CIFF
Centro Internacior
de Formación Fina

CIFF Trustees:

Santander

Universidad de Alcalá

Social Network Analysis

Lecture 5: Data Modelling in Neo4j

David Martín-Corral Calvo
dmartincc@smartvel.com

Abril / Mayo

MASTER EN BA & BD

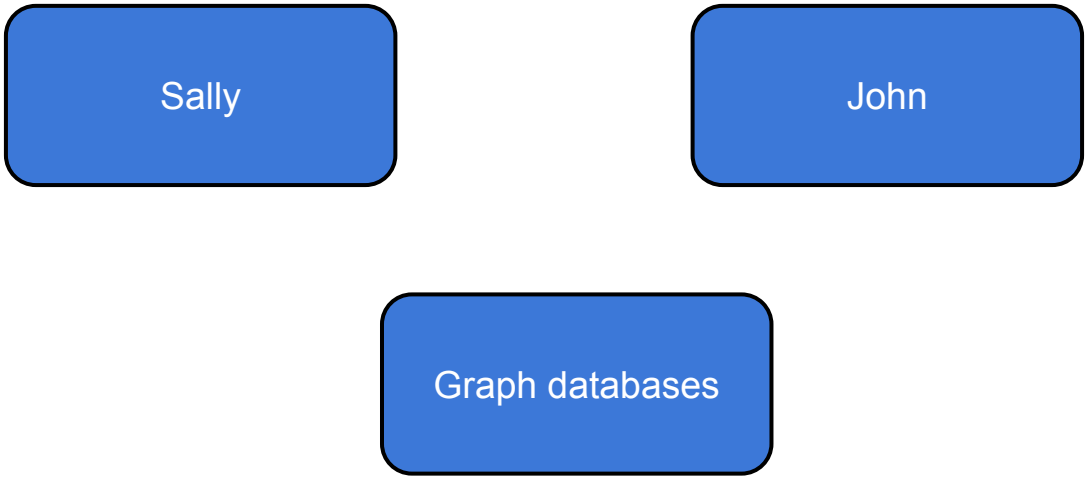
Recap:

- The world is small if you look at it as a network
- It has lots of interesting structures.
- Motifs: Given a particular structure, search for it in the network, e.g. complete triads

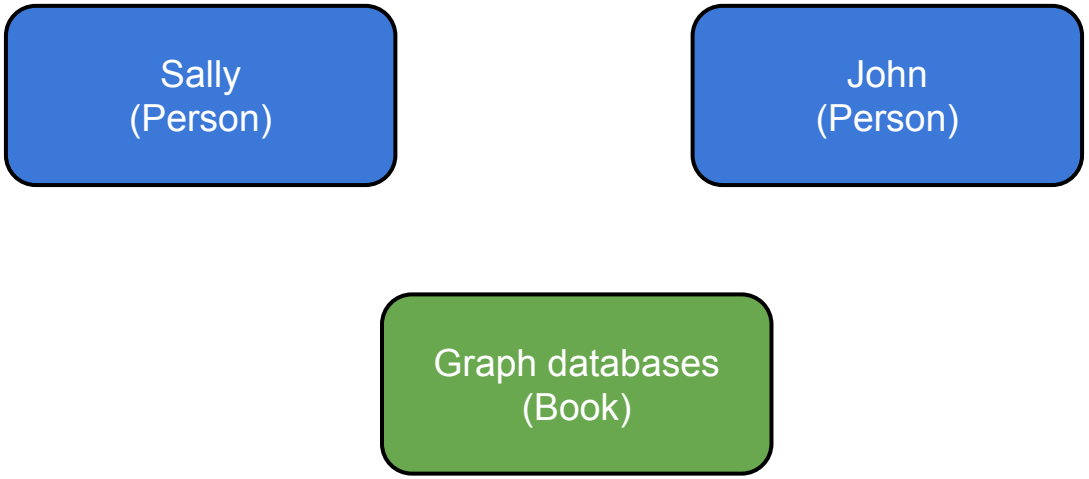
Outline:

- Data Models
- Questions to queries
- Technical analysis

Nodes



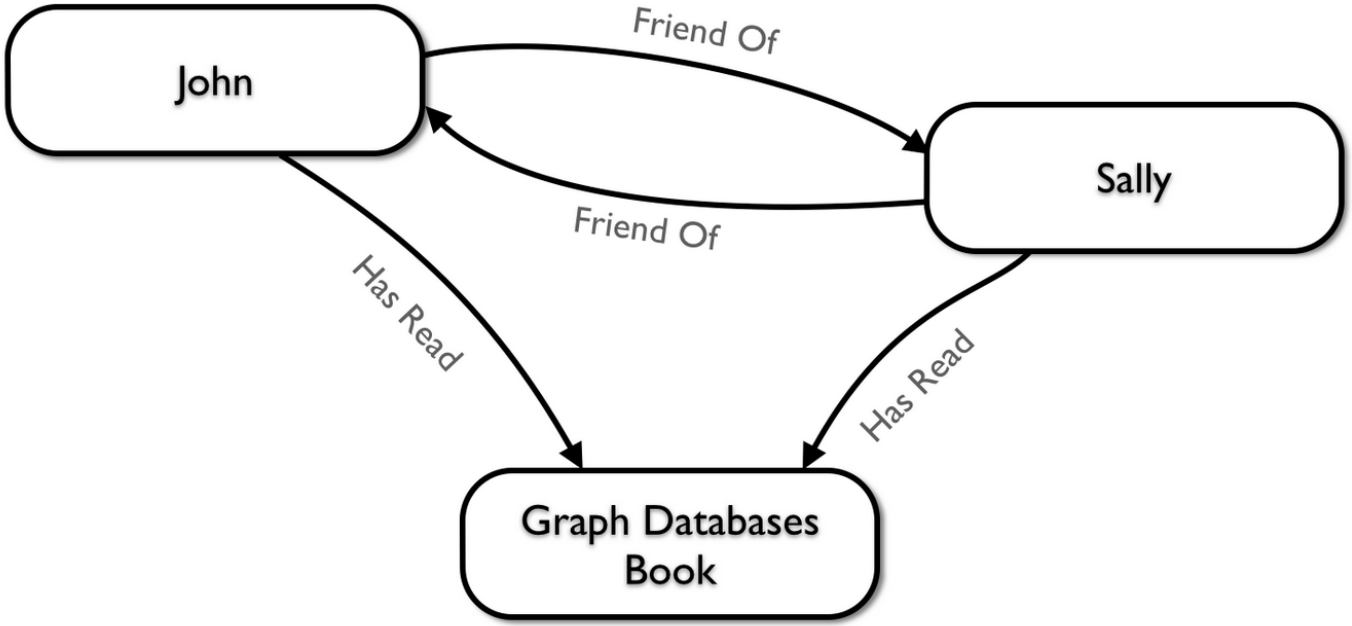
Labels



Relationships (Edges)

- John is a friend of Sally
- Sally is a friend of John
- John has read Graph Databases
- Sally has read Graph Databases

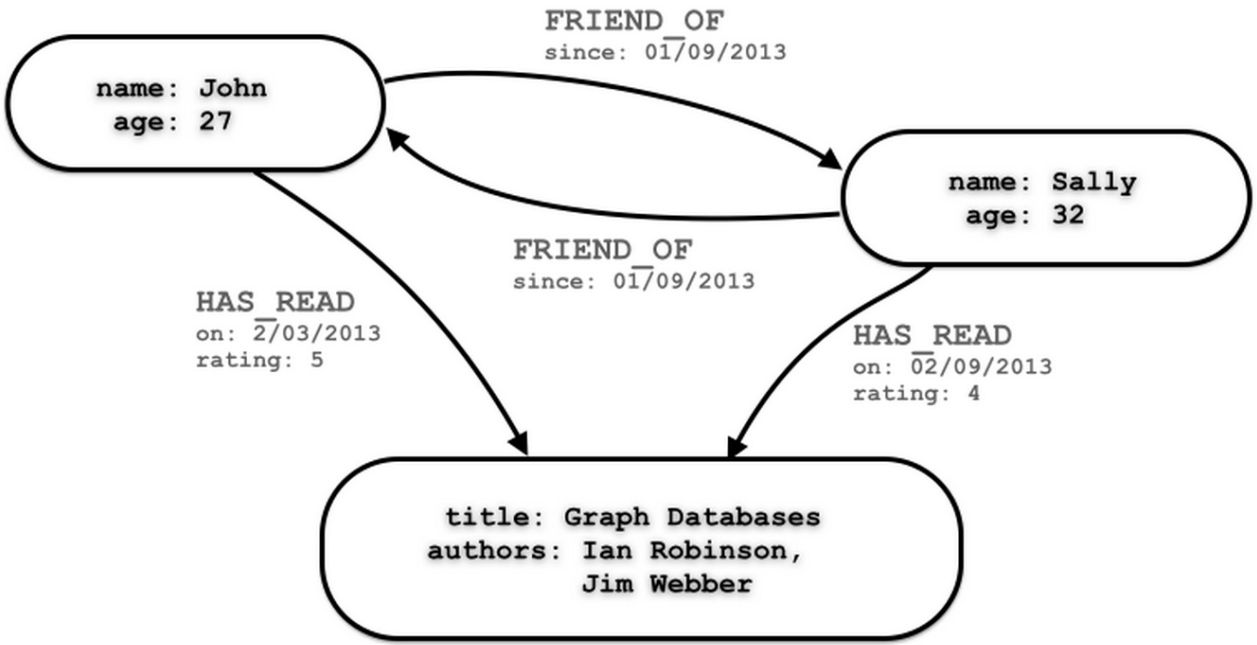
Data Model


John
Friend Of
Sally
Friend Of
Has Read
Has Read
Graph Databases
Book

Questions that can be answered

- When did John and Sally become friends?
- What is the average rating of the book Graph Databases?
- Who is the author of the book Graph Databases?
- How old is Sally?
- How old is John?
- Who is older, Sally or John?
- Who read the book *Graph Databases* first, Sally or John?

Attributes


FRIEND_OF
since: 01/09/2013
name: John
age: 27
name: Sally
age: 32
FRIEND_OF
since: 01/09/2013
HAS_READ
on: 2/03/2013
rating: 5
HAS_READ
on: 02/09/2013
rating: 4
title: Graph Databases
authors: Ian Robinson,
Jim Webber

1. Go to GrapheneDB
2. Create database
3. Neo4j web admin interface

Dataset

```
// Create Sally
CREATE (sally:Person { name: 'Sally', age: 32 })

// Create John
CREATE (john:Person { name: 'John', age: 27 })

// Create Graph Databases book
CREATE (gdb:Book { title: 'Graph Databases',
                   authors: ['Ian Robinson', 'Jim Webber'] })

// Connect Sally and John as friends
CREATE (sally)-[:FRIEND_OF { since: 1357718400 }]->(john)

// Connect Sally to Graph Databases book
CREATE (sally)-[:HAS_READ { rating: 4, on: 1360396800 }]->(gdb)

// Connect John to Graph Databases book
CREATE (john)-[:HAS_READ { rating: 5, on: 1359878400 }]->(gdb)
```

When did John and Sally became friends?

```
MATCH (sally:Person { name: 'Sally' })
MATCH (john:Person { name: 'John' })
MATCH (sally)-[r:FRIEND_OF]-(john)
RETURN r.since as friends_since
```

What is the average rating of Graph Database?

```
MATCH (gdb:Book { title: 'Graph Databases' })
MATCH (gdb)<-[r:HAS_READ]-()
RETURN avg(r.rating) as average_rating
```

How old is Sally?

```
MATCH (sally:Person { name: 'Sally' })
RETURN sally.age as sally_age
```

Who is older?

```
MATCH (people:Person)
WHERE people.name = 'John' OR people.name = 'Sally'
RETURN people.name as oldest
ORDER BY people.age DESC
LIMIT 1
```

What is the average rating of Graph Database?

```
MATCH (gdb:Book { title: 'Graph Databases' })
MATCH (gdb)<-[r:HAS_READ]-()
RETURN avg(r.rating) as average_rating
```

Recap:

- Data model
- Nodes
- Edges
- Attributes

Activity: Load csv to neo4j

Python script

nodes.csv

relationships.csv

Reading list:

Data Modelling Guidelines for Neo4j

Neo4j Tutorial

CIFF

Centro Internacional
de Formación Financiera

CIFF Trustees:

Santander

Universidad de Alcalá