

Graph Data in R with RNeo4j

Jesús Barrasa

London, 30th Nov 2015

London R Meeting

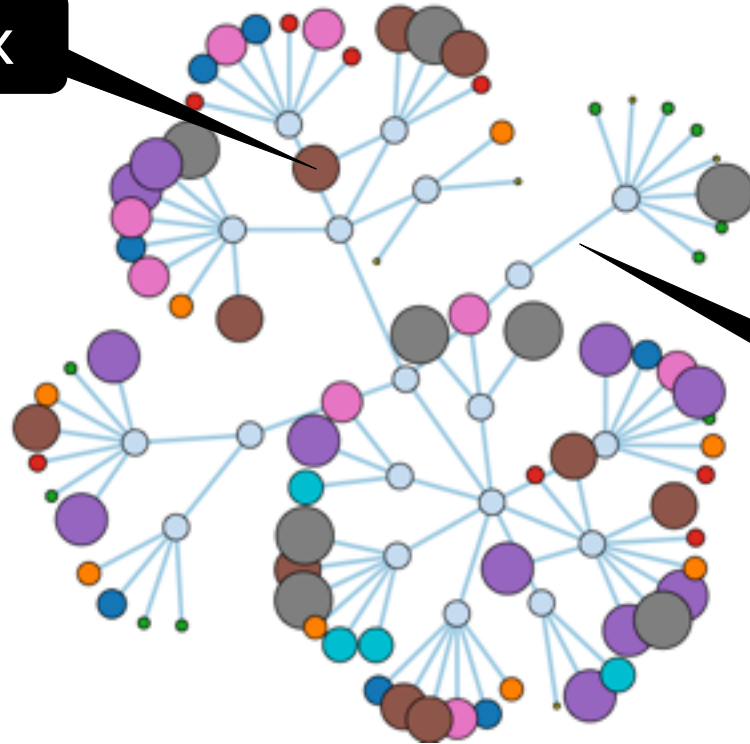


What is a graph?



Vertex

Edge



R packages

iGraph

others?



Graphs are everywhere!



Social Network



Network Impact Analysis



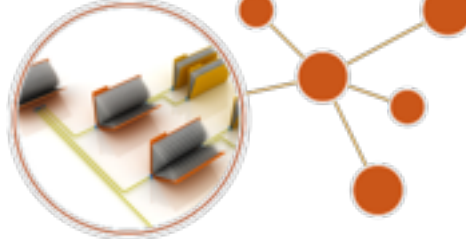
Recommendations



Logistics



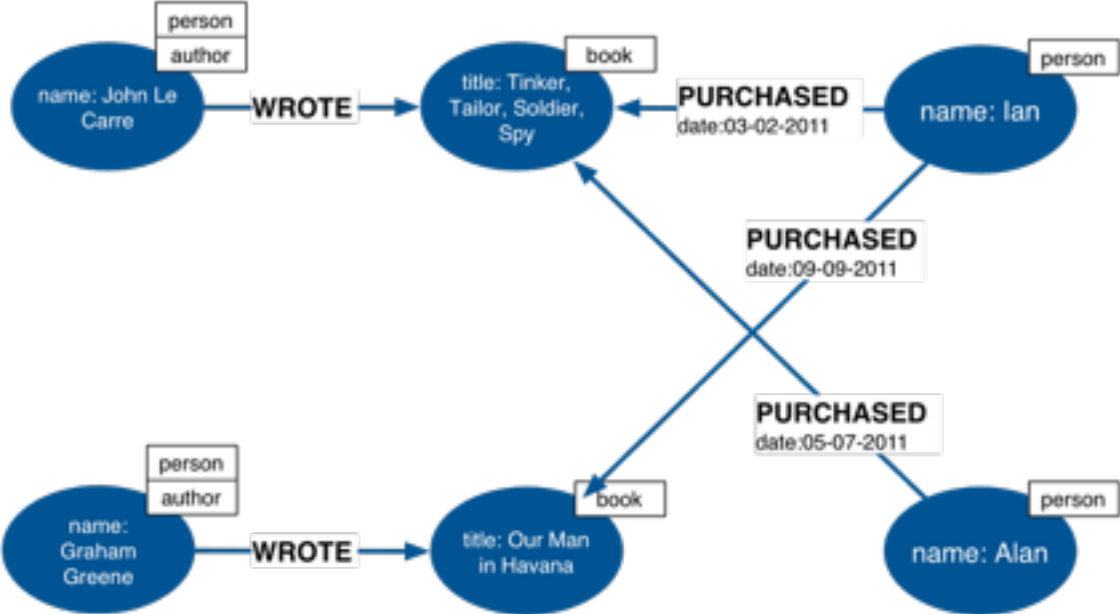
Access Control



Fraud Analysis



Labeled Property Graph



3 minute crash course on Graph Databases



The screenshot shows a web browser window at localhost:7474/browser/. The address bar contains the URL. The main content area displays a Cypher query:

```
1 // Create a node... and query it
2 CREATE (n:Person {name:"Jim"}) RETURN "hello", n.name
```

Below the query, there is a "play start" button. The main content area features the Neo4j logo (version 2.2.5) and two columns of links:

- Learn about Neo4j**
A graph epiphany awaits you.
What is a graph database?
How can I query a graph?
What do people do with Neo4j?
[Start Learning](#)
- Jump into code**
Use Cypher, the graph query language.
Code walk-throughs
RDBMS to Graph
Query templates
[Write Code](#)

At the bottom, there is a link: [The 5-Minute Interview: Andrew Royer, Independent Neo4j Hacker - Neo4j Graph Database](#). The footer contains the text: Copyright © Neo Technology 2002-2015.



Installing and loading RNeo4j

```
install.packages("RNeo4j")
```

```
library(RNeo4j)
```

Connect to your Neo4j instance

```
graph = startGraph("http://localhost:7474/db/data/",  
                  username="neo4j", password="neo")
```


Create nodes and connect them with relationships

```
#nodes
nicole = createNode(graph, "Person", name="Nicole",
                    userId="nicole@neo4j.com")
greta = createNode(graph, "Person", name="Greta",
                   userId="greta@neo4j.com")
kenny = createNode(graph, "Person", name="Kenny",
                   userId="kenny@gmail.com")

#rels
r1 = createRel(greta, "LIKES", nicole,
              when="30/11/2015", how="iPhone App", weight=7)
r2 = createRel(nicole, "LIKES", kenny,
              when="29/11/2015", weight=2)
```

Querying the Neo4j Graph

```
query = "  
MATCH (nicole:Person)-[r:LIKES]-(p:Person)  
WHERE nicole.name = 'Nicole'  
RETURN nicole.name, r.weight, p.name  
"
```

```
cypher(graph, query)
```

```
##  nicole.name r.weight p.name  
## 1      Nicole         2  Kenny  
## 2      Nicole         7  Greta
```

Path queries

```
p = shortestPath(greta, "LIKES", kenny, max_depth=4)
n = nodes(p)
sapply(n, "[", "name")
```

```
## [1] "Greta" "Nicole" "Kenny"
```

Data set from Movie Database

```
query = "  
MATCH (p:Person)-[:ACTED_IN]->(movie)<-[:ACTED_IN]-(coActor)  
RETURN p.name AS actor,  
       p.born AS yob,  
       count(distinct coActor) AS coActorCount,  
       count(distinct movie) AS movieCount  
"  
movieDataSet <- cypher(graph, query)  
  
head(movieDataSet, n=3)
```

##	actor	yob	coActorCount	movieCount
## 1	Tom Skerritt	1933	5	1
## 2	Geena Davis	1956	5	1
## 3	Clint Eastwood	1930	2	1

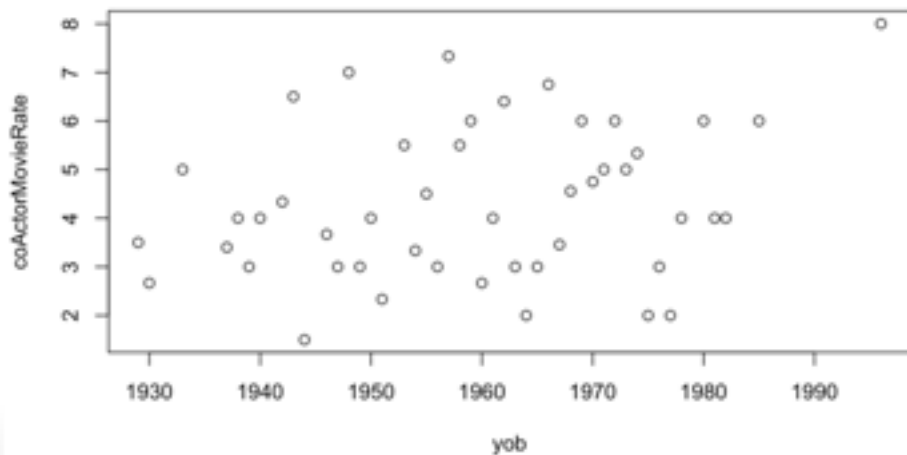
Co-actor - Movie rate data set

```
query = "  
MATCH (p:Person)-[:ACTED_IN]->(m)<-[:ACTED_IN]-(coActor)  
RETURN p.born AS yob,  
      toFloat(count(distinct coActor)) /  
      toFloat(count(distinct m)) AS coActorMovieRate"
```

```
coActorMovieRateDataSet <- cypher(graph, query)
```

Plot

```
plot(coActorMovieRateDataSet)
```



More R-esources



- RNeo4j

<https://github.com/nicolewhite/RNeo4j>

- Nicole White's talks:

Python, R and Neo4j - The Data Science Stack

<https://skillsmatter.com/skillscasts/7099-python-r-and-neo4j-the-data-science-stack>

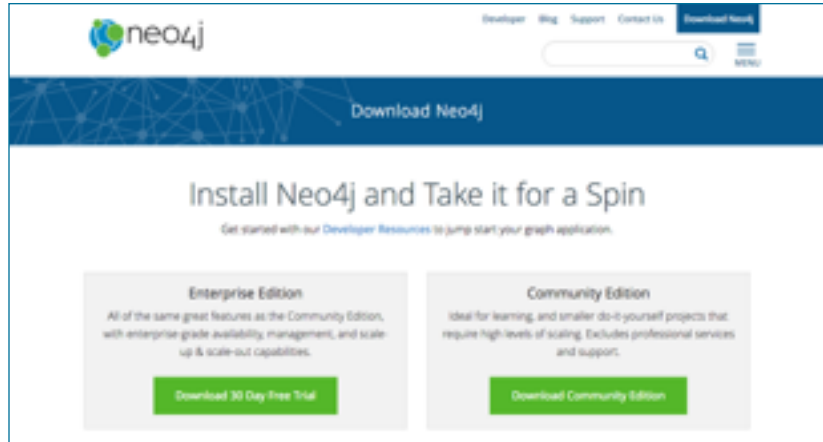
Neo4j, Graphs R Cool

<https://www.youtube.com/watch?v=bdQ90y9Pefo>

What next?



<http://neo4j.com/download/>



- **Tomorrow!!** Neo4j London Meetup. Session by Mark Needham on graph driven real time recommendations with Neo4j @SkillsMatter

