Investigating the Panama Papers with Open-Source Tools like Neo4j

Michael.Hunger@neo4j.com
Graphs in Humanities - Mainz
January 2017
A GLOBAL INVESTIGATION
THE PANAMA PAPERS
Politicians, Criminals and the Rogue Industry That Hides Their Cash

#PanamaPapers
(Michael)-[:WORKS_FOR]->(Neo4j)

michael@neo4j.com | @mesirii | github.com/jexp | jexp.de/blog

Michael Hunger - Caretaker General @Neo4j
Offshore Companies have legal uses.

There is no suggestion that parties listed in the documents have broken the law or acted improperly.
Organization of ca. 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.”

icij.org/about
Panama Papers series named

*Investigation of the Year*

Thu, June 16th 2016, Vienna, Austria

Project by International Consortium of Investigative Journalists wins Data Journalism Award
You may remember them from...
Latest Leak: #BahamasLeak
Former EU Official Among Politicians Named in New Leak of Offshore Files from The Bahamas

A cache of leaked documents provides names of politicians and others linked to more than 175,000 Bahamian companies registered between 1990 and 2016.

For years, Neelie Kroes traveled Europe as one of the continent’s senior officials, warning big corporations that they couldn’t “run away” from the European Union’s rules.

The Dutch politician sympathized with average citizens who worried they’d been left to pay the bills “as infringers cream off the extra profits.”

As the EU’s commissioner for competition policy from 2004 until 2010, she was Europe’s top corporate enforcer and made Forbes magazine’s annual list of the “World’s 100 Most Powerful Women” five times.

What Kroes never told audiences – and didn’t tell European Commission officials in mandatory disclosures – was that she had been listed as a director of an offshore company in the Bahamas, the Caribbean tax haven whose secrecy and tax structures have attracted multinational companies and criminals alike.

How did it start?
214,000
letterbox companies
Hello. This is John Doe. Interested in data?
We're very interested.
There are a couple of conditions. My life is in danger.
We will only chat over encrypted files.
No meeting, ever.
The choice of stories is obviously up to you.

Why are you doing this?
I want to make these crimes public.
Core Team of the ICIJ

ICIJ leadership

US-based team

Data & Research Unit

Africa desk
+370 journalists in 80 countries
Biggest leak in the history of data journalism just went live, and it's about corruption.
panamapapers.sueddeutsche.de/en/
Amount of data compared to previous leaks

1.7 GB Cablegate/Wikileaks (2010)

260 GB Offshore-Leaks/ICIJ (2013)

4 GB Luxemburg-Leaks/ICIJ (2014)

3.3 GB Swiss-Leaks/ICIJ (2015)

≈2.6 TB Panama Papers/ICIJ (2016)
Inside the 2.6 TB of Data

- Emails: 4,804,618
- Database formats: 3,047,306
- PDFs: 2,154,264
- Images: 1,117,026
- Text documents: 320,166
- Other: 2,242
While this might not be “Big Data” for us, it definitely is for Journalists.
Who is working on it?

+370 journalists
+100 media organizations
80 countries
1 Year
Data Team:
3 Data Journalists +
3 Developers!
1) Collaboration Supported by OSS Tools & Encryption
2) Using OSS to allow non-developers to make sense of highly connected data
Data Processing

- Raw Files
- Database
- Meta-Data

Raw Text

- OCR
- STORE
- POWER

Search

Discovery
Investigators used Nuix’s optical character recognition to make millions of scanned documents text-searchable. They used Nuix’s named entity extraction and other analytical tools to identify and cross-reference the names of Mossack Fonseca clients through millions of documents.
Lucene syntax queries with proximity matching!
Stack

Unstructured data extraction
- Nuix professional OCR service
- ICIJ Extract (open source, Java: https://github.com/ICIJ/extract), leverages Apache Tika, Tesseract OCR and JBIG2-ImageIO.

Structured data extraction
- A bunch of Python

Database
- Apache Solr (open source, Java)
- Redis (open source, C)
- **Neo4j (open source, Java)**

App
- Blacklight (open source, Rails)
- Linkurious (closed source, JS)
Exposed the offshore holdings of 12 current and former world leaders.

Dealings of 128 more politicians and public officials around the world.
Exclusive
The secret $2bn trail of deals that lead all the way to Putin

From Panama to the Kremlin, the offshore network that made Russian president and his friends fabulously wealthy

Panama Papers

113,000
113
26

26

Bigger leak in history

"Dreck über dem Deck"

From a source with knowledge

214 488

214 488

"Pien de Luxe"

Nederlandse vennootschappen mogelijk gebruikt voor betaling ongeweld Fifa

How secret offshore money helps fuel Miami’s luxury real-estate boom

The Secret Shell Game

Massive leak exposes how the wealthy and powerful hide their money

#panamapapers
Main goals, achieved:

1) Uncover the truth
2) Assure whistle blower safety
Disconnected Documents
Context is King

name: “John”
last: „Miller“
role: „Negotiator“

name: "Jose"
last: "Pereia"
position: "Governor"

name: “Alice”
last: „Smith“
role: „Advisor“

name: “Some Media Ltd”
value: “$70M”

name: "Maria"
last: "Osara"
Context is King

name: “John”
last: “Miller”
role: “Negotiator”
since: Jan 10, 2011

name: “Jose”
last: “Pereia”
position: “Governor”

name: “Alice”
last: “Smith”
role: “Advisor”

name: “Some Media Ltd”
value: “$70M”

name: "Maria"
last: "Osara"
In today's news

The New York Times  |  By SUSANNE CRAIG

Trump's Empire: A Maze of Debts and Opaque Ties

A review of the Republican presidential nominee’s holdings uncovered more debt than what is apparent on his federal election filing, and partnerships with even more liabilities.

Photo credits: Andrew Toth/Getty, Sylvain Gaboury/PMC (©Patrick MCMullan), Danny Moloshok/Invision/AP, Michael Vadon/Flickr (CC BY-SA)

https://www.publicintegrity.org/2016/10/07/20305/robert-mercer-connections-stephen-k-bannon
The world is a graph – everything is connected

- people, places, events
- companies, markets
- countries, history, politics
- sciences, art, teaching
- technology, networks, machines, applications, users
- software, code, dependencies, architecture, deployments
- criminals, fraudsters and their behavior
Why Graphs? The World is a Graph
We need to store and query our meta-data!

Real, inferred and integrated
Property Graph Model

Nodes
• The entities in the graph
• Can have name-value properties
• Can be labeled

Relationships
• Relate nodes by type and direction
• Can have name-value properties
Neo4j

An *open-source graph database*

- Manage and store your connected data as a graph
- Query relationships easily and quickly
- Evolve model and applications to support new requirements and insights
- Built to solve *relational pains*
What is Neo4j?

An open-source, native graph database

- Built for Connected Data
- Easy to use
- Optional Schema
- Highly Scalable Performance
- Transactional ACID-Database
- Clustering for HA and Scale
- Built in Security
Value from Data Relationships
Some Use Cases

Internal Applications
Master Data Management
Network and IT Operations
Fraud Detection

Customer-Facing Applications
Real-Time Recommendations
Graph-Based Search
Identity and Access Management

http://neo4j.com/use-cases
Neo4j: All about Patterns

(:Person { name:"Dan"}) -[:KNOWS]-(:Person {name:"Ann"})
MATCH (:Person { name:"Dan"}) -[:KNOWS]-> (who:Person) RETURN who
Cypher: Clauses

CREATE pattern
MERGE pattern
ADD
DELETE
Cypher: Clauses

MATCH pattern
WHERE predicate
ORDER BY expression
SKIP ... LIMIT ...
RETURN expression AS alias ...
WITH expression AS alias, ...
UNWIND list AS item
LOAD CSV FROM „url“ AS row
Getting Data into Neo4j

• Bulk Load from CSV Files

• Update Graph from
  • Web APIs (JSON, XML)
  • Other Databases
  • CSV Files
  • User Activity (Logs, Callbacks)
Cypher-Based “LOAD CSV”
• Transactional (ACID) writes
• Initial and incremental loads of up to 10 million nodes and relationships

```cypher
LOAD CSV WITH HEADERS FROM "url" AS row
MERGE (:Person {name:row.name, age:toInt(row.age)});
```
Getting Data into Neo4j

Load JSON with Cypher

- Load JSON via procedure
- Deconstruct the document
- Into a non-duplicated graph model

CALL apoc.load.json("url") yield value as doc
UNWIND doc.items as item
MERGE (:Contract {title:item.title,
amount:toFloat(item.amount)});
Getting Data into Neo4j

**CSV Bulk Loader** *neo4j-import*

- For initial database population
- For loads with 10B+ records
- Up to 1M records per second

```
bin/neo4j-import --into people.db
--nodes:Person people.csv
--nodes:Company companies.csv
--relationship:STAKEHOLDER stakeholders.csv
```
===>  /Users/mh/Downloads/panama/import/Addresses.csv

address,icij_id,valid_until,country_codes,countries,node_id:ID,sourceID

27 ROSEWOOD DRIVE #16-19 SINGAPORE 737920,6991059DFFF057DF31089BF31CC4A0E6,The Panama Papers data is current through 2015,SGP,Singapore,14000001,Panama Papers

===>  /Users/mh/Downloads/panama/import/Entities.csv

ame,name,original_name,former_name,jurisdiction,jurisdiction_description,company_type,address,internal_id,incorporation_date,inactivation_date,struck_off_date,dorm_date,status,service_provider,ibcRUC,country_codes,countries,note,valid_until,node_id:ID,sourceID

"TIANSHENG INDUSTRY AND TRADING CO., LTD.","TIANSHENG INDUSTRY AND TRADING CO., LTD.",SAM,Samoa,,ORION HOUSE SERVICES (HK) LIMITED ROOM 1401; 14/F.; WORLD COMMERCE CENTRE; HARBOUR CITY; 7-11 CANTON ROAD; TSIM SHA TSUI; KOWLOON; HONG KONG,1001256,23-MAR-2006,18-FEB-2013,15-FEB-2013,,Defaulted,Mossack Fonseca,25221,HKG,Hong Kong,,The Panama Papers data is current through 2015,10000001,Panama Papers

===>  /Users/mh/Downloads/panama/import/Intermediaries.csv

name,internal_id,address,valid_until,country_codes,countries,status,node_id:ID,sourceID

"MICHAEL PAPAGEORGE, MR.",10001,MICHAEL PAPAGEORGE; MR. 106 NICHOLSON STREET BROOKLYN PRETORIA 0002; GAUTENG (PWV) SOUTH AFRICA,The Panama Papers data is current through 2015,ZAF,South Africa,ACTIVE,11000001,Panama Papers

===>  /Users/mh/Downloads/panama/import/Officers.csv

name,icij_id,valid_until,country_codes,countries,node_id:ID,sourceID

KIM SOO IN,E72326DEA50F1A9C2876E112AAEB42BC,The Panama Papers data is current through 2015,KOR,"Korea, Republic of",12000001,Panama Papers

===>  /Users/mh/Downloads/panama/import/all_edges.csv

node_id:START_ID,rel_type:TYPE,node_id:END_ID

11000001,intermediary_of,10208879
IMPORT DONE in 20s 747ms. Imported:
  839434 nodes
  1253582 relationships
  8211010 properties

<table>
<thead>
<tr>
<th>labels(n)</th>
<th>count(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[&quot;Officer&quot;]</td>
<td>344455</td>
</tr>
<tr>
<td>[&quot;Entity&quot;]</td>
<td>319150</td>
</tr>
<tr>
<td>[&quot;Address&quot;]</td>
<td>151054</td>
</tr>
<tr>
<td>[&quot;Intermediary&quot;]</td>
<td>23636</td>
</tr>
</tbody>
</table>
The Steps Involved in the Document Analysis

1. **Acquire** documents

2. **Classify** documents
   - Scan / OCR
   - Extract document metadata

3. Whiteboard **domain** and **questions**, determine
   - **entities** and their **relationships**
   - potential entity and relationship **properties**
   - **sources** for those entities and their properties
The Steps Involved in the Document Analysis

4. Develop analyzers, rules, parsers and named entity recognition
5. Parse and store metadata, document and entity relationships
   • Parse by author, named entities, dates, sources and classifications
6. Infer entity relationships
7. Compute similarities, transitive cover and triangles
8. Analyze data using graph queries and visualizations
Journalists say: „It’s like Magic“
Journalists say: „It’s like Magic“

- Find interesting spots with full-text and fuzzy search
- See neighbourhoods of suspects and interesting facts
- Find connections and shortest paths between seemingly disconnected information
- Jointly add new knowledge as relationships
- Stories emerge from the collaboration
- Add more information from other sources
We need a Data Model

Either based on our use cases & questions
On the entities present in our meta-data and data.

Meta Data Entities
• Document, Email, Contract, DB-Record
• Meta: Author, Date, Source, Keywords
• Conversation: Sender, Receiver, Topic
• Money Flows

Actual Entities
• Person
• Representative (Officer)
• Address
• Client
• Company
• Account
Data Model – Relationships

Meta-Data
- sent, received, cc’ed
- mentioned, topic-of
- created, signed
- attached
- roles
- family relationships

Activities
- open account
- manage
- has shares
- registered address
- money flow
The Basic ICIJ Data Model
The Basic ICIJ Data Model
The Real ICIJ Data Model
The ICIJ Data Model

- Simplistic Datamodel with 4 Entities and 5 Relationships
- We only know the published model
- Missing
  - Documents, Metadata
  - Family Relationships
  - Connections to Public Record Databases
- Contains Duplicates
- Relationship information stored on entities
- Could use richer labeling
Data initially exposed as interactive Visualization

- Public figures and leaders
- Different shell companies & involvements

The family of Azerbaijan President Ilham Aliyev leads a charmed, glamorous life, thanks in part to financial interests in almost every sector of the economy. His wife, Mehriban, comes from the privileged and powerful Pashayev family that...
Offshore Leaks Database

Find out who’s behind almost 320,000 offshore companies and trusts from the Panama Papers and the Offshore Leaks investigations

**Search by country**
People, companies and addresses connected to offshore entities

**Search by jurisdiction**
Offshore companies and trusts

Enter any name here

All countries

SEARCH

To search, enter a term in the box. Leave it blank and select a country or jurisdiction to see a full list of names.
People Love the #PanamaPapers Database!

Since May 9, 2016:

- 7.4 million sessions
- 44.2 million page views
- Visitors coming from (top 5 countries): US, Japan, Spain, Canada, UK
We have also made a **distribution** of Neo4j available with the data in it. This will allow you to **query** the database to fully explore from your computer the connections between people and companies. The package also includes a **guide** that explains how to use Neo4j.
Let’s go, get it!
Next steps for the ICIJ and all of us

- Data integration with other sources
- Entity extraction
- Email pattern analysis
- Content & Data mining
- Machine learning
- Alerts with real time news / social media
- Investigative recommendations
- Active search for new sources ...
Visual Graph Search

Make data accessible for non-developers
Put it on a Map
(did it this morning in 10’)

- Ingredients
  - Neo4j
  - Cypher
  - Bolt Javascript Driver
  - APOC procedures (geocode)
  - MapBox

https://github.com/jexp/panama-map
Steps

Geocode
MATCH (a:Address) WHERE a.country_codes = "DEU"
CALL apoc.spatial.geocodeOnce(a.address) YIELD location
SET a += location

Query + Render
MATCH (a:Address)<--(officer:Officer)-[role]->(entity:Entity)
WHERE a.country_codes = "DEU" AND exists(a.longitude) AND
distance(pos, point(a)) < ({distanceInKm} * 1000)
RETURN a.latitude, a.longitude,
{officer : officer.name, entity: entity.name, role: type(role),
 address : a.address, country: entity.country_codes} AS data
Linkurious.js

- JS library based on sigma.js
- Integrates with Neo4j using Cypher

https://github.com/Linkurious/linkurious.js/
Popoto.js

• JS library based on d3.js
• Uses Graph Metadata to offer visual search
• Categories to filter Instances
• Component based extensions
• Zero Config with Web Extension

http://www.popotojs.com/
Visual Search Bar

- Based on visualsearch.js
- Uses graph metadata for parametrization
- Limit suggestions by selected items

» You searched for: **Actor.name**: "Zach Grenier". (1 node)

maxdemarzi.com/2013/07/03/the-last-mile/
Facebook Graph Search

• Natural Language to Cypher
• Ruby TreeTop Gem for NLP
• Convert phrases to Cypher Fragments

Cypher Query:
START me = node({me}), thing = node:things({thing})
MATCH me <-[friends] > people, people <-[likes] > thing
RETURN DISTINCT people, people.uid, people.name, people.image_url
LIMIT 100
Parameters: {"me"=>"me", "thing"=>"name: Neo4j"}
Users Love Neo4j

Performance
"The Neo4j graph database gives us drastically improved performance and a simple language to query our connected data”
- Sebastian Verheugher, Telenor

Scale and Availability
"As the current market leader in graph databases, and with enterprise features for scalability and availability, Neo4j is the right choice to meet our demands.”
- Marcos Wada, Walmart

“We found Neo4j to be literally thousands of times faster than our prior MySQL solution, with queries that require 10 to 100 times less code. Today, Neo4j provides eBay with functionality that was previously impossible.”
- Volker Pacher
Senior Developer
Walmart
Graphs in Humanities

The Human World is connected
The Codex – Iian D. Neil

The Codex
An Atlas of History

TWITTER  BITBUCKET  LINKEDIN
An atlas of history

The Codex is a digital humanities project built in Neo4j and .NET that breaks history down into a series of semantically-tagged events.

It is an ‘atlas’ of history because it enables events to be viewed and contrasted in various ways, by subject, by person, and by place.

Relationships between historical figures and their semantic attributes are also recorded.
Goals of the Codex

- Refactoring of project to be more collaborative & XML standards-compliant
- Public API to enable querying of events by time, place, person, etc.
- Data entry/classification: expansion beyond Italian Renaissance to 19th century England & other information-rich periods
- Integration of diverse data sets
  - astronomical events, census data, climate
- Want to be involved? I’m looking for designers, history buffs, or anyone interested in the concept of using Neo4j to map history
Literature Studies – Vanderbilt University

- Analysing literary works
- Analysing lives of artists
- Student practices / thesis
- Visual appeal
- Workshops for students and staff
- Related projects
- Interview

http://gallery.library.vanderbilt.edu/exhibits/show/artists-books/neo4j
Historiana.eu

Your Portal to the Past is an on-line educational multimedia tool that offers students multi-perspective, cross-border and comparative historical sources to supplement their national history textbooks.

Historiana might be considered as a digital alternative to a European textbook, however the website does not attempt to present a comprehensive ...
Libraries and Museums

- D:Swarm
  - Open source library management system
  - Initial work on data import / modeling
  - SLUB + Avantgarde Labs
  - dswarm.org
- Natural History Museum London
  - Example for asset portal neo4j.com/blog/graphstarter-neo4j-rails-application
Neo4Art – Lorenzo Speranzoni

• Started Mapping Vincent Van Goghs journey
• dd
  • Places
  • Influence(r)s
  • Art
  • People

Traversing Art through its connections
Example Connections
Color Analysis
Life journey van Gogh
Medieval Research

• Graphdatenbanken für Historiker. Netzwerke in den Registern der Regesten Kaiser Friedrichs III. mit neo4j und Gephi
• Andreas Kuczera
• http://mittelalter.hypotheses.org/5995
• 3rd step after image digitization and fulltext document search
GraphCommons

- Graph based modeling for researchers and journalists
- Intuitive, collaborative Graph creation
- Embedding in Websites Encourage Sharing
- Interactive Graph Creation
- Graph based storytelling
- Tabular Data Entry
- Embedding in Websites
- Encourage Sharing

Montoro y compañía: un canje entre hermanos

Viejos conocidos
Summary
**Summary - Use the Right Database for the Job**

- **Discrete Data**
  - Minimally connected data

- **Connected Data**
  - Focused on Data Relationships

**Graph Databases are designed for data relationships**

- **Development Benefits**
  - Easy model maintenance
  - Easy query

- **Deployment Benefits**
  - Ultra high performance
  - Minimal resource usage
Real-Time Query Performance
Graph Versus Relational and Other NoSQL Databases

Relational and Other NoSQL Databases

Neo4j is 1000x faster “Minutes to milliseconds”

Neo4j

Response Time

Connectedness and Size of Data Set

0 to 2 hops
0 to 3 degrees
Thousands of connections

Tens to hundreds of hops
Thousands of degrees
Billions of connections

0 to 3 degrees
Thousands of connections

Billions of connections
Users Love Neo4j – Will you too?

John Resig
@jeresig

Really digging @neo4j. What use to be a bunch of complicated analysis scripts are now a handful of simple Cypher queries.

Marc Kuo
@k uomarc

loving @neo4j Browser -- what a beauty! Any DB should come bundled with such a slick interface #outofthebox

Sourabh Jain
@ja insourabh2

Just got my hands on @neo4j and it simply rocks!!!!! Amazingly easy to install, understand and code... Kudos to the Team..

Guillermo Szeliga
@g szeliga

I can't believe that @neo4j is actually real. Seems like a dream come true
What will YOU connect?

• Twitter & Facebook?
• Your family tree?
• Books, Songs, Movies, Series?
• Stories, Characters, Plots?
• Money, Accounts, Contracts?
• Products, Prices, Reviews, Tags?
• Software, Machines, Devices, Sensors?
What will YOU connect?

- Literature, Works, Authors?
- Language, Words, Phonemes?
- Research(ers), Publications, Citation?
- History, Events, Places?
- Catalogues, Items, Keywords, Annotations?
- Images, Documents, Letters, Books?
- Ontologies, Vocabulary, Translations?
- ...

neo4j
Get started with Neo4j today – Discover Value in Your Relationships

neo4j.com/developer
Want to learn more?

Free e-book download

For the best question
Let’s not stop here!
Support the work of organizations that help keep the world habitable
Neo4j Data Journalism Accelerator Program

Graph Databases are an amazing tool to discover connections in real-world data. Neo4j has a long history of being used by journalists, including the International Consortium of Investigative Journalists (ICIJ). The ICIJ used the Panama Papers and Swiss Leaks datasets to investigate and draw connections between people, shell companies and firms which helped create them. (Database Download, Why Graph Databases)

What is the Accelerator Program?
The Neo4j Developer Relations team wants to work with interested data journalists to help them understand how their data would be modeled as a graph database and what types of queries might be helpful to explore the graph. Drawing new connections between existing information and new sources allows new stories to be uncovered and told.

Short Application Process
We have a quick application process for us to understand your goals and your data. Our Engineers will work with selected applicants over the coming weeks to help you start being successful using Neo4j – either the community or enterprise editions. The only thing we ask is that you mention your usage of Neo4j in any resulting stories.

Get Started Now

More Insight

- Neo4j Blog
  - http://neo4j.com/blog/panama-papers/
  - http://neo4j.com/blog/analyzing-panama-papers-neo4j/
- ICIJ
  - https://panamapapers.icij.org/
  - https://panamapapers.icij.org/the_power_players/
  - https://panamapapers.icij.org/graphs/
- SZ
- Guardian
Source Material

taken from

- the ICIJ presentation
- the Reddit AMA
- online publications (SZ, Guardian, TNW et.al.)
- the ICIJ website
  - https://panamapapers.icij.org/
  - The Power Players
  - Key Numbers & Figures
Thanks! Questions?

Learn more: neo4j.com/developer
Me: @mesirii | @neo4j
Why should I care?

Because Relationships Matter
What is it with Relationships?

- World is full of connected people, events, things
- There is “Value in Relationships”!
- What about Data Relationships?
- How do you store your object model?
- How do you explain JOIN tables to your boss?
Neo4j – allows you to connect the dots

• Was built to efficiently
  • store,  
  • query and
  • manage highly connected data
• Transactional, ACID
• Real-time OLTP
• Open source
• Highly scalable already on few machines
Relational DBs Can’t Handle Data Relationships Well

- Cannot model or store data and relationships without complexity
- Performance degrades with number and levels of relationships, and database size
  - Query complexity grows with need for JOINs
- Adding new types of data and relationships requires schema redesign, increasing time to market

... making traditional databases inappropriate when data relationships are valuable in real-time

Slow development
Poor performance
Low scalability
Hard to maintain
NoSQL Databases *Don’t* Handle Data Relationships

- **No data structures** to model or store relationships
- **No query constructs** to support data relationships
- **Relating data requires “JOIN logic”** in the application
- Additionally, **no ACID support** for transactions

... making NoSQL databases [inappropriate](#) when data relationships are valuable in real-time
Largest Ecosystem of Graph Enthusiasts

- 1,000,000+ downloads
- 27,000+ education registrants
- 25,000+ Meetup members in 29 countries
- 100+ technology and service partners
- 170+ enterprise subscription customers including 50+ Global 2000 companies
Value from Data Relationships

Common Use Cases

Internal Applications
- Master Data Management
- Network and IT Operations
- Fraud Detection

Customer-Facing Applications
- Real-Time Recommendations
- Graph-Based Search
- Identity and Access Management
Graph Database Fundamentals
Find all reports and how many people they manage, up to 3 levels down.

**Cypher Query**

MATCH (boss)-[:MANAGES*0..3]->(sub), (sub)-[:MANAGES*1..3]->(report)
WHERE boss.name = "John Doe"
RETURN sub.name AS Subordinate, count(report) AS Total

**SQL Query**

```
SELECT T.directReports AS directReports, sum(T.count) AS count
FROM
(SELECT manager.pid AS reports, T.count AS count
FROM person_reports AS T
WHERE T.pid = manager.pid)
GROUP BY directReports
UNION
SELECT T.directReports AS directReports, count(T.reports) AS count
FROM person_reports AS T
WHERE T.pid = manager.directly_manages"
```