Neo Technology

The World's Leading Graph Database

NOSQL Roadshow

Dirk Möller
dirk.moeller@neotechnology.com
Cell: +49 151 40136308

Agenda

1. About Neo Technology
2. Graph Momentum & Relevance
3. Graph Databases & The Neo4j Graph DB
4. Neo Customers
5. Q&A

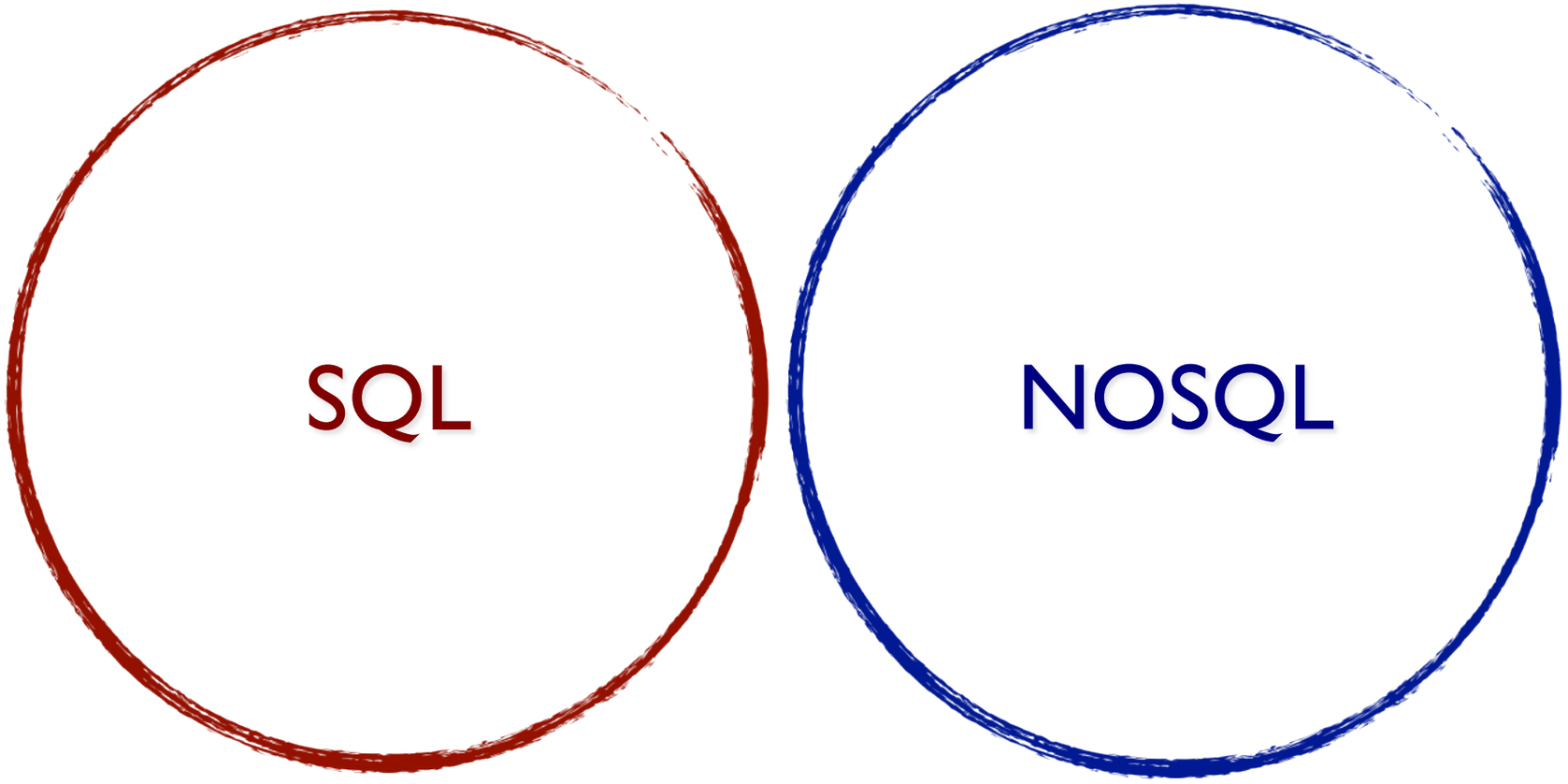
Our Mission:

Help the world to make sense of data

Our Vision for Database use in 2020:

NOSQL Dominated by Graph DBMSs

neo technology
value in relationships



Stores 50% of data
$45B market*

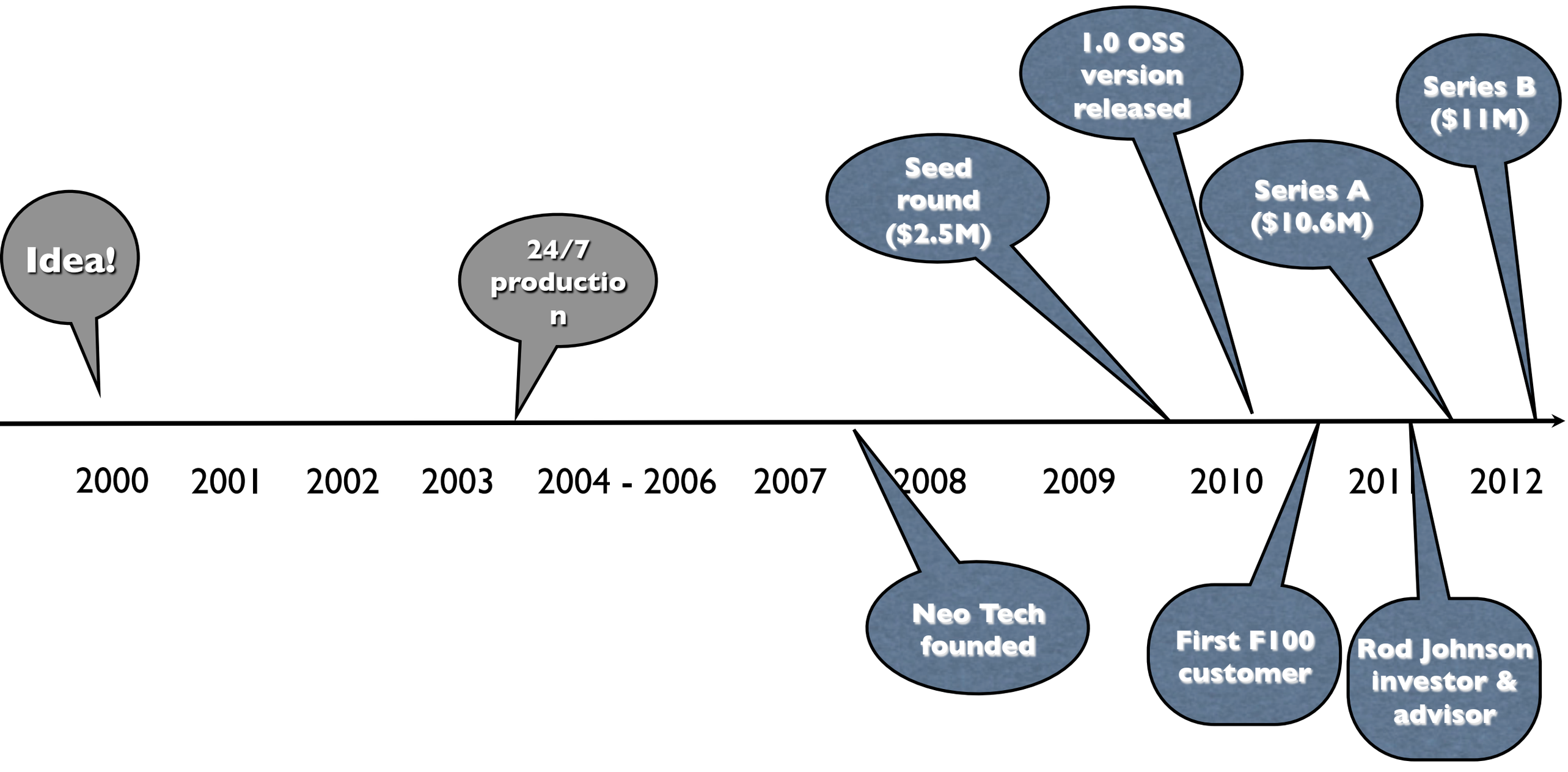
Stores 50% of data
? market

* Source: Gartner

Neo Technology
Fact Sheet

- **Numbers:**
  - **50 people** / Nine countries / Four continents
  - $24M raised (seed round Oct 2009 + $10M series A led by **Fidelity** Oct 2011 + $11M in Series B led by **Sunstone** Nov 2012)
  - **30,000+** downloads per month
  - **~ 100 commercial** customers
- **Team:**
  - Selected Leadership: Emil Eifrem (CEO) Rod Johnson (Chairman - SpringSource/VMW), Lars Nordwall (COO - ex SugarCRM, Pentaho), Philip Rathle (Sr Dir of Products - ex Accenture, Embarcadero), Jim Webber (Chief Scientist - ex ThoughtWorks)
- **Product:**
  - Development of Neo4j started in **2000** in Sweden
  - Put in 24/7 production in **2003**
  - Open sourced in **2007**
  - **Today the leading graph database**

Neo Technology story



neo technology
value in relationships

Graph Momentum & Relevance

The Early Adopters

neo technology
value in relationships

facebook

Google™

Facebook

Facebook

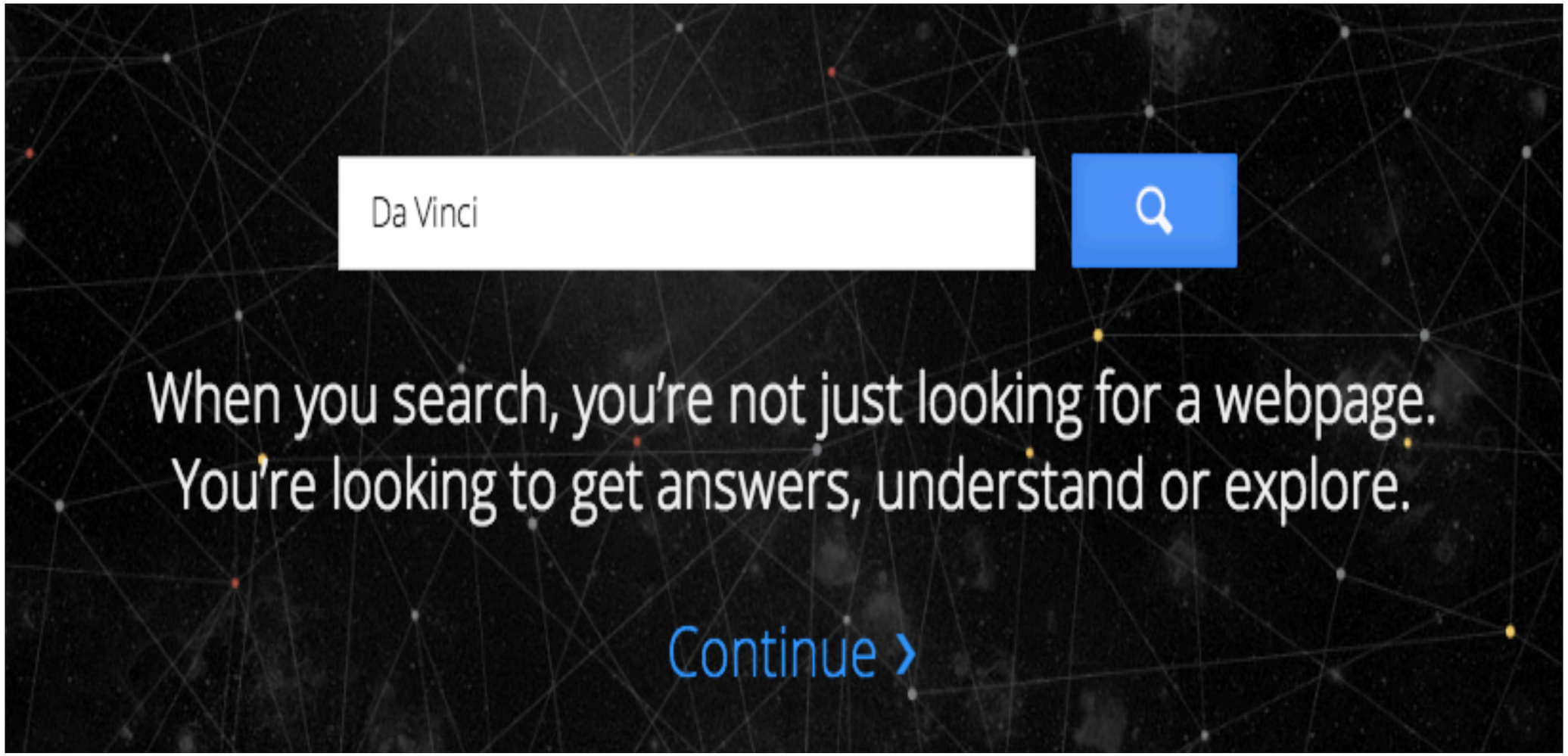
https://www.facebook.com/about/graphsearch

Google

Google

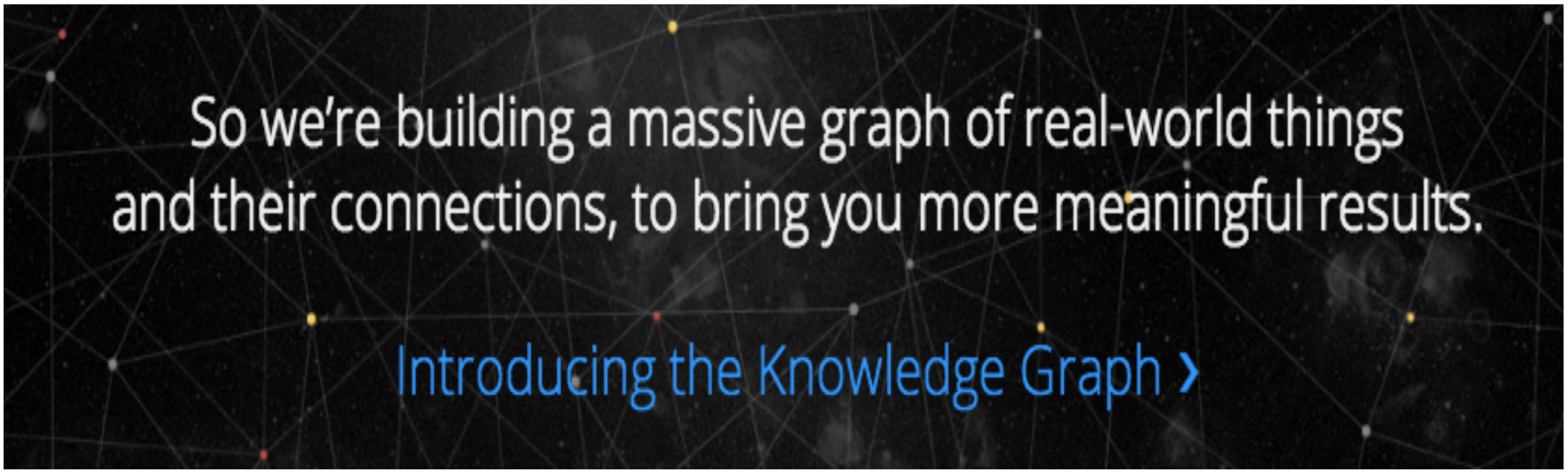
http://www.google.com/insidesearch/features/search/knowledge.html

Google



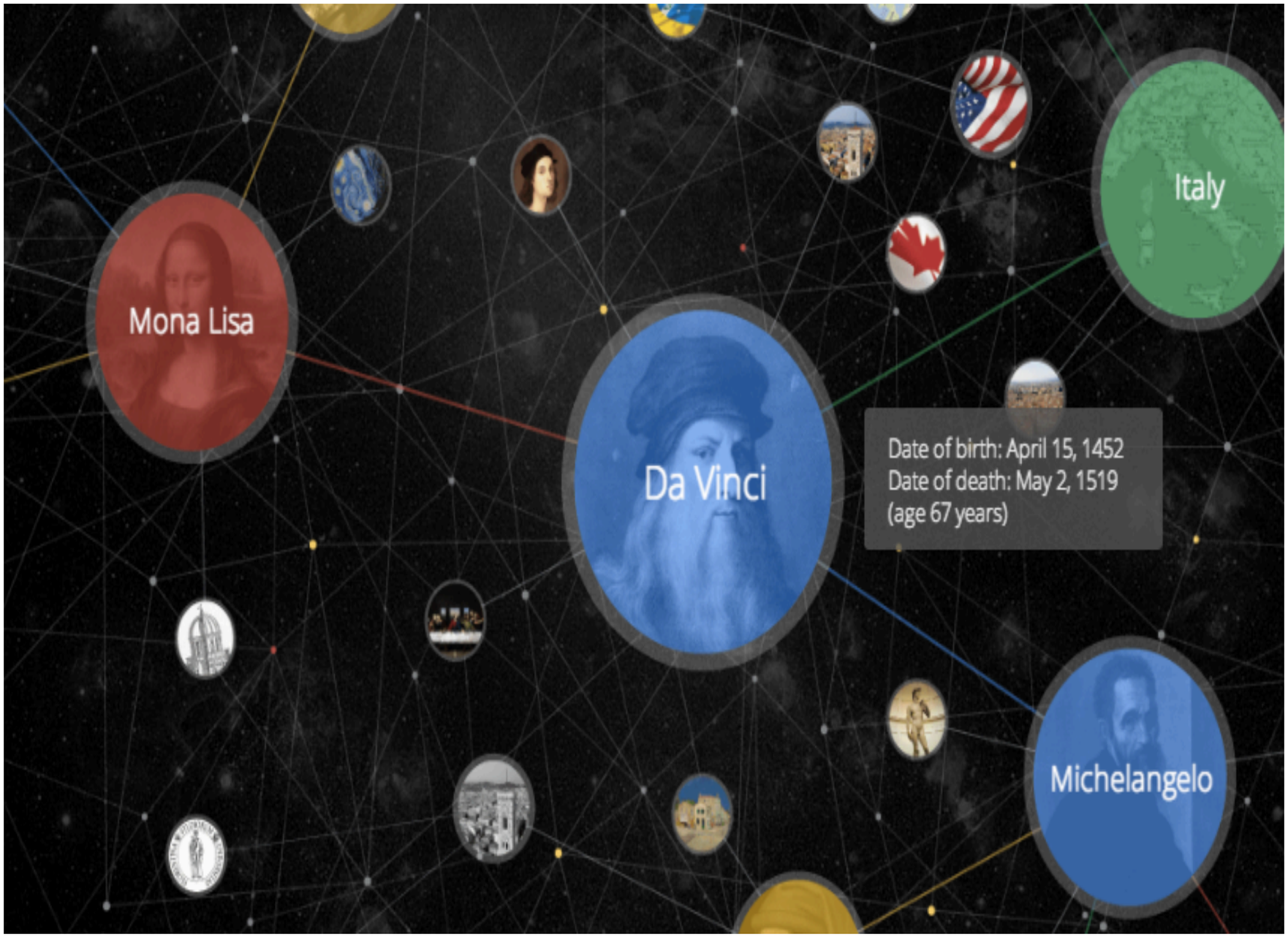
http://www.google.com/insidesearch/features/search/knowledge.html

Google



http://www.google.com/insidesearch/features/search/knowledge.html

Google



http://www.google.com/insidesearch/features/search/knowledge.html

Twitter

Twitter

Who to follow

Twitter accounts suggested for you based on who you follow and more.

Search using a person's full name or @username

Search Twitter

Lessig @lessig

law professor, reformer.

Followed by Manya Ellenberg, William Cook and Juan Pablo Buritica.

Follow

Jeremy Kemper @bitsweat

Building it: 37signals, Ruby on Rails

Follow

Sarah Mei @sarahmei

Ruby developer at Pivotal Labs, RailsBridge co-founder, dancer, gamer, friend, mom...etc.

Followed by Jeff Casimir, Mike Gehard and Vanessa Hurst.

Follow

Tom Preston-Werner @mojombo

GitHub Cofounder - Social Coding (http://github.com)

Followed by Derek P., Corey Haines and Paul Dix.

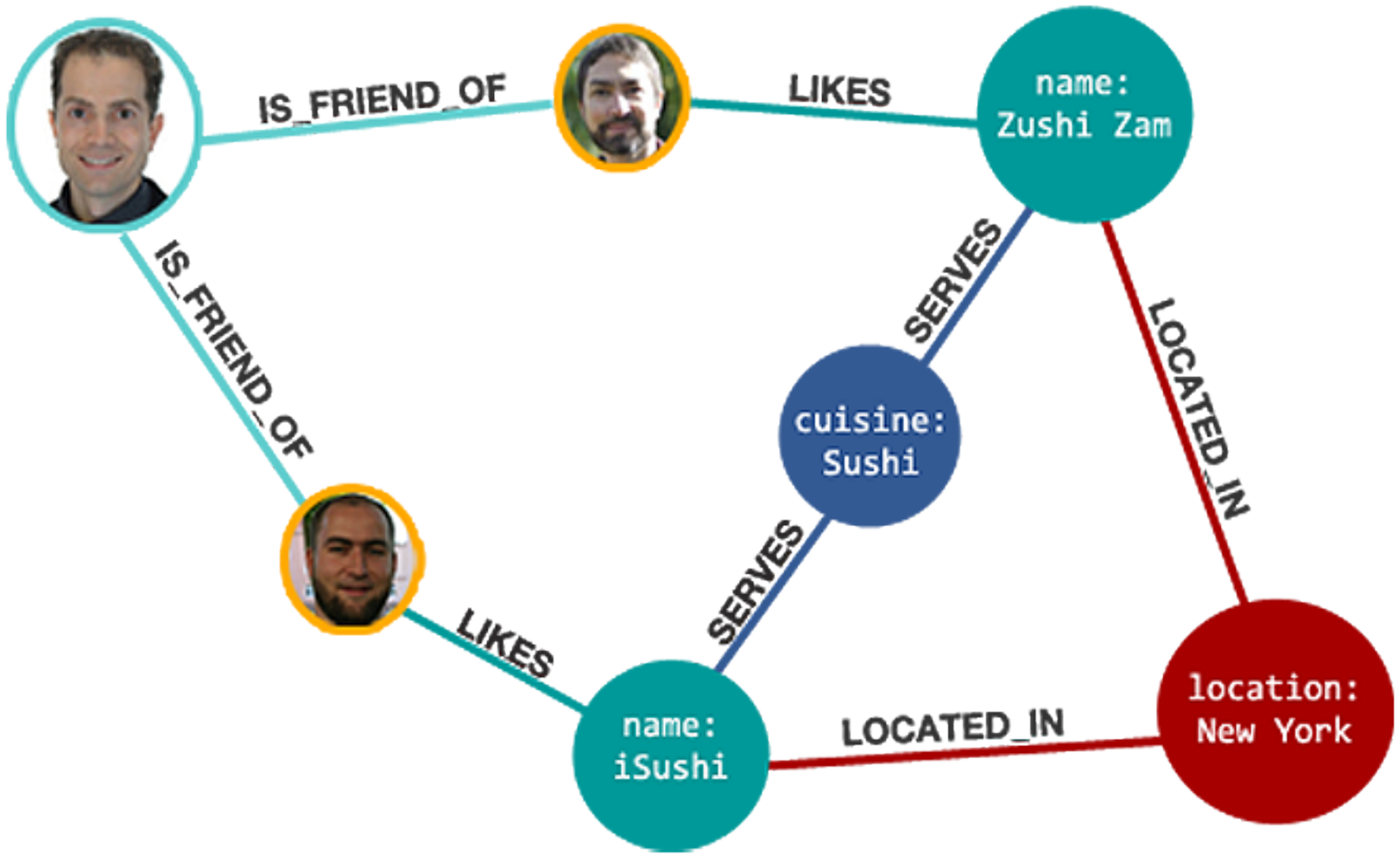
Follow

For the Facebook Graph Question:

What sushi restaurants in NYC do my friends like?

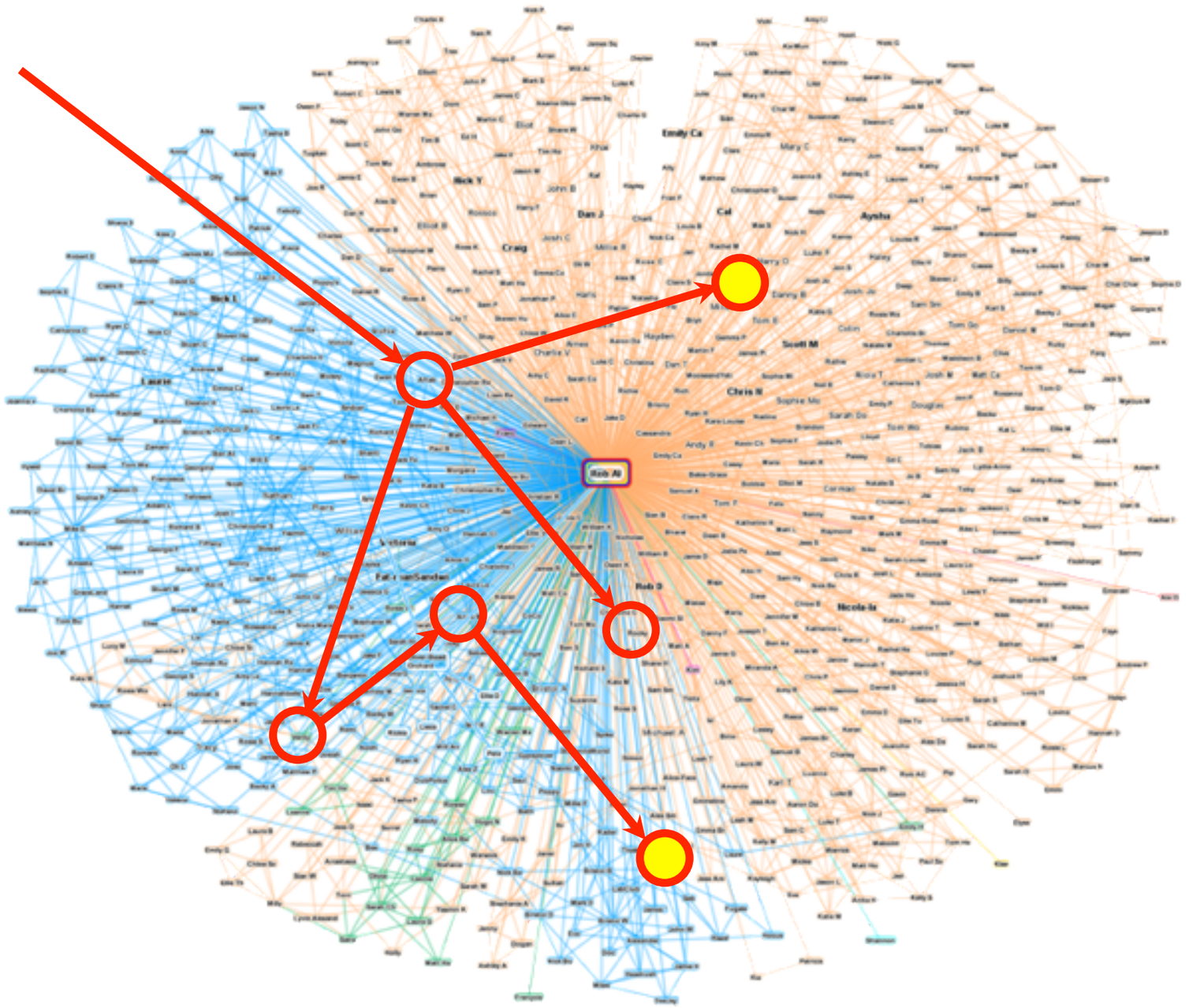
What the Graph Looks Like:

What sushi restaurants in NYC do my friends like?



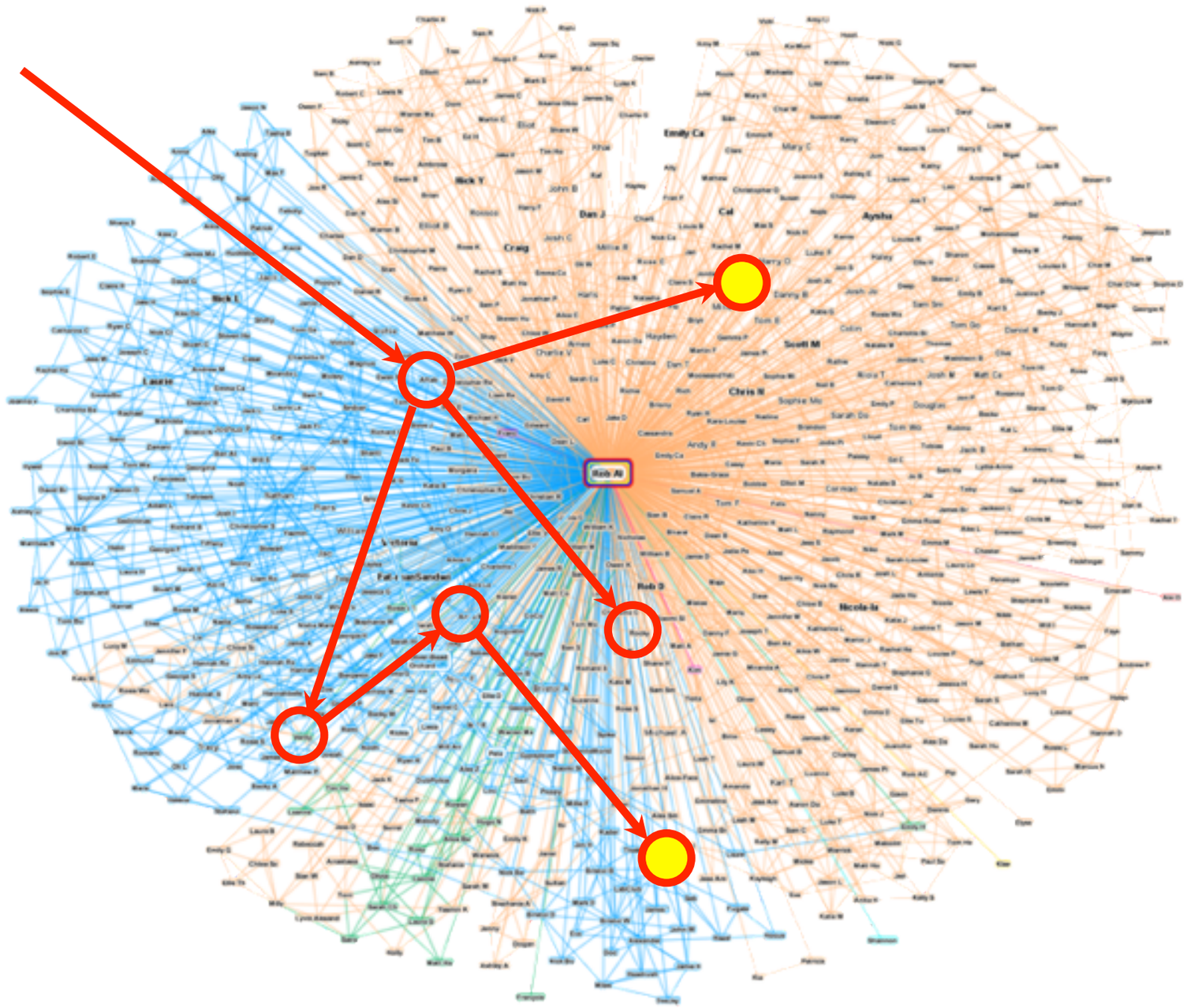
What the Search Looks Like:

What sushi restaurants in NYC do my friends like?



What Other Graph Searches Look Like

What drugs will bind to protein X and not interact with drug Y?



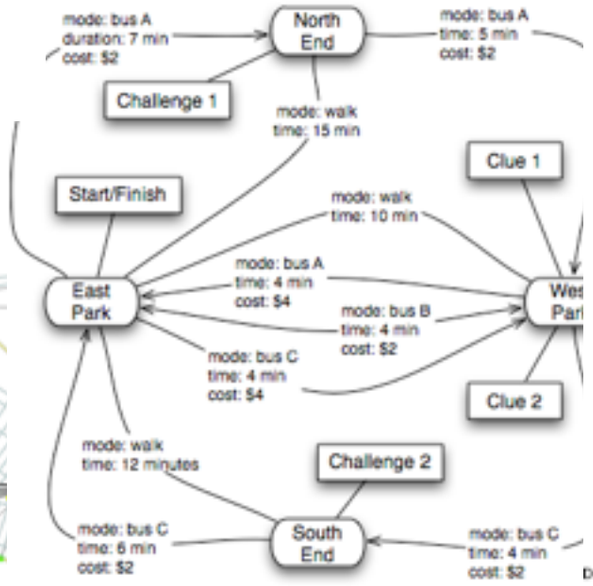
Beyond Buzz

Some Actual Neo4j Graphs

Content Management & Access Control

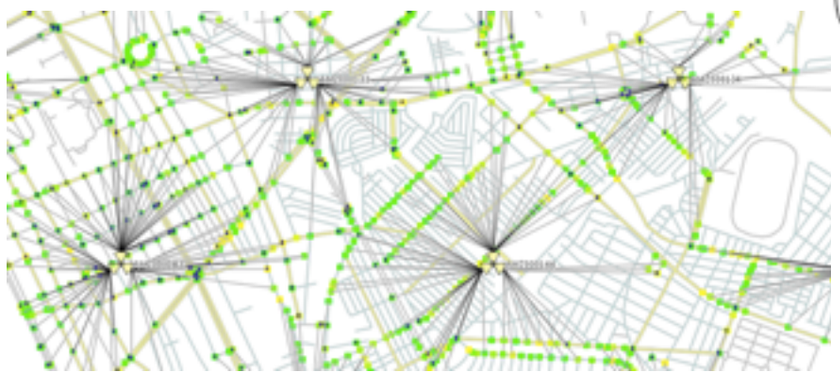
Geo Routing

(Public Transport)

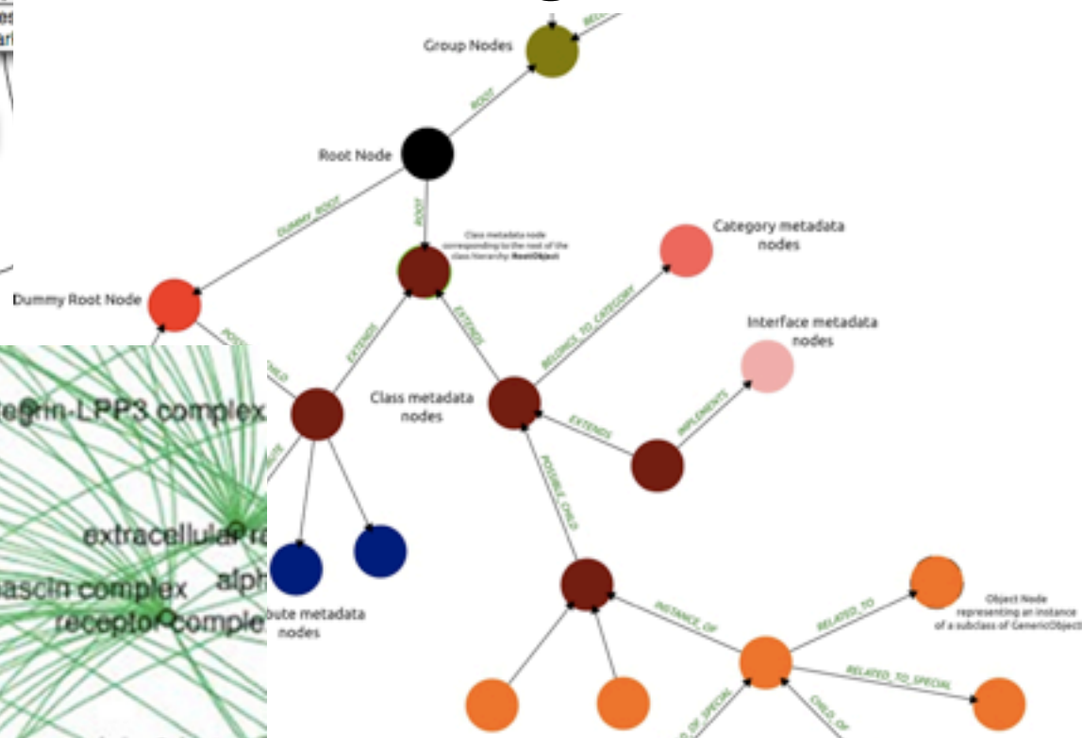


Insurance Risk Analysis

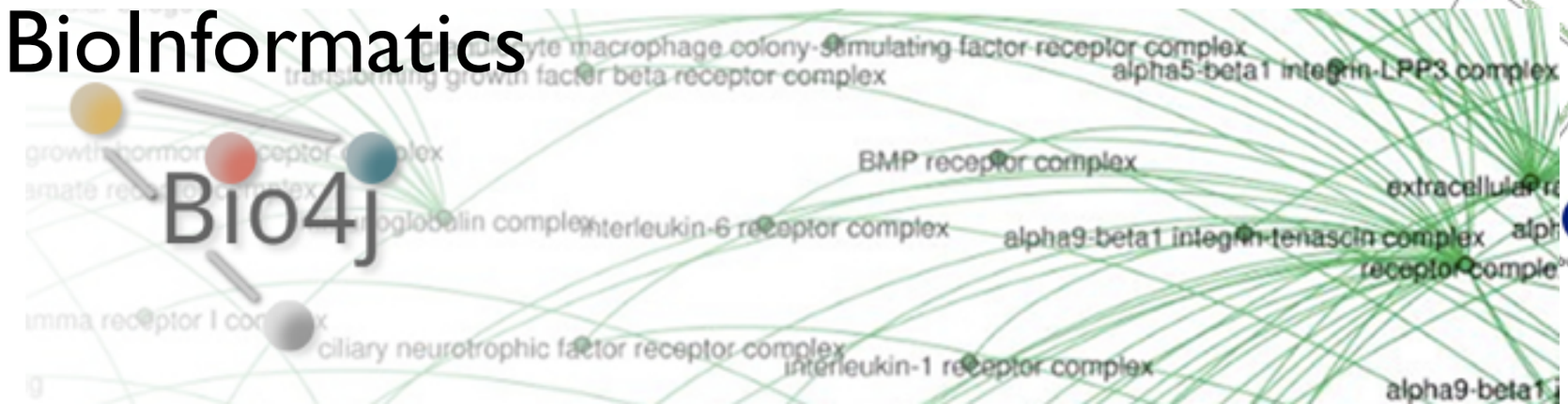
Network Cell Analysis



Network Asset Management



BioInformatics



Beyond Buzz

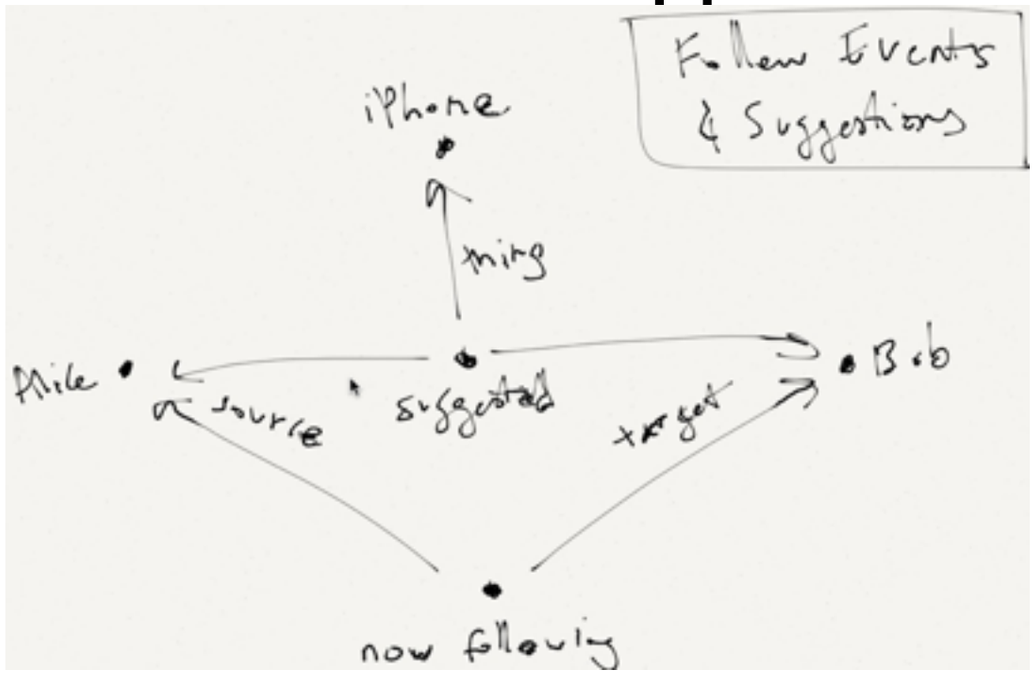
Some Actual Neo4j Graphs

Web Browsing

Portfolio Analytics

Gene Sequencing

Mobile Social Application



Early Adopter Segments

neo technology
graphs are everywhere

Core Industries & Use Cases:	Web / ISV	Finance & Insurance	Datacom / Telecom
Network Management			
MDM			
Social			
Geo			

Actual Commercial Demand

neo technology
graphs are everywhere

Core Industries & Use Cases:	Web / ISV	Finance & Insurance	Datacom / Telecom
Network Management	AXON ACTIVE Focusing on your decisions SERENA Juni sphere gen E	Bloomberg	SFR
MDM	Pitney Bowes		CISCO
Social	viadeo careerbuilder 53 bright glassdoor.com	ice Global markets in clear view	Deutsche Telekom maaii Let's connect
Geo			Justdial.com India's No.1 local search engine

Actual Commercial Demand

neo technology
graphs are everywhere

Core Industries & Use Cases:	Web / ISV	Finance & Insurance	Communi-cations	Logistics	Life Sciences	Media & Publishing	Education, Not-for-Profit	Government, Aerospace, Gaming, Other
Network Management	AXON ACTIVE, SERENA, Juni sphere, gen, Bloomberg		SFR					
MDM	Pitney Bowes		CISCO					
Social	viadeo, careerbuilder, bright, glassdoor.com, ice, Deutsche Telekom, maaii Let's connect							
Geo			Justdial.com India's No.1 local search engine					
Authorization & Access Control	identropy		telenor				teachscape	
Content Management	Adobe				SevenBridges genomics			
Recommend-ations	moviepilot, InfoJobs							
Fraud Detection, Other	iDMISSION empowering identity, DRW TRADING GROUP							

Accenture, shuti, ZEPHYR HEALTH INC, indiatimes, SQUIDOO, LIFECHURCH.TV, gamesys, zeebox, DOSB NEW MEDIA GMBH, Deutscher Olympischer SportBund, LAUREATE INTERNATIONAL UNIVERSITIES, <fuseworks/>, CHIP, research now, compete, LOCKHEED MARTIN, SA·NT DataLink, Netherlands Chamber of Commerce

Why Customers Choose Neo4j

1. Order-of-magnitude improvements in **query performance** for complex, connected data
2. Drastically accelerated **application development cycles**
3. **Maintainability** and **extensibility** of the data model
4. Relative **maturity** of the community and product

Benefits of a commercial relationship

- Commercial license vs open-source
- Service offerings, OEM
- Features
  - HA, DR, reporting
  - Turbo Cache
  - Online backup
  - Monitoring

Overview of Editions & Licenses

Edition	Features	Primary Licensing Model	...	(Special Cases)
Community	Full-featured graph database	Open Source (GPLv3)		Commercial (OEM)
Advanced	Full-featured graph database + Advanced Monitoring	Commercial License available through Neo Technology *Includes 5X10 Support		Open Source (AGPLv3) - avail for Open Source Projects
Enterprise	Full-featured graph database + Advanced Monitoring + Robust, fault-tolerant replicated cluster for demanding production use + Online backups + Turbo (aka GCR) cache	Commercial License available through Neo Technology *Includes 7X24 Support		Open Source (AGPLv3) - avail for Open Source Projects

Any edition can be run as Server or Embedded

neo technology
value in relationships

Use Cases & Industries

Use Cases

Network/Cloud Management: Management & tracking of physical or virtual computing resources outside of a data center, connected by wired or wireless networking elements.

Master Data Management: Storage and management of master data, esp. that which is hierarchical and variable in structure, such as organization and product.

Social: Systems that enable users to express social behavior, such as ratings, reviews, and discovery of content and other users.

Geo: Use of Neo4j to solve problems related to geographic locations, such as best and shortest path routing and nearest neighbor

Use Cases

Resource Authorization & Access Control: Storing information about resources and parties, as well as rules governing access to resources. Facilitating efficient execution of those rules.

Content Management: Processes and technologies that support the collection, managing, and publishing of information assets.

Recommendations: Methods for recommending relevant content to a user, based on known information about users and content, including other users' preferences and activities.

Data Center Management: Management & tracking of physical and virtual computing assets inside the data center, and how they are connected.

Fraud Detection: Refers to any one of various methods for detecting fraudulent behavior: either as it occurs, or after the fact

Neo Customers

Adobe

social networking, recommendations, access control

Description

Identifies which collections a user has access to
Finds third-party assets that are *like* a user's assets
Infers professional relations based on user-generated content

Background

Creative Cloud, announced 2011, is a cloud-based offering for professional users of Adobe's creative suite
Collaborative Cloud is the social element of the Creative Cloud, connecting professional users around the world

Benefits

Fit

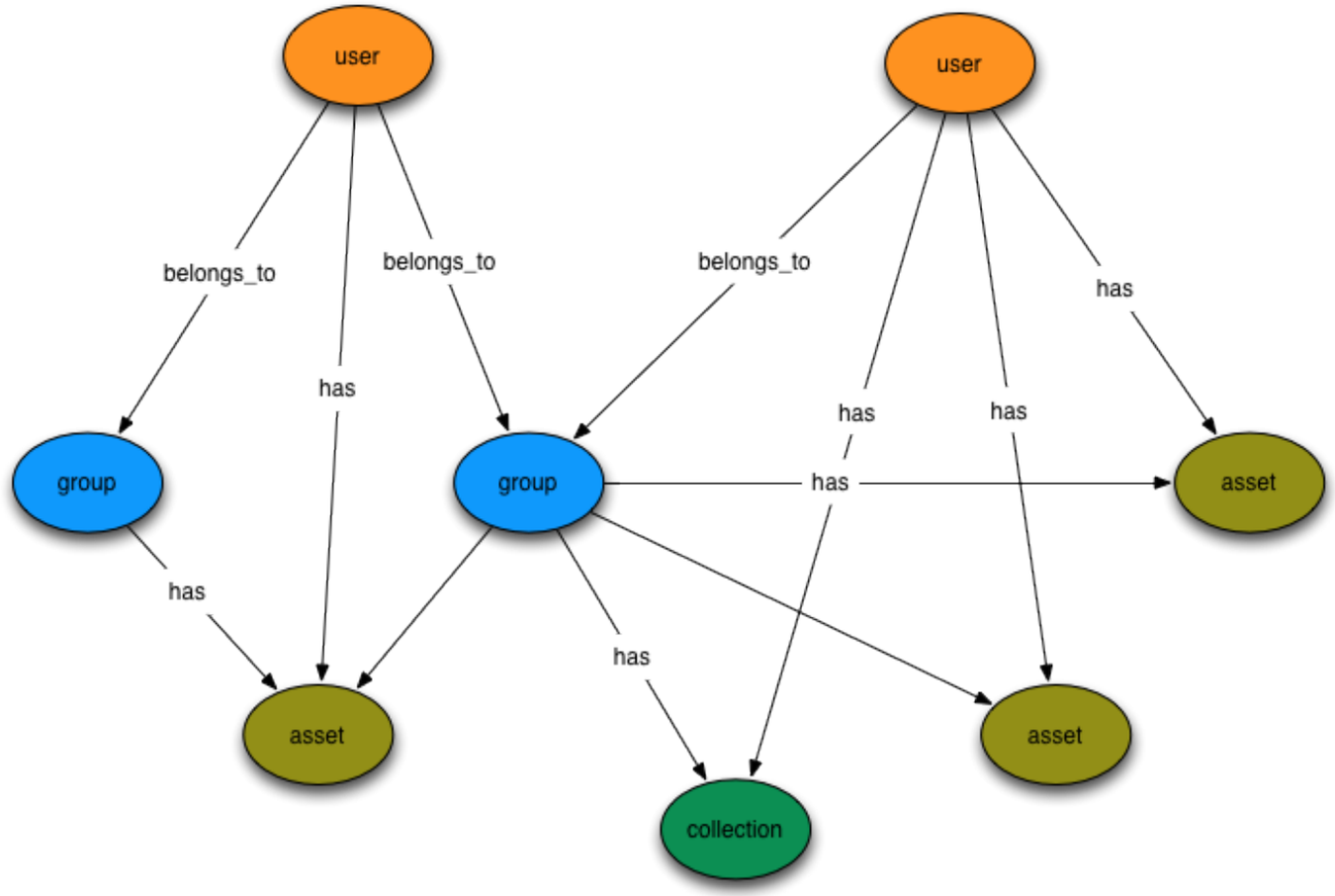
Graph model is a natural fit for social network
Collaborative user experience adds competitive advantage to Adobe offering

Flexibility

Datamodel can be easily evolved to support permissions and more sophisticated recommendation strategies

Perfomance

Sub-second results for large, densely-connected datasets



Cisco

master data management, sales compensation management, online customer support

Description

Real-time conflict detection in sales compensation management. Business-critical "P1" system. Neo4j allows Cisco to model complex algorithms, which still maintaining high performance over a large dataset.

Background

Neo4j replaces Oracle RAC, which was not performant enough for the use case.

Architecture

3-node Enterprise cluster with mirrored disaster recovery cluster
Dedicated hardware in own datacenter
Embedded in custom webapp

Sizing

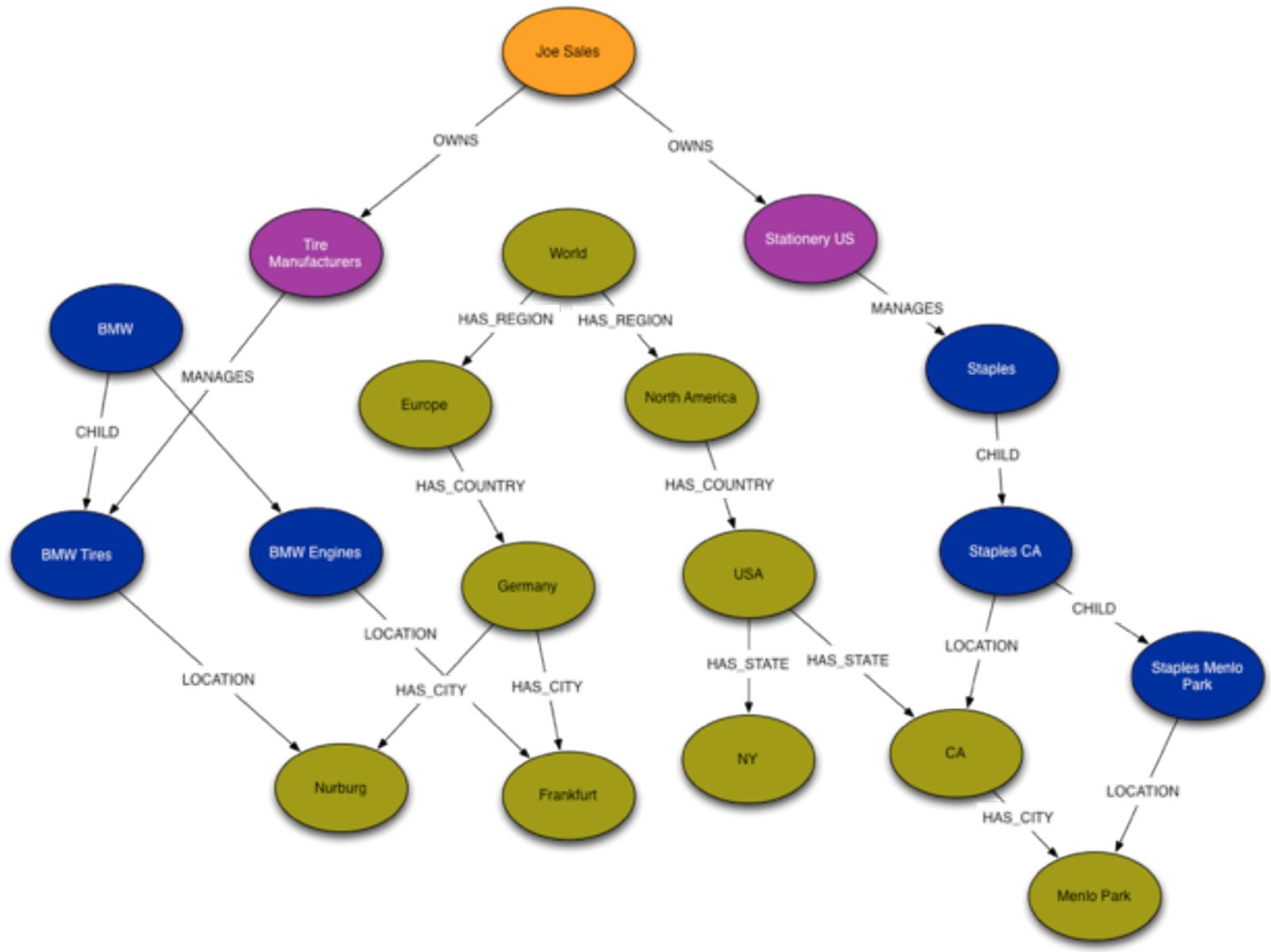
35 million nodes
50 million relationships
600 million properties

Benefits

Performance : "Minutes to Milliseconds"
Outperforms Oracle RAC, serving complex queries in real time

Flexibility
Allows for Cisco to model interconnected data and complex queries with ease

Robustness
With 9+ years of production experience, Neo4j brings a solid product.



Use case description - Cisco

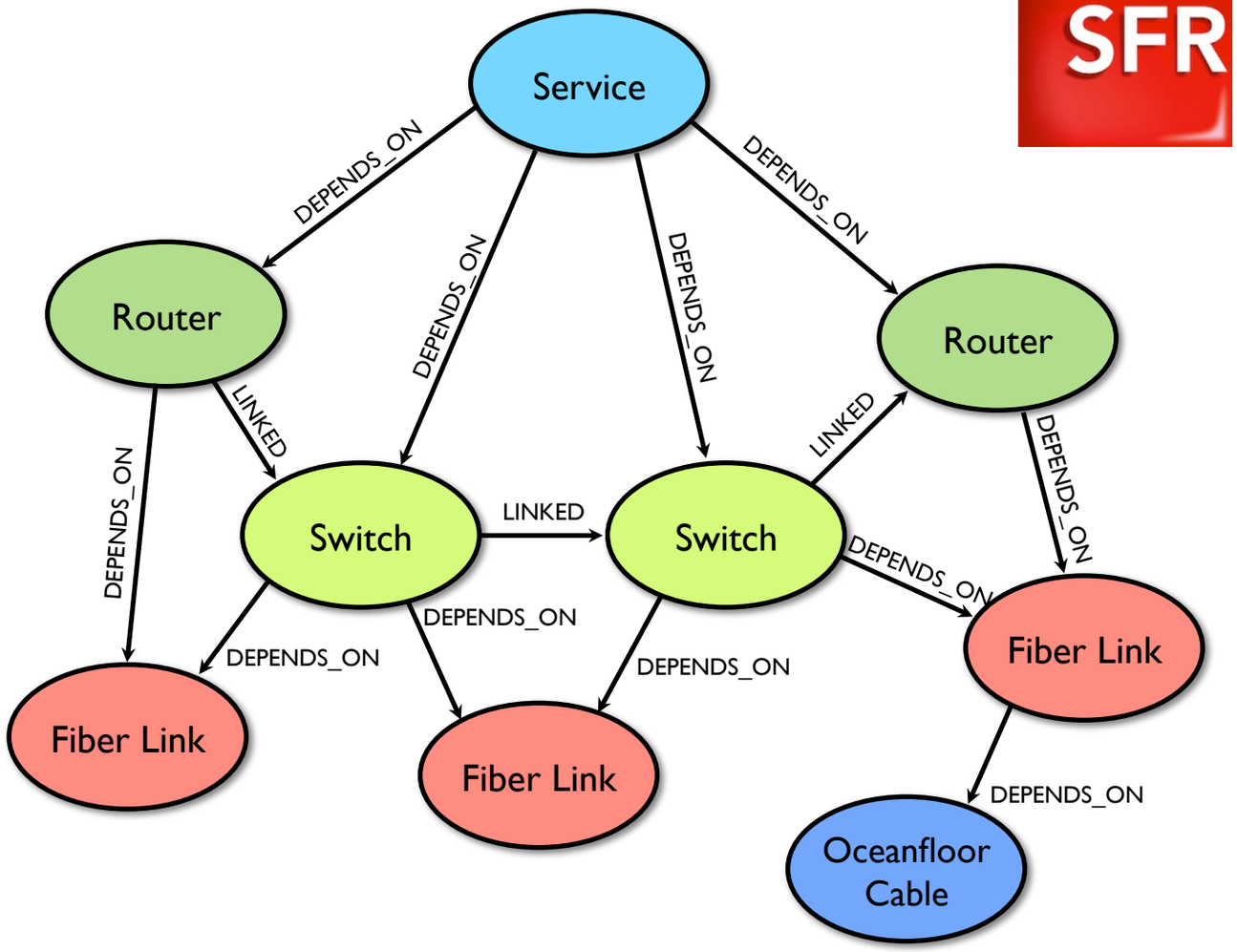
Response Time	Real time planning and resource management is possible when the application responds to 'what if scenario' in less than 5 seconds
Volume of Data	Sales account management identifies every Cisco customer with one sales representative for a given date. This association is query intensive operation spanning across more than 15 database tables scanning through 15-100 M records
Performance	Oracle database performance degrades when multiple tables are joined even after accounting for query tuning , partitioning and building multiple indexes
Concurrency	Oracle performance further degrades when multiple read / write operations are performed. Such behavior seriously impacts year round planning where users can do many changes during 'what if analysis' scenario in less than 5 seconds

SFR

Industry: *Datacom/Telecom*
Use case: *Network Management*

Background

- Second largest Telco in France
- Part of Vivendi Group, partnering with Vodaphone



Business problem

- Need for flexible network inventory management, aggregation, and troubleshooting
- Impact analysis of planned and unplanned network outages, so that affected services can be notified or receive increased redundancy
- Highly volatile network structure changing daily, with business requirements changing as well

Solution & Benefits

- Neo4j Enterprise with a highly available cluster
- Dynamic system allowing for new applications to tie into network structure data
- Near 1:1 mapping of real world to graph, greatly reducing modeling work
- High adaptability to changing business requirements

neo technology
graphs are everywhere

Telenor

organizational modeling, access control

Description

Identifies which customers, accounts and subscriptions an (administrative) user has access to

Identifies which users can modify a customer, account or subscription

Background

Current Sybase solution is too slow and too difficult and costly to replace

Big self-service customers require better user experience and improved performance

Telenor on track to exceed batch window - sought a move to real time

Benefits

Cost

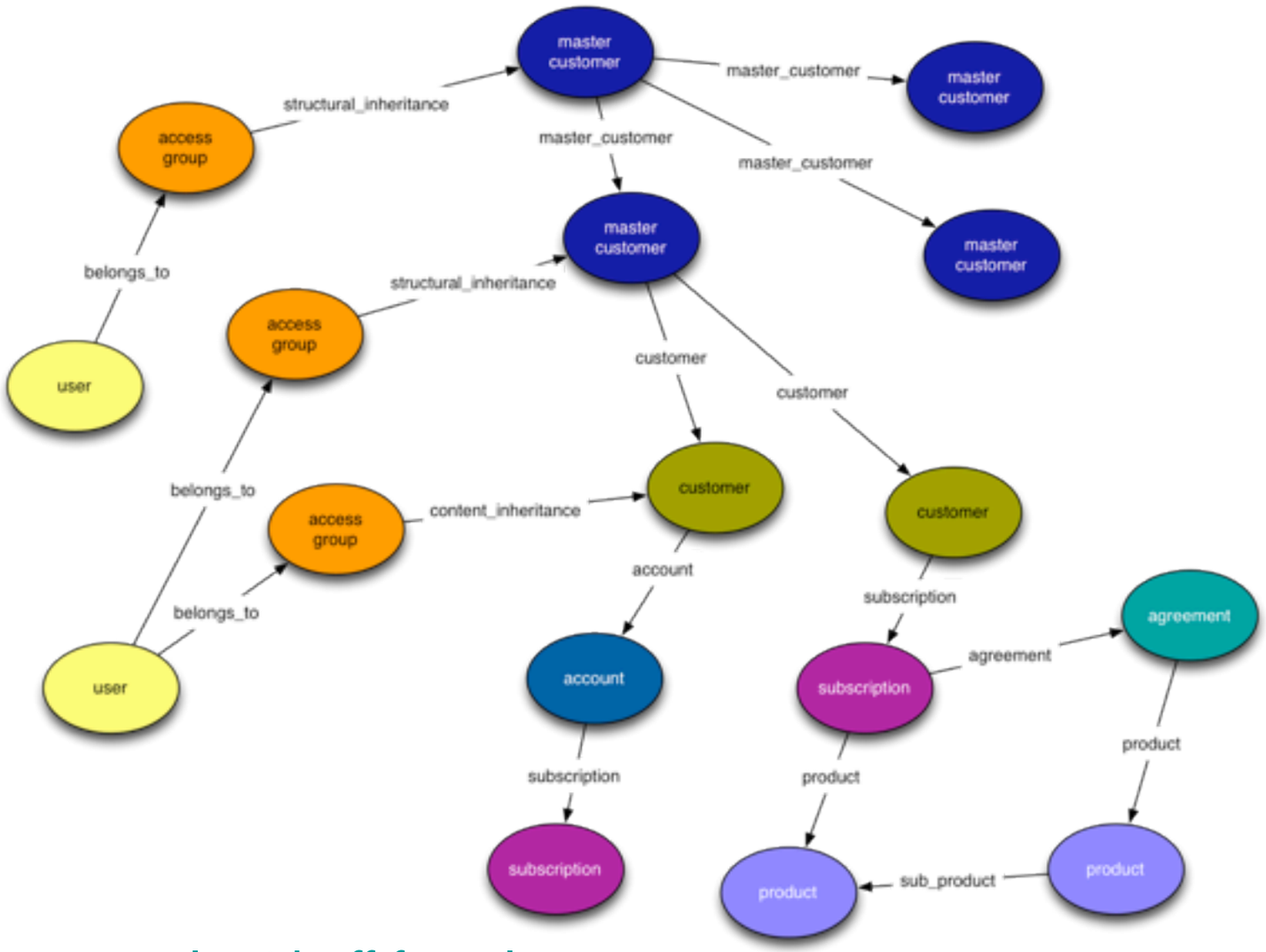
Low cost compared to alternatives

Flexibility

Extensible data model allows new information to be attached to the organisational structure

Perfomance

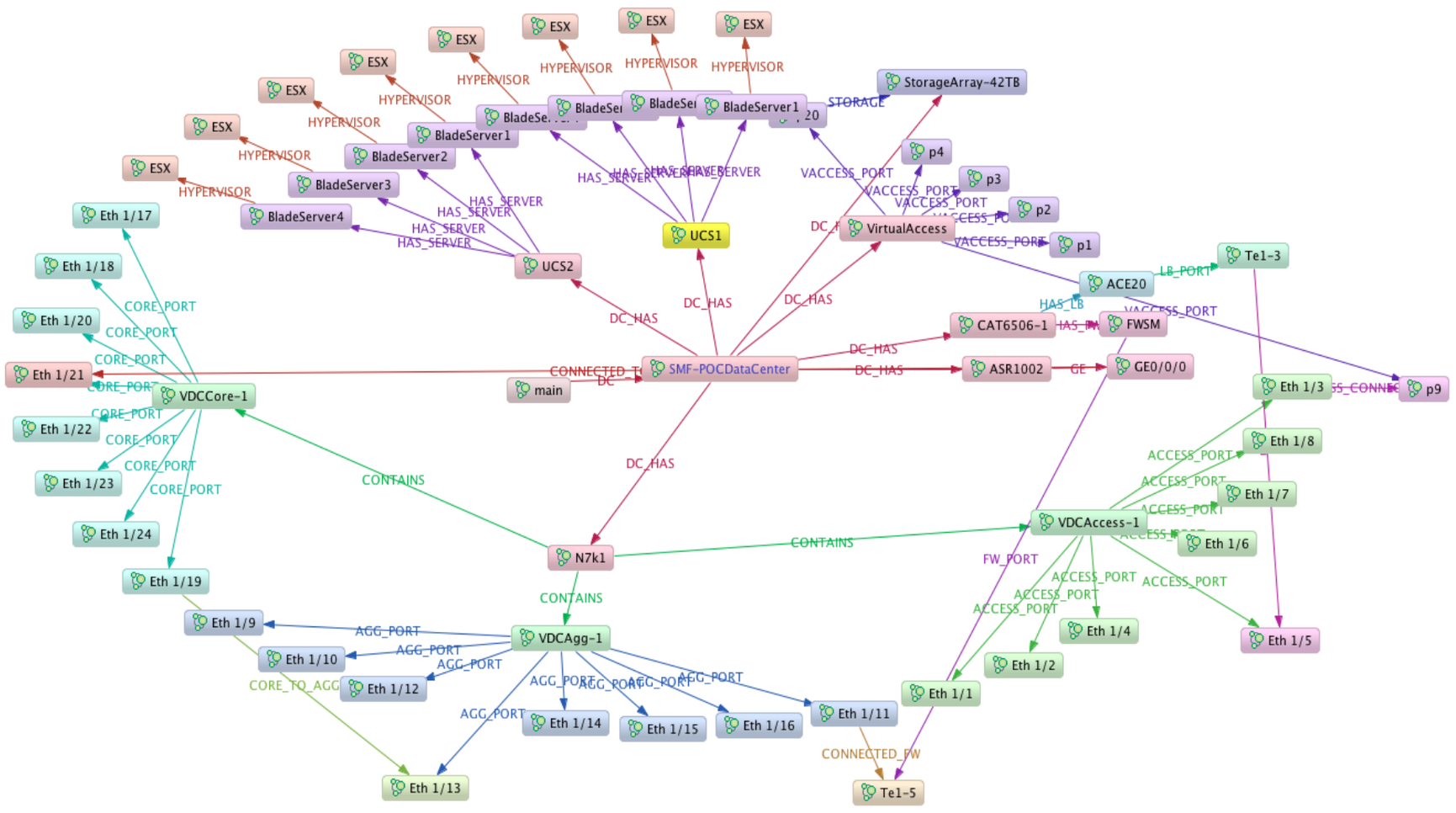
Sub-second results for complex queries across broad and deep graphs



http://www.slideshare.net/verheughe/how-nosql-paid-off-for-telenor

Network Management

- Modeling of physical and logical networks
  - impact analysis
  - configuration management
  - network inventory



Case study: Social networking
Viadeo

Problem definition

- Real-time recommendation imperative to attract new user and to maintain positive user retention
- Clustered MySQL solution not scalable and fast enough to support real-time requirements
- 8M nodes; 35M relationships
-Complex batch calculations with week old data
- Real-time traversals of the social graph not possible

Benefits & time frame

- Scalable solution with real-time end-user experience
- Low mainteance and reliable architecture with High-Availability (HA) and master failover
- 8 week implementation
- Three technical resources involved part-time

Solution

- Clustered Neo4j Enterprise architecture
- Solr for search; Tomcat; Memcache; etc.
- MySQL for general data storage; neo4j for social graph relationship characteristics

Company

- The French LinkedIn
- 35 million members
- 30,000 new members daily

Get involved in the community

neo technology
graphs are everywhere

Stack Overflow

Find answers or reach to fellow developers with questions.

Ask Neo4j questions »

http://stackoverflow.com/questions/tagged/neo4j

Neo4j Google Group

Share your experiences and expertise with fellow graphistas.

Join now »

http://groups.google.com/group/neo4j

GitHub Issues

Encountered an issue with Neo4j?
Submit it here.

https://github.com/neo4j/neo4j/issues

Meetups / User Groups

Neo4j meetups are worldwide. Make a connection or start a new group.

Join a Meetup »

http://neo4j.meetup.com/

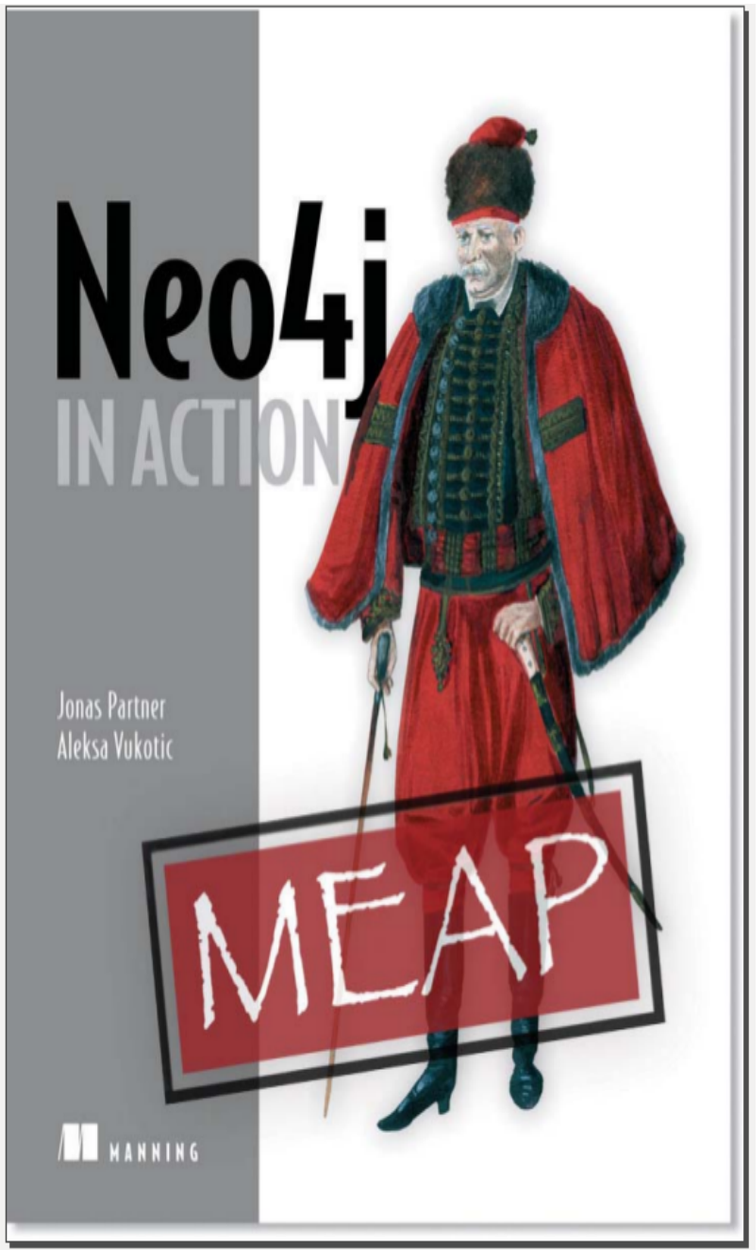
Graphistas World Map

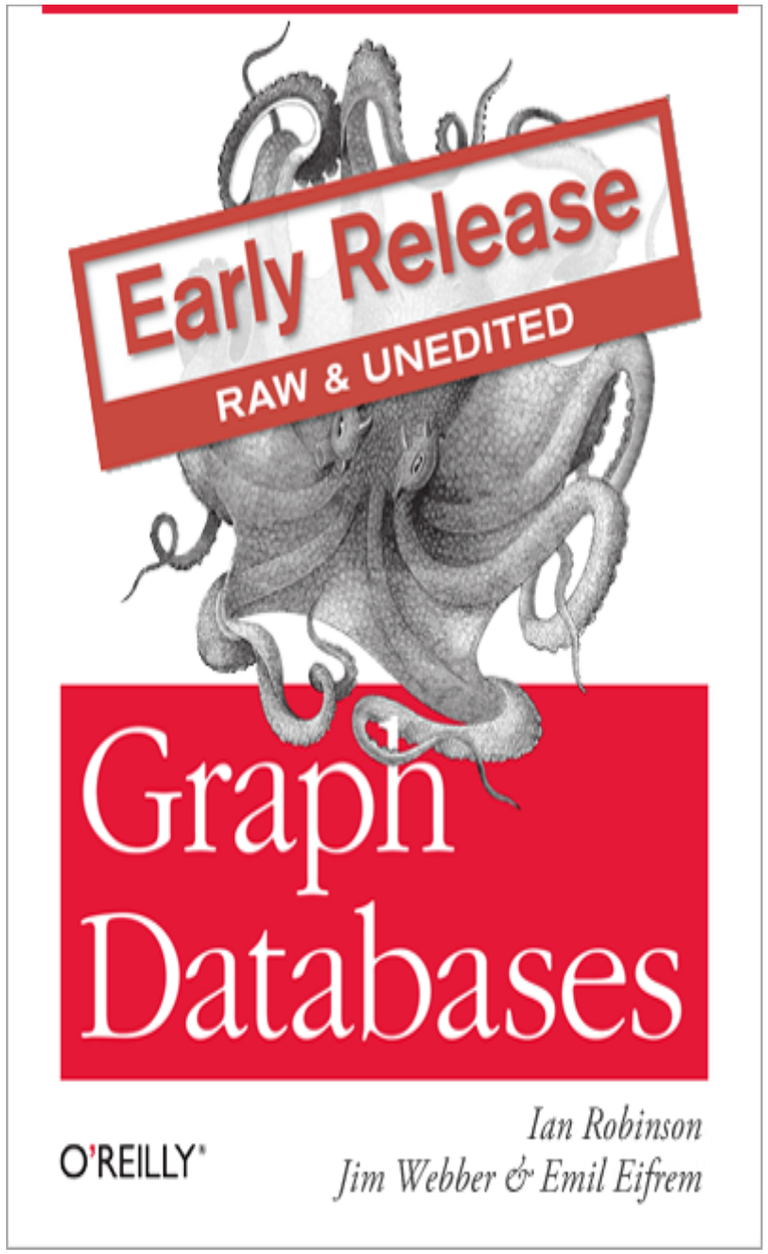
Add yourself to the graphistas world map and let it become a smaller place.

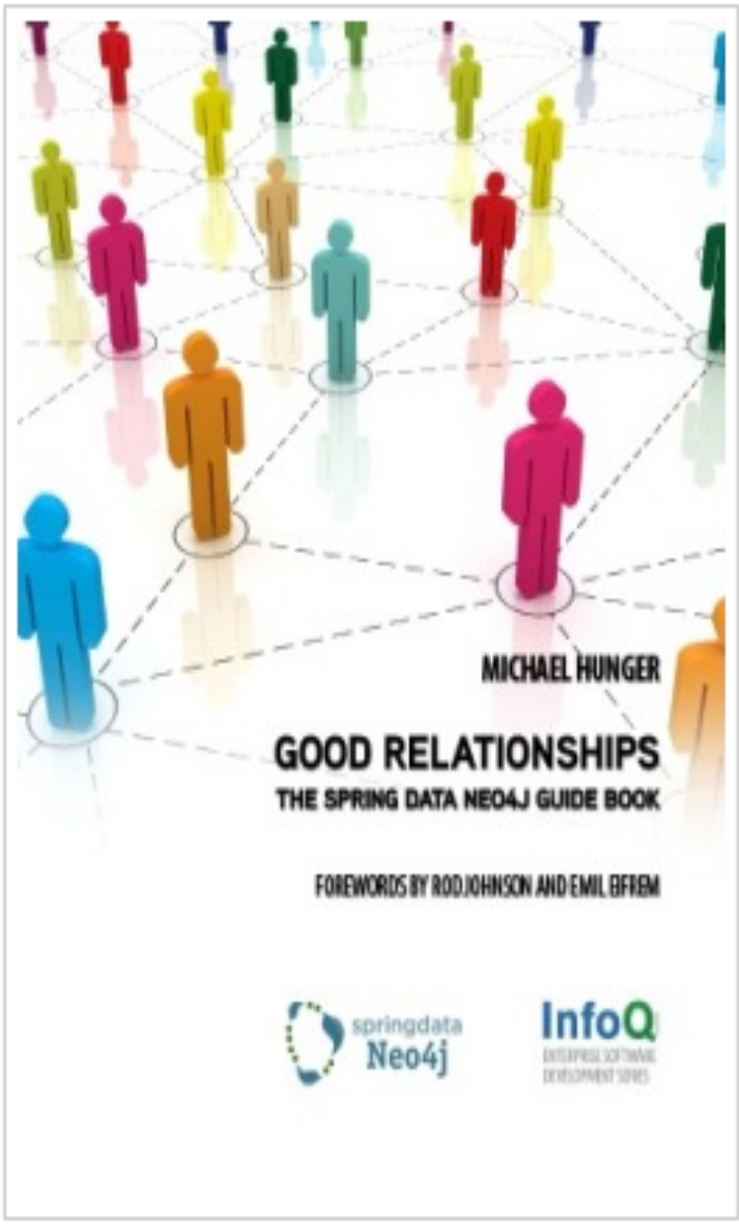
Add yourself »

http://www.neo4j.org/participate/contributors#map

Books


Neo4j
IN ACTION
Jonas Partner
Aleksa Vukotic
MEAP
MANNING


Early Release
RAW & UNEDITED
Graph Databases
O'REILLY
Ian Robinson
Jim Webber & Emil Eifrem


MICHAEL HUNGER
GOOD RELATIONSHIPS
THE SPRING DATA NEO4J GUIDE BOOK
FOREWORDS BY ROD JOHNSON AND EMIL EFREM
springdata
Neo4j
InfoQ

How to reach us:

Web:

- www.neo4j.org,
- www.neotechnology.com

Twitter: @Neo4j, @Neo4jDE

Email: dirk.moeller@neotechnology.com

Phone: +49 800 723 6231

Q&A

http://www.neo4j.org
http://groups.google.com/group/neo4j
http://www.neo4j.org/develop

Thank you!