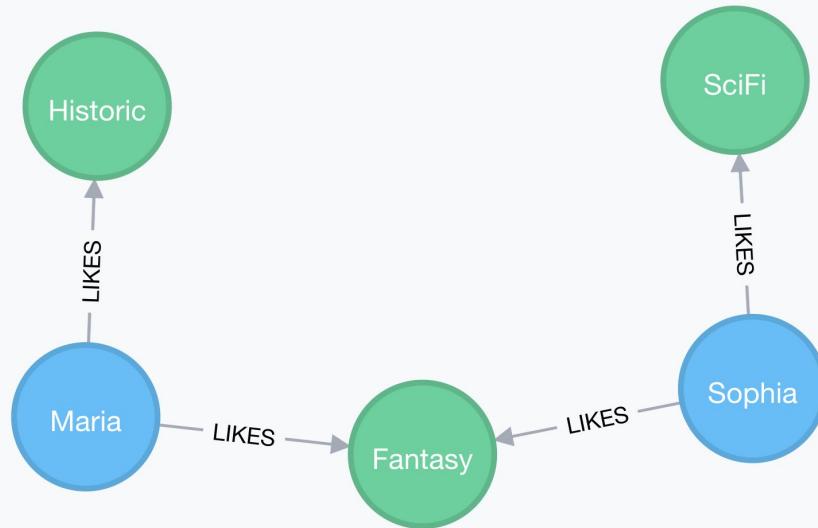
Use Cases

Real-Time Recommendations



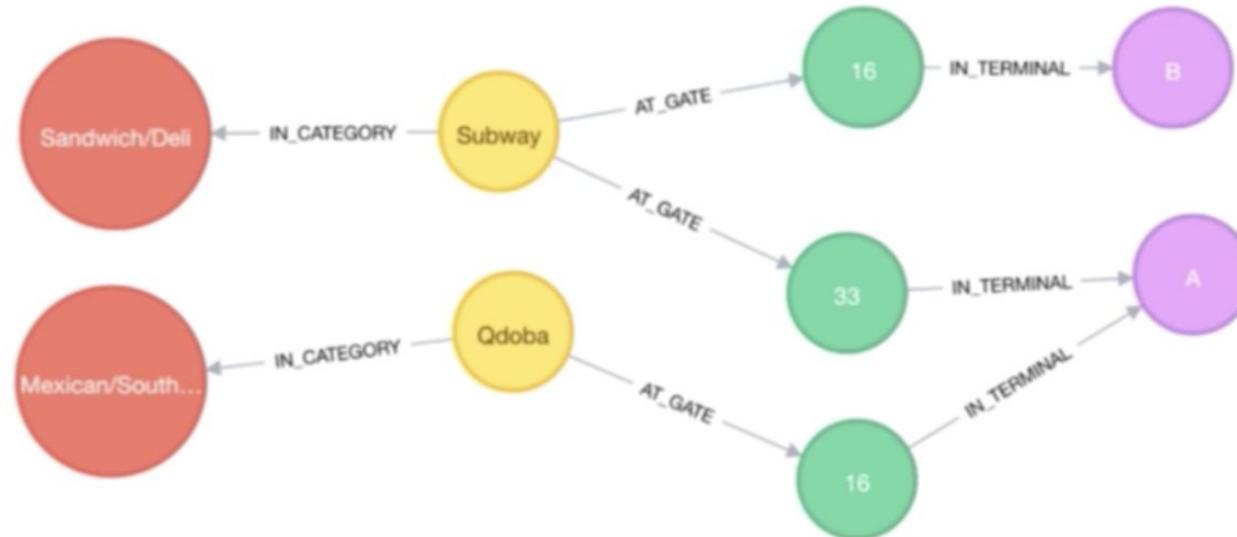
Use Cases

Real-Time Recommendations



Use Cases

Real-Time Recommendations

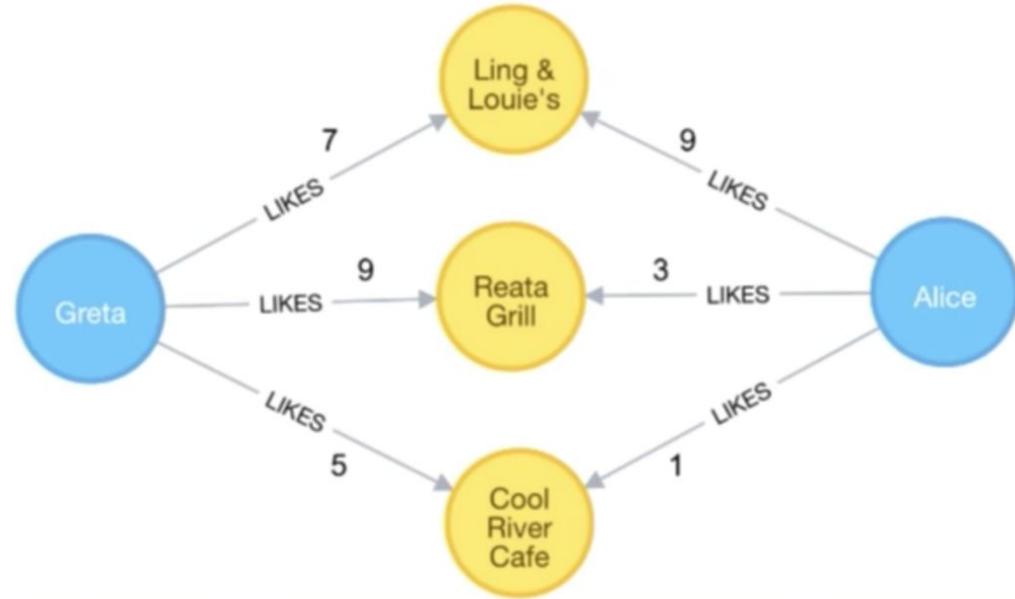


Article: <https://dzone.com/articles/building-a-real-time-recommendation-engine-with-da>

Video: <https://youtu.be/60E2WV4iwIg>

Use Cases

Real-Time Recommendations



Use Cases

Data Journalism



The International Consortium of Investigative Journalists (ICIJ) logo is displayed, consisting of the acronym "ICIJ" in white on a red square followed by the full name in large, bold, black letters. Below the logo, a dark blue background features text about the consortium's size and mission. At the bottom, there is a small graphic of a network graph with several colored nodes (red, purple, yellow, green) connected by white lines.

**Almost 200 journalists
Based in 65 countries**

*"Our aim is to bring journalists from different countries together
in teams - eliminating rivalry and promoting **collaboration**.
Together, we aim to be the
world's best **cross-border investigative team**."*

<https://www.youtube.com/watch?v=S20XMQyvANY>

Blog post: <https://neo4j.com/blog/icij-neo4j-unravel Panama-papers/>

Resources

- <https://neo4j.com/resources/>
- <https://neo4j.com/use-cases/>
- <https://neo4j.com/sandbox-v2/>
- <https://neo4j.com/graph-databases-book/>
- <https://neo4j.com/download/community-edition/>
- Neo4j Browser & Browser Guides
- <https://neo4j.com/docs/>
- <http://www.lyonwj.com>
- <https://maxdemarzi.com>
- Recorded talks at <http://graphconnect.com/>

