

Neo4j overview

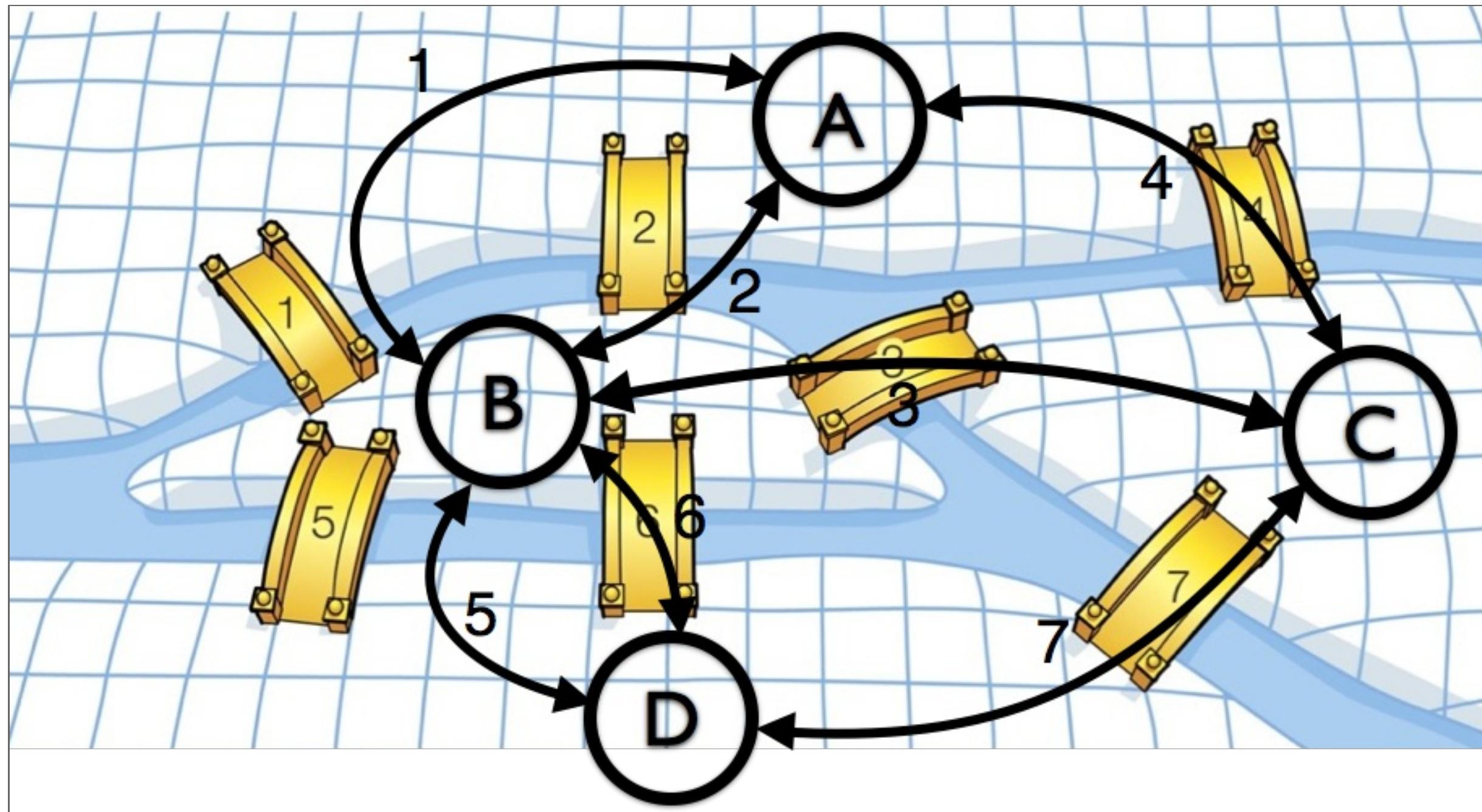
Benoit Simard (@logisma)

History

Leonhard Euler 1707-1783

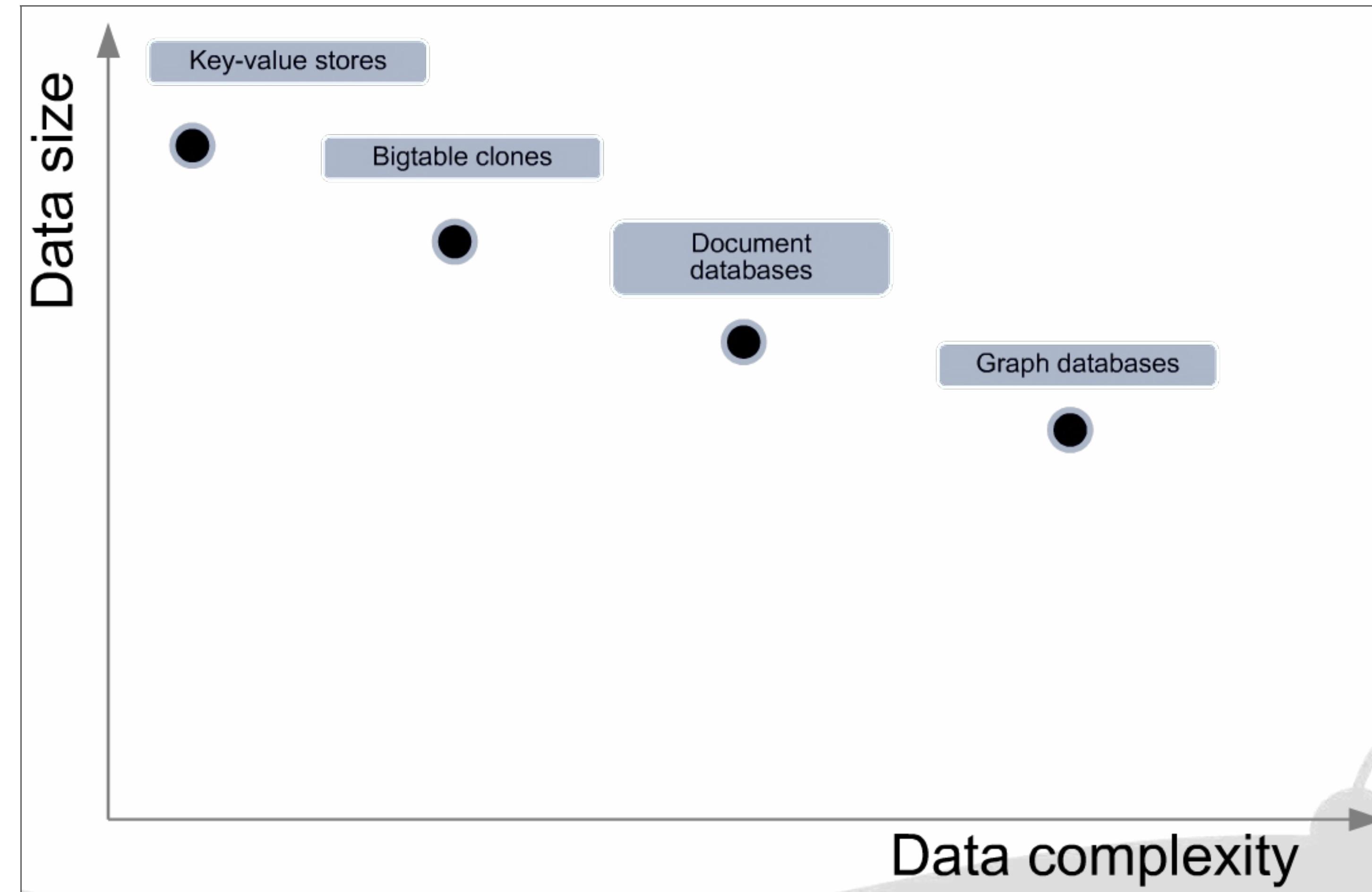


Konigsbergs



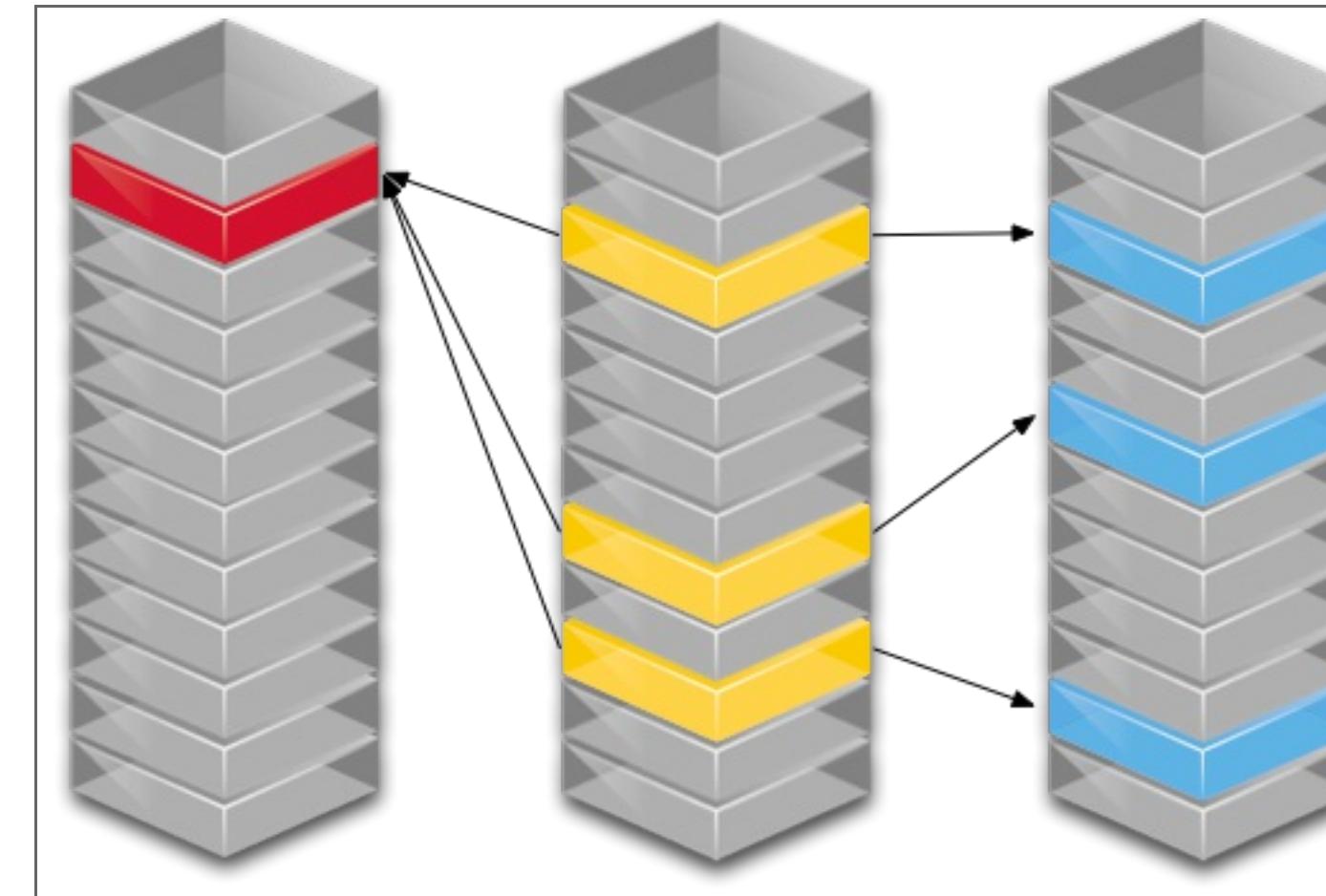
Why Graph database ?

Data complexity



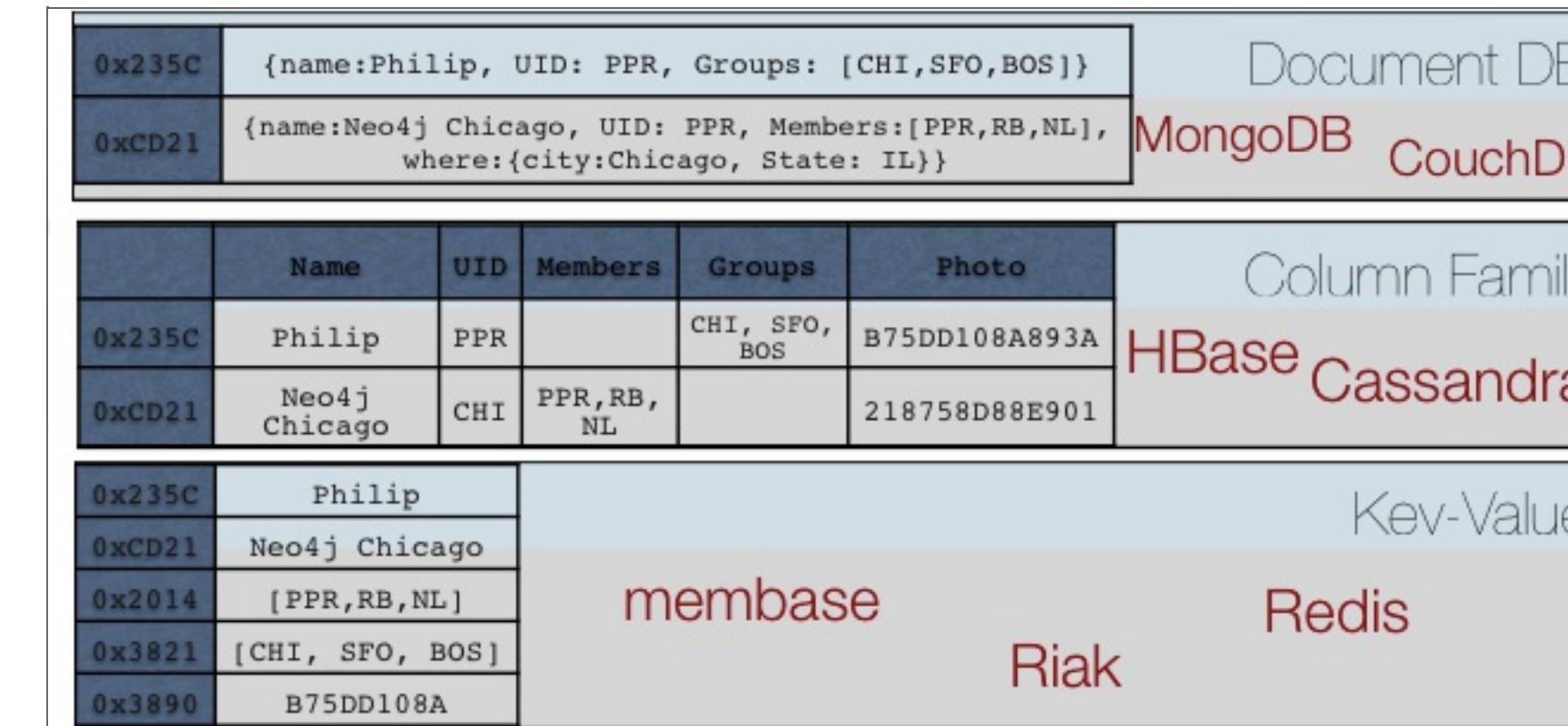
Linked Data

Relation in RDBMS

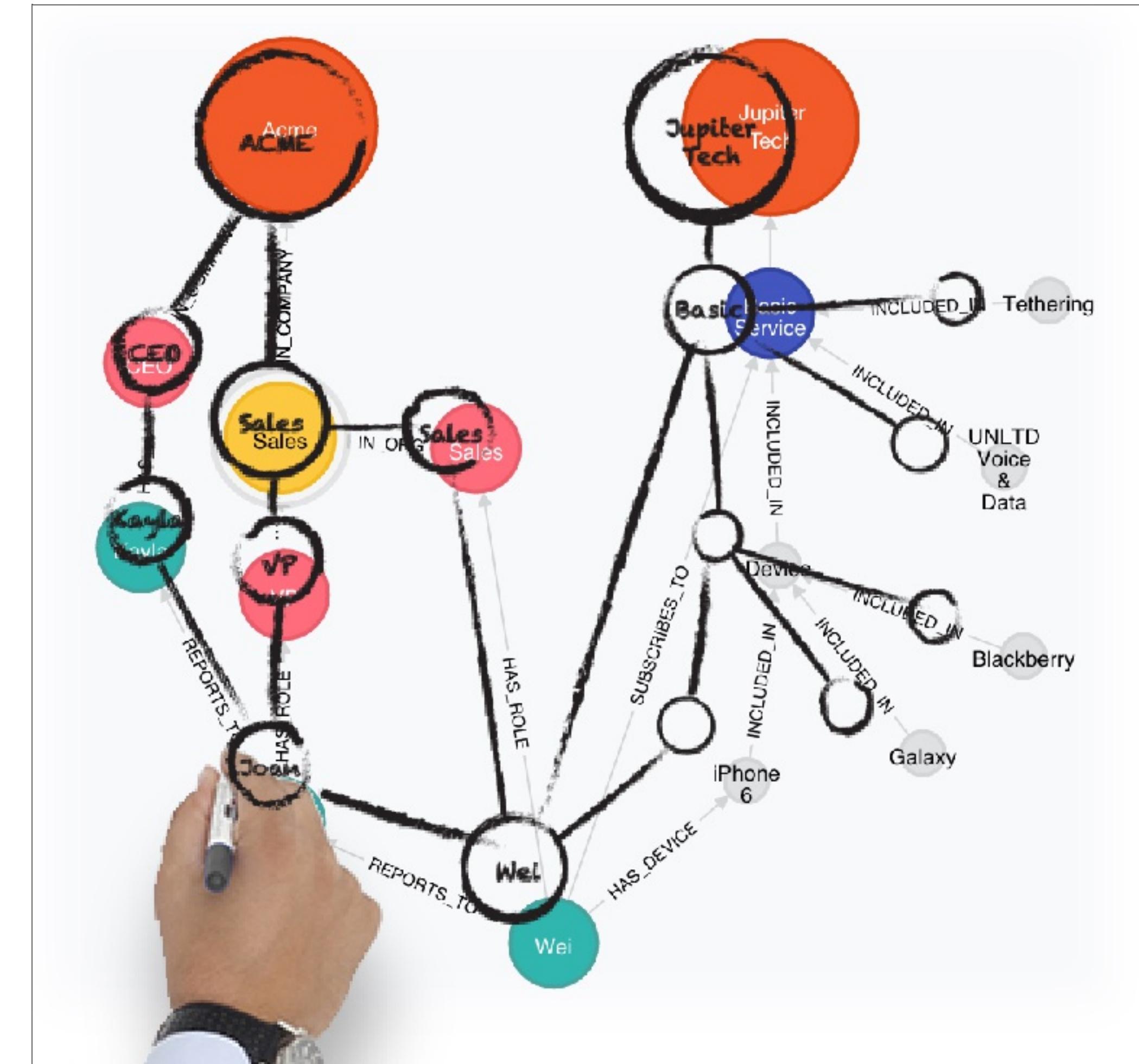


- schema not flexible
- poor performances on linked data (@see snowflake schema)
- complexe queries

Relation in No-SQL



Graph database



When to use it ?

- Do you need a flexible schema ?
- Do your data are highly connected ?
- Do you need near real time ?

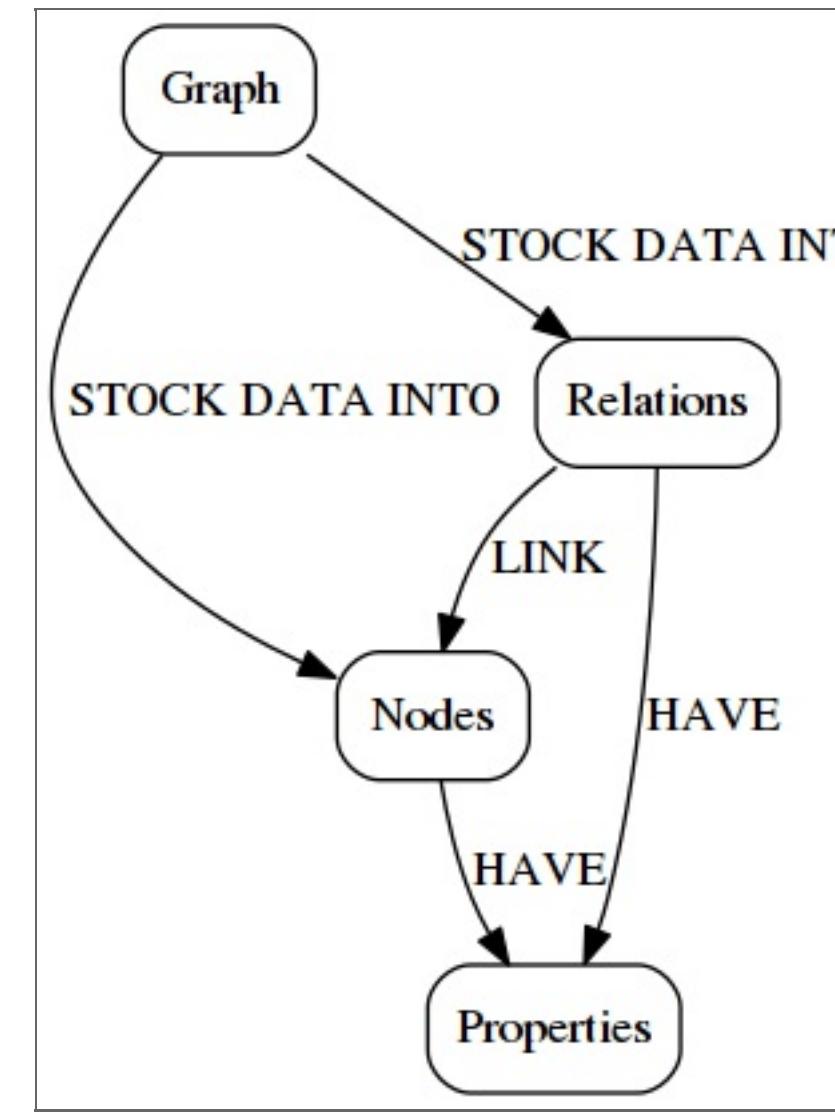
Recommendation, fraud detection, MDM, network analysis, impact analysis, ACL, ...

What is a Graph database ?

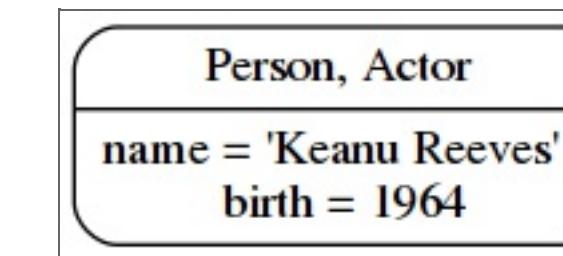
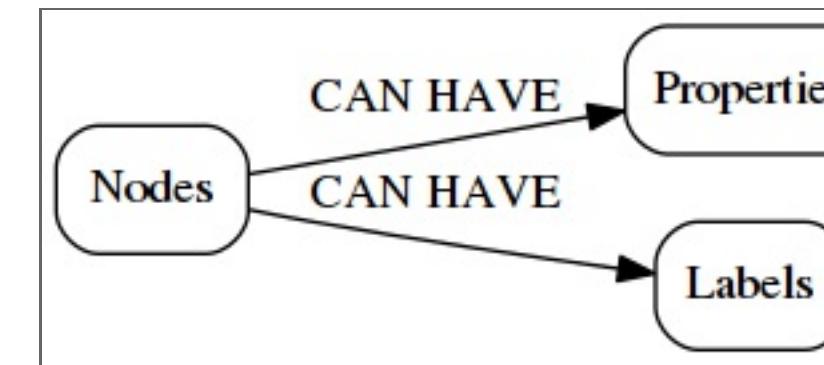
Neo4j

- Since 2010
- License under GPLv3
- Full ACID
- HA
- Schemaless

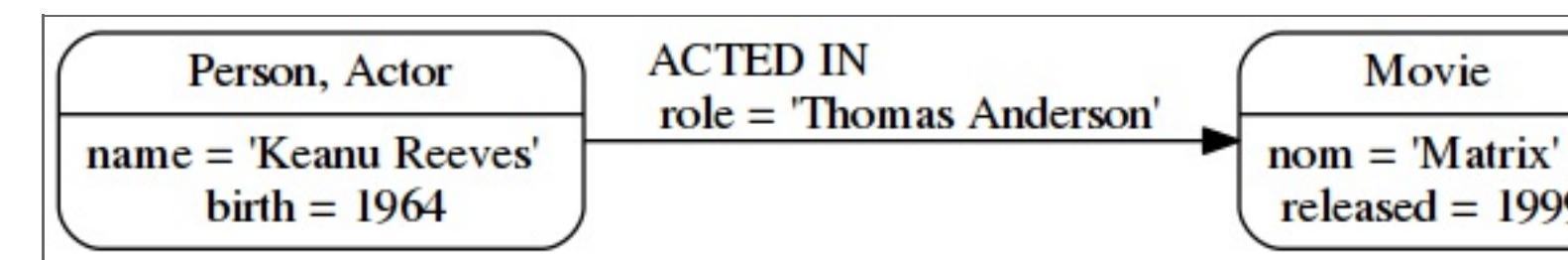
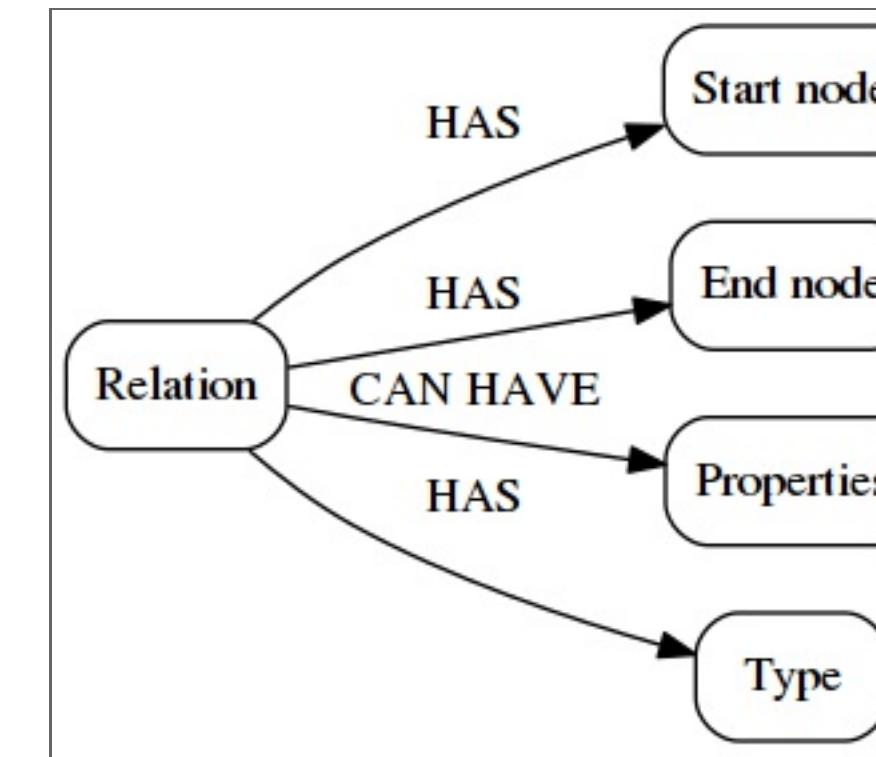
Concept



Nodes



Relationships



Index & constraint

Indexes

- on a property for all nodes that have a label

Constraint

- for the uniqueness of a property
- for the existence of a property

Top use-cases

Yeah !

(Graph)-[:ARE]→(Everywhere)

Moteur de recommandation en temps réel



La recommandation

Des réseaux sociaux, à la vente au détail, de service ou dans le secteur des médias, la recommandation est partout !

Objectif : offrir une recommandation en temps réel, hautement ciblés et contextuelles afin quelle soit pertinentes pour l'utilisateur

LES CLIENTS ONT ÉGALEMENT ACHETÉ...

Produit	Prix
AIM IR605Q...	14,50 €
FX-AUDIO...	125,00 €
Dayton Audio...	2,20 €
Gaine Thermo...	2,90 €
JJ...	13,50 €
Kiwame 5W -...	1,90 €
Kiwame 5W -...	1,90 €
Borniers...	7,50 €
JJ...	12,90 €
Condensateur...	1,40 €

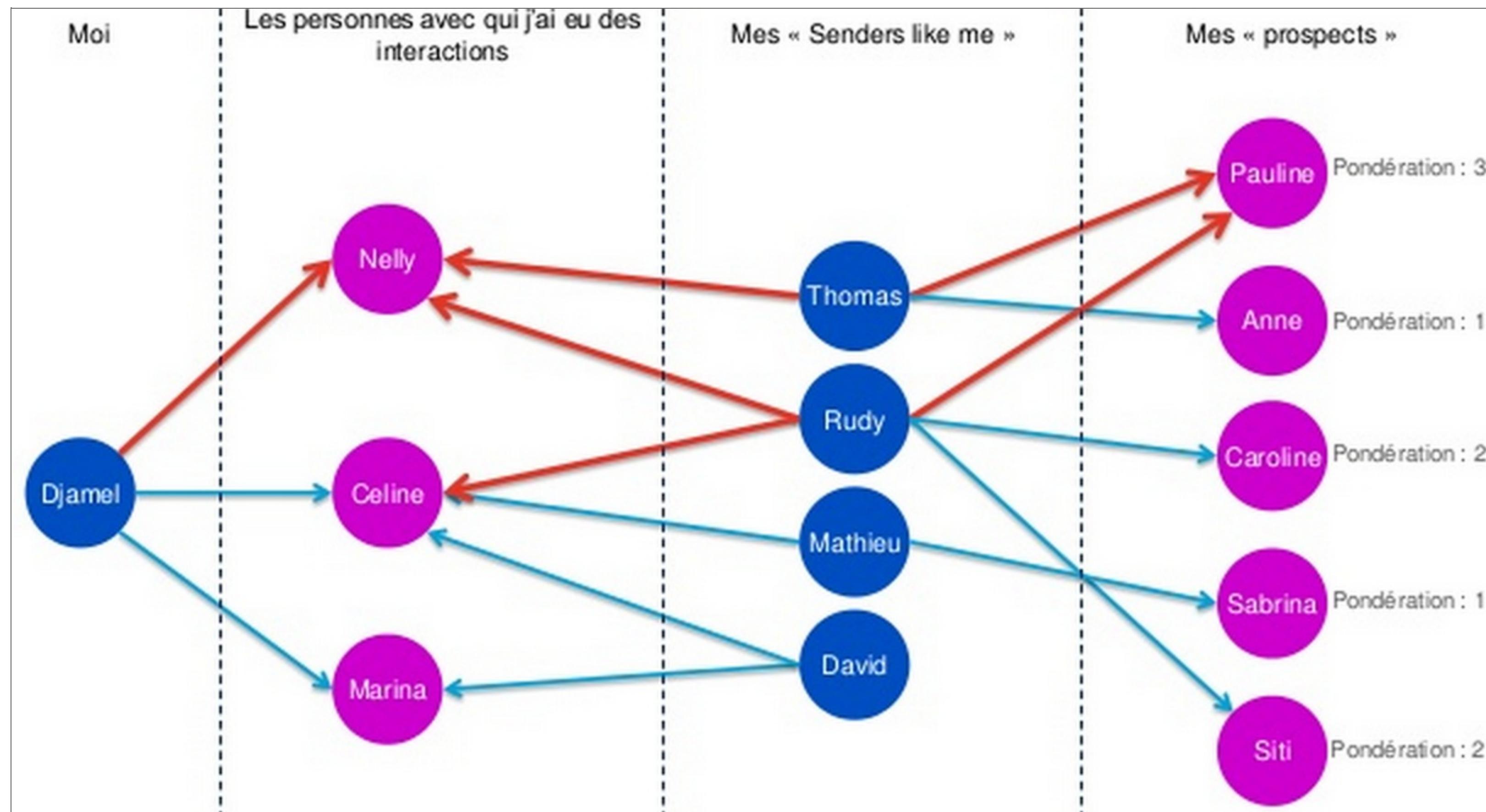
Suggestions · Actualiser · Tout afficher

 linkurious @Linkurious ×
Suivre

 Philip Rathle @prathle ×
Suivre

Trouver des amis

Etude de cas : Meetic

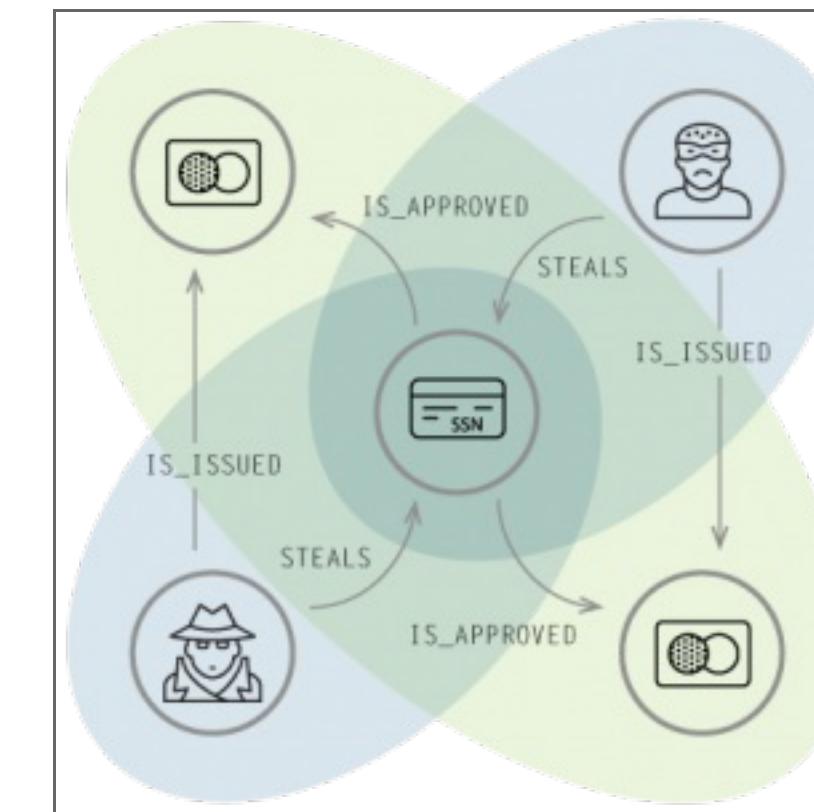


La détection de fraudes

Pourquoi du graphe ?

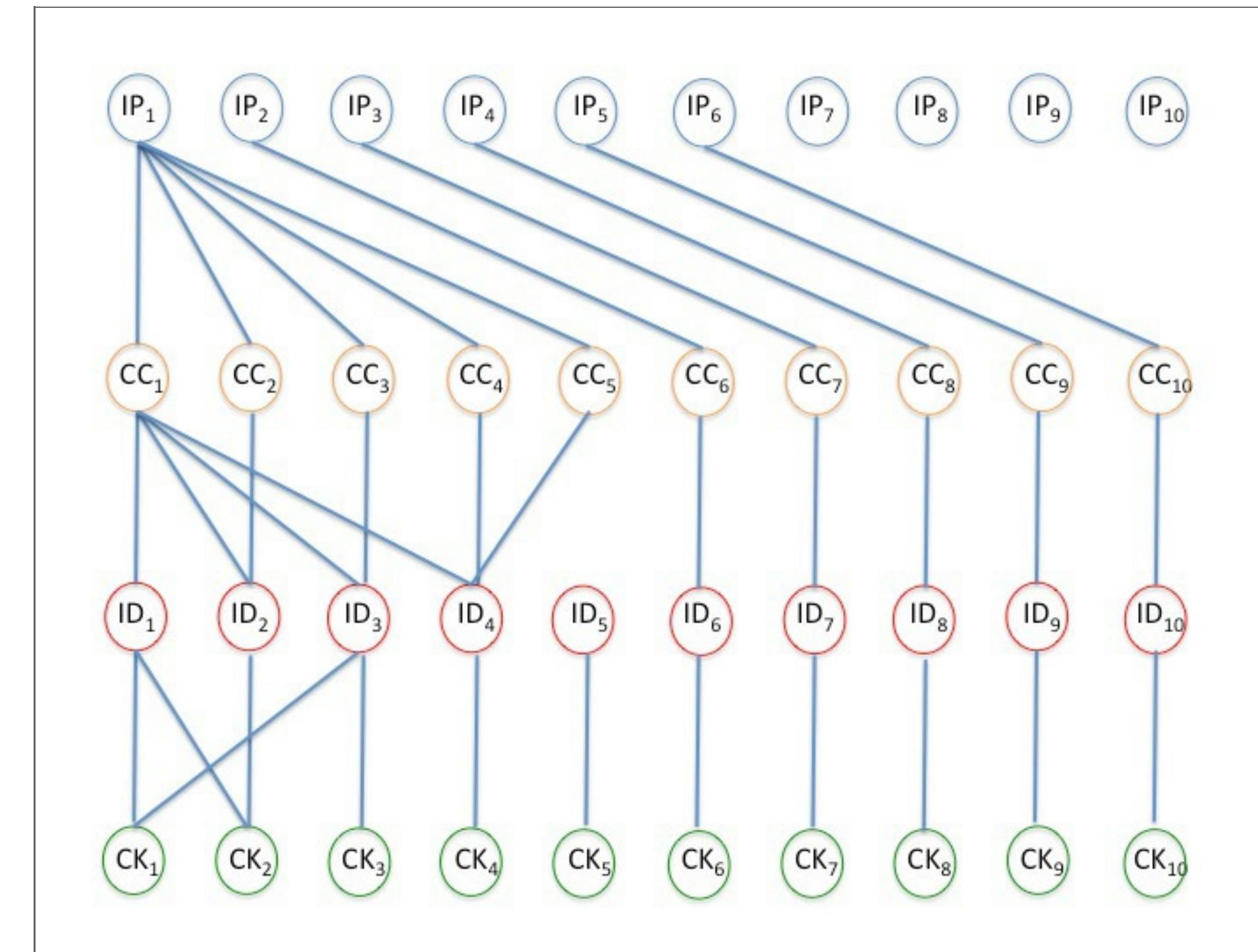
Bien qu'aucune mesure de détection de fraudes ne soit parfaite, il est possible d'introduire des améliorations considérables si l'on se penche sur les connexions qui relient les points de données individuels.

Il faut étudier l'interconnexion des données



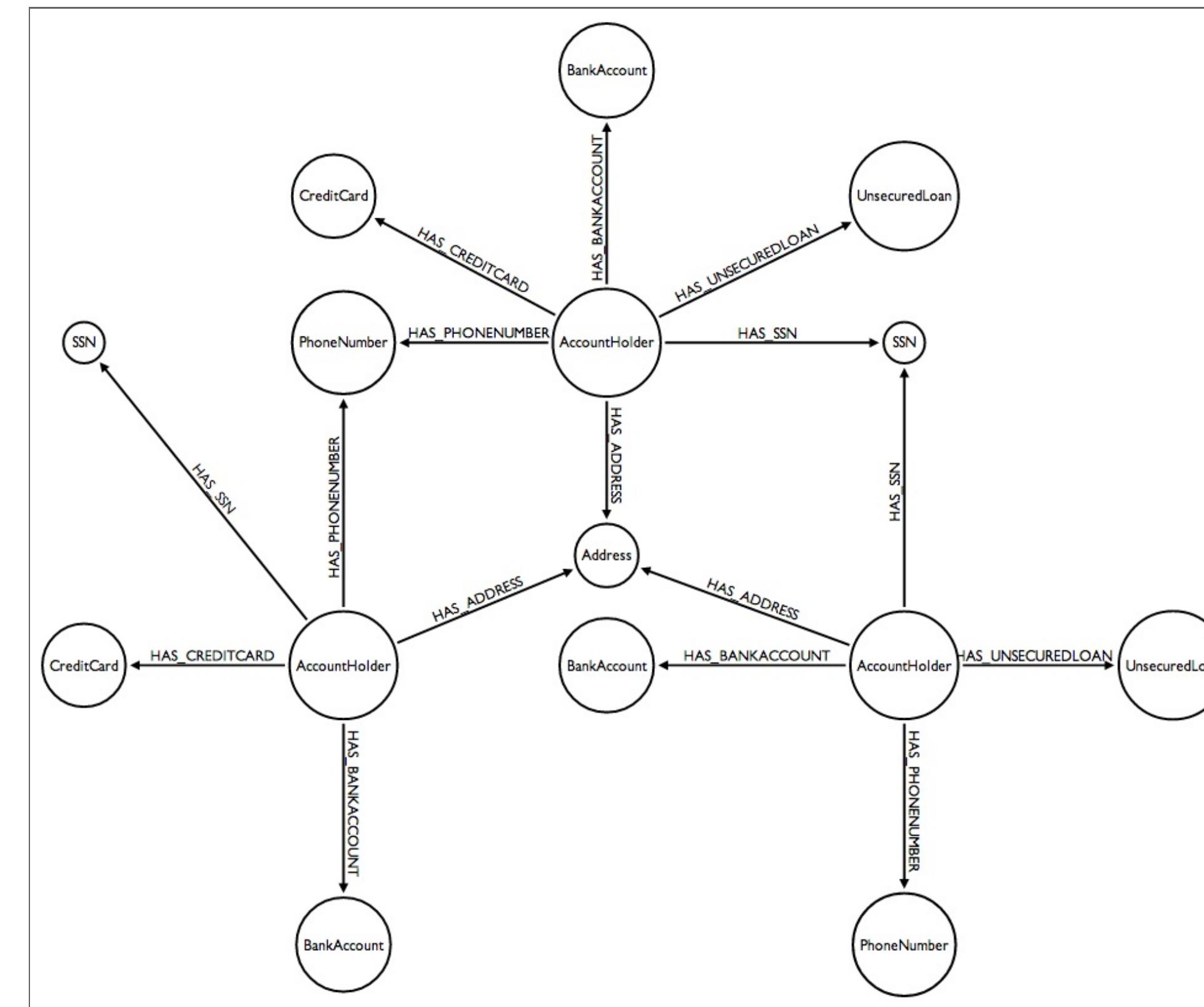
Etude de cas : Fraude dans le e-commerce

Graphe d'une série de transactions effectuée avec fraude probable se produisant à l'adresse IP1.
(IP : Adresse IP, CC : carte de crédit, ID : Identifiant de l'utilisateur, CK : Adresse de livraison)



Etude de cas : La détection d'anneau

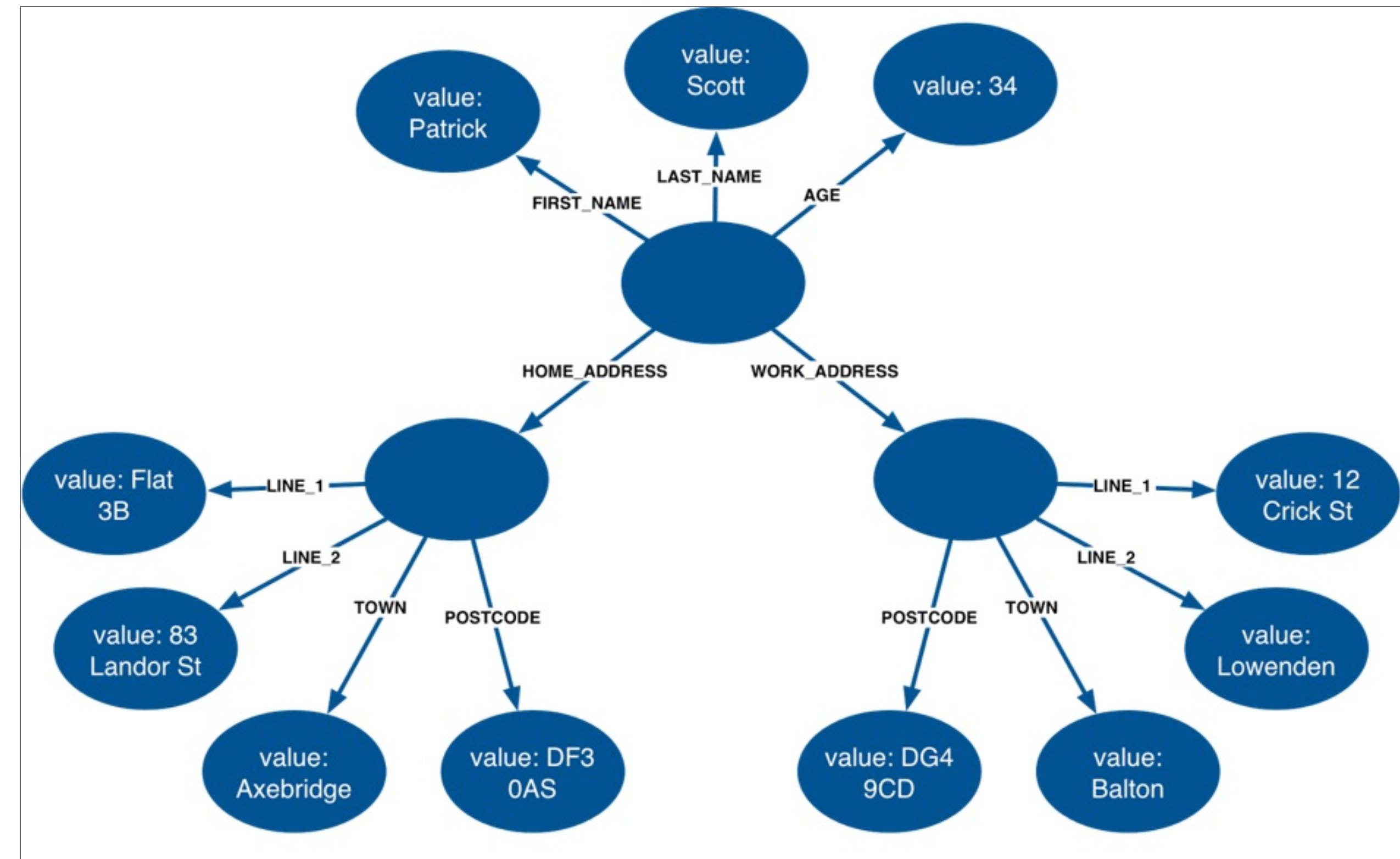
Constituer un ensemble de fausses identités à partir d'identités volées, dans l'objectif de commettre un acte frauduleux



La gestion de données de référence

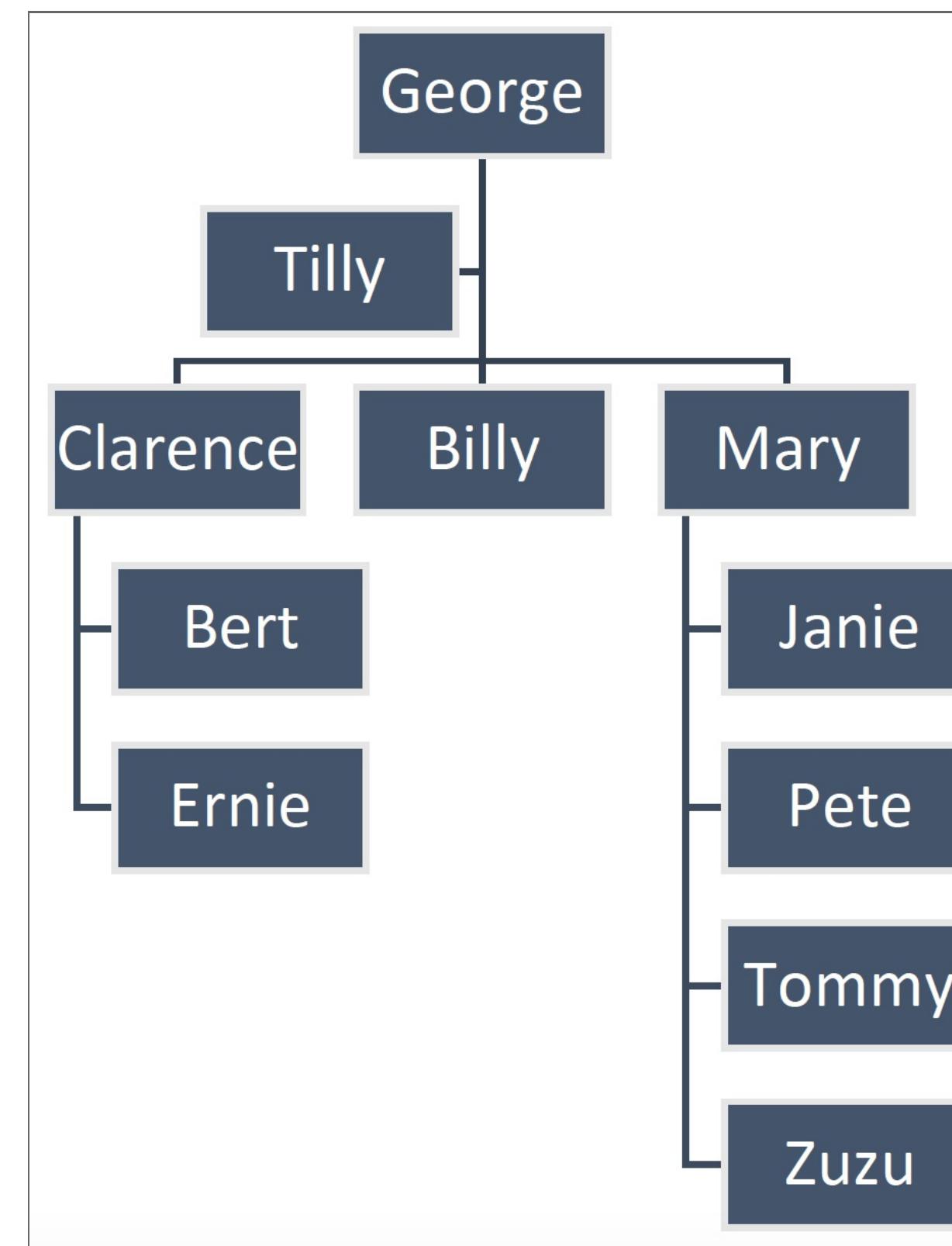
Reconciliation des données

Réconciliation des données par degré de similitude



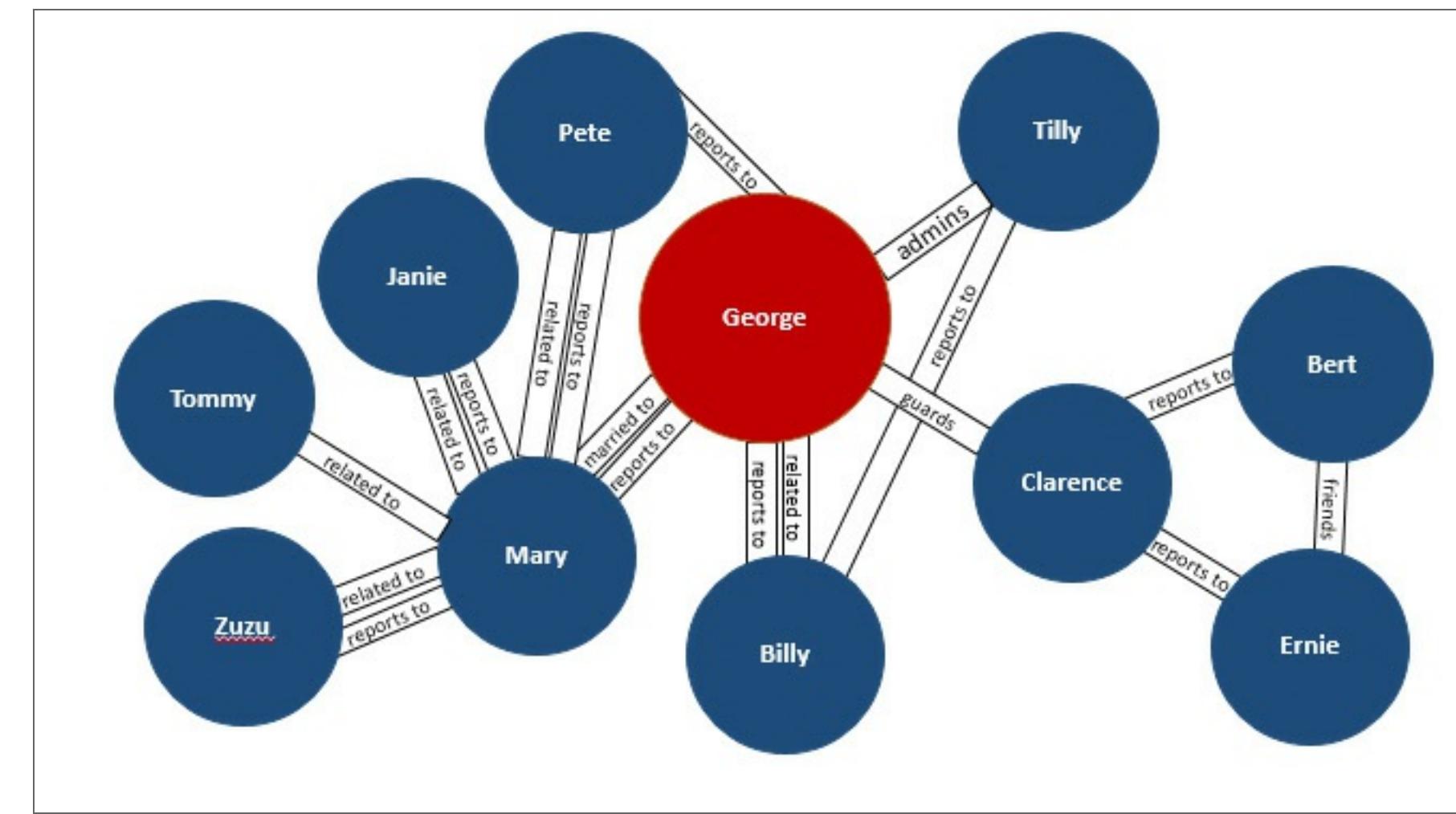
La modélisation RH d'une entreprise

Hiérarchie de données de référence illustrant des structures de rapport et de supervision d'employé.



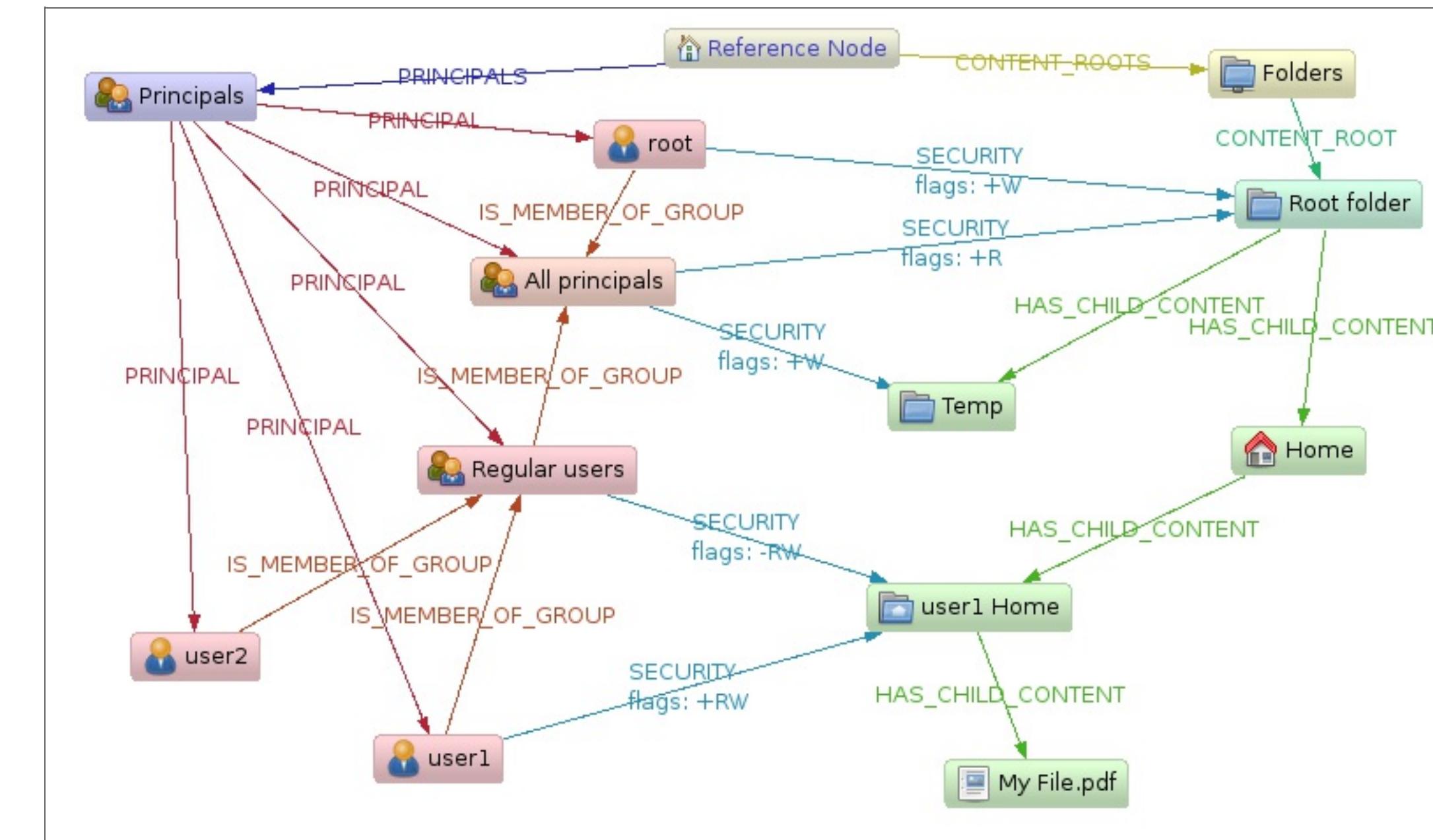
La modélisation RH d'une entreprise

Réseau de données de référence détaillant les relations de rapport et de supervision d'employé, ici avec une plus grande complexité du **monde réel**.



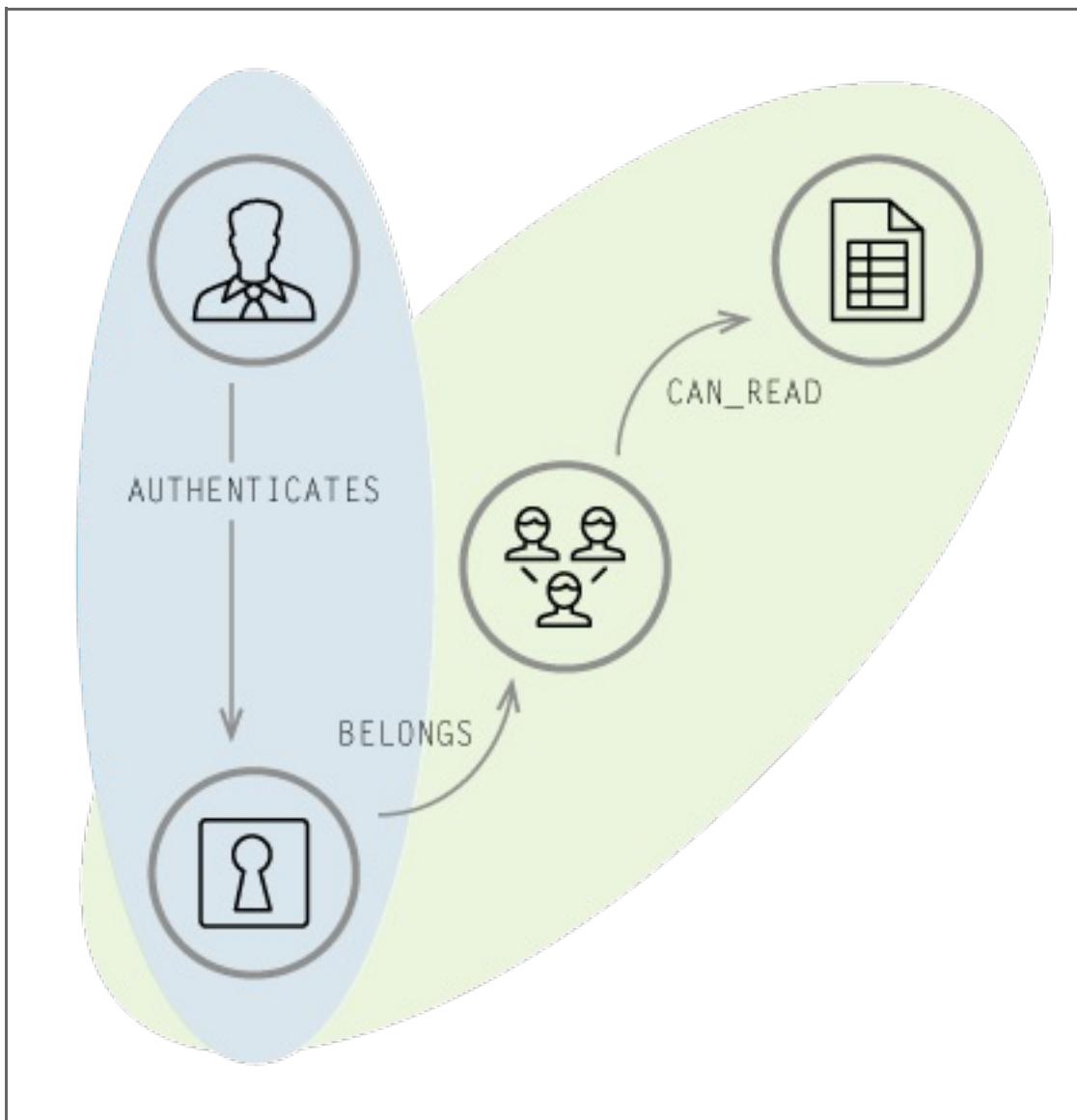
La gestion des identités et des accès

La gestion des identités et des accès



Pourquoi du graphe ?

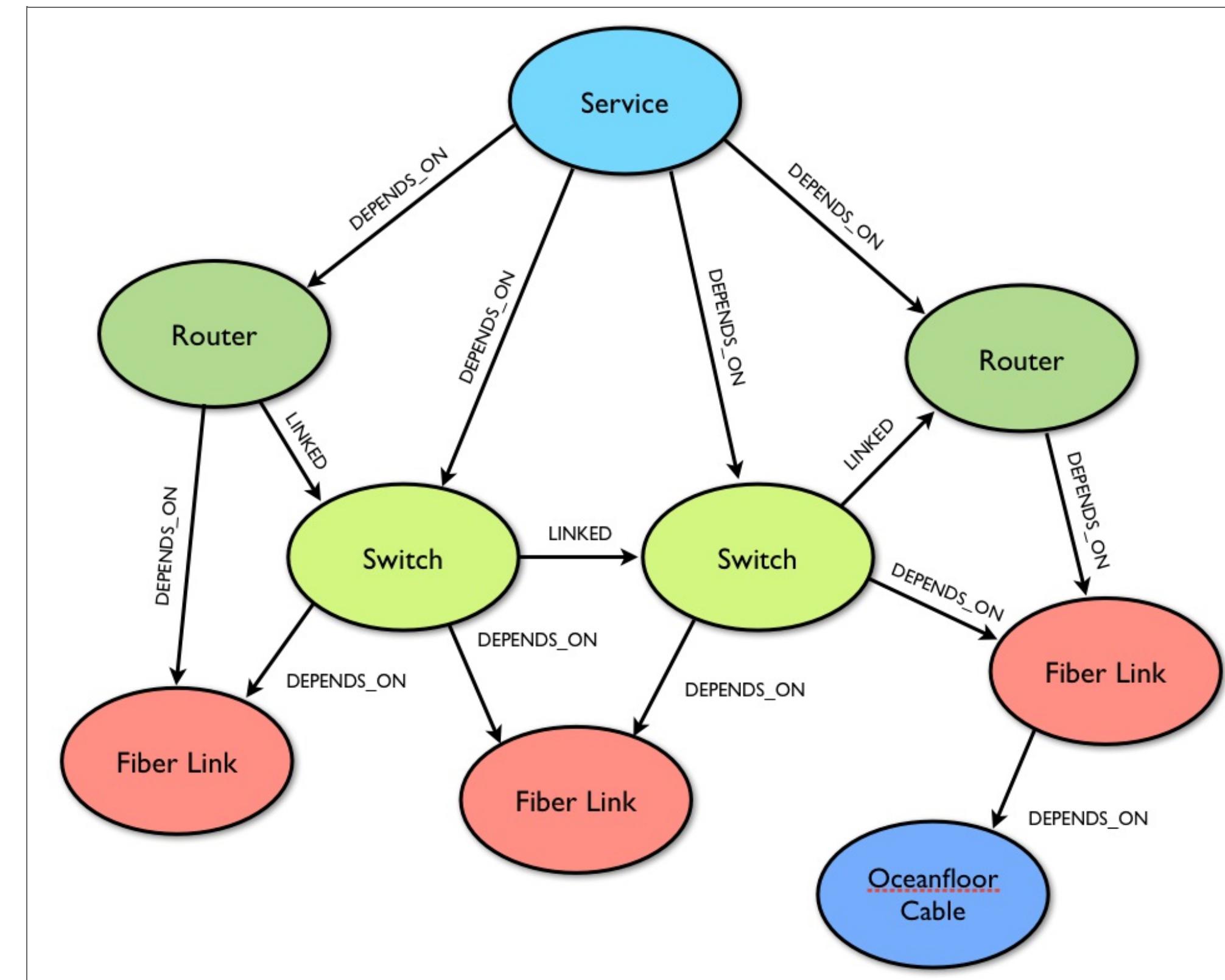
C'est un graphe !

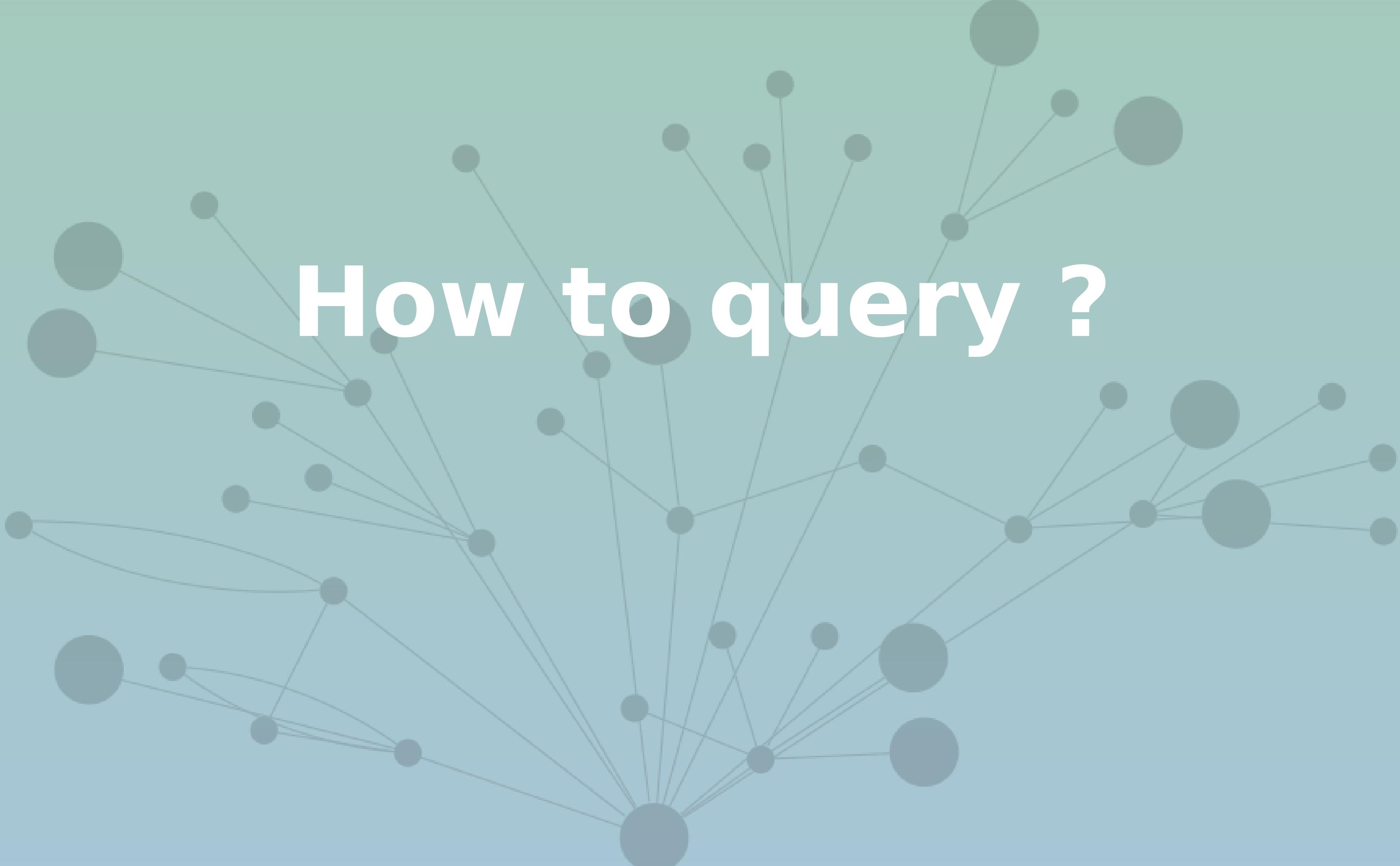


Les réseaux informatiques

Analyse d'impact chez SFR

Modélisation de leur réseau physique dans Neo4j.





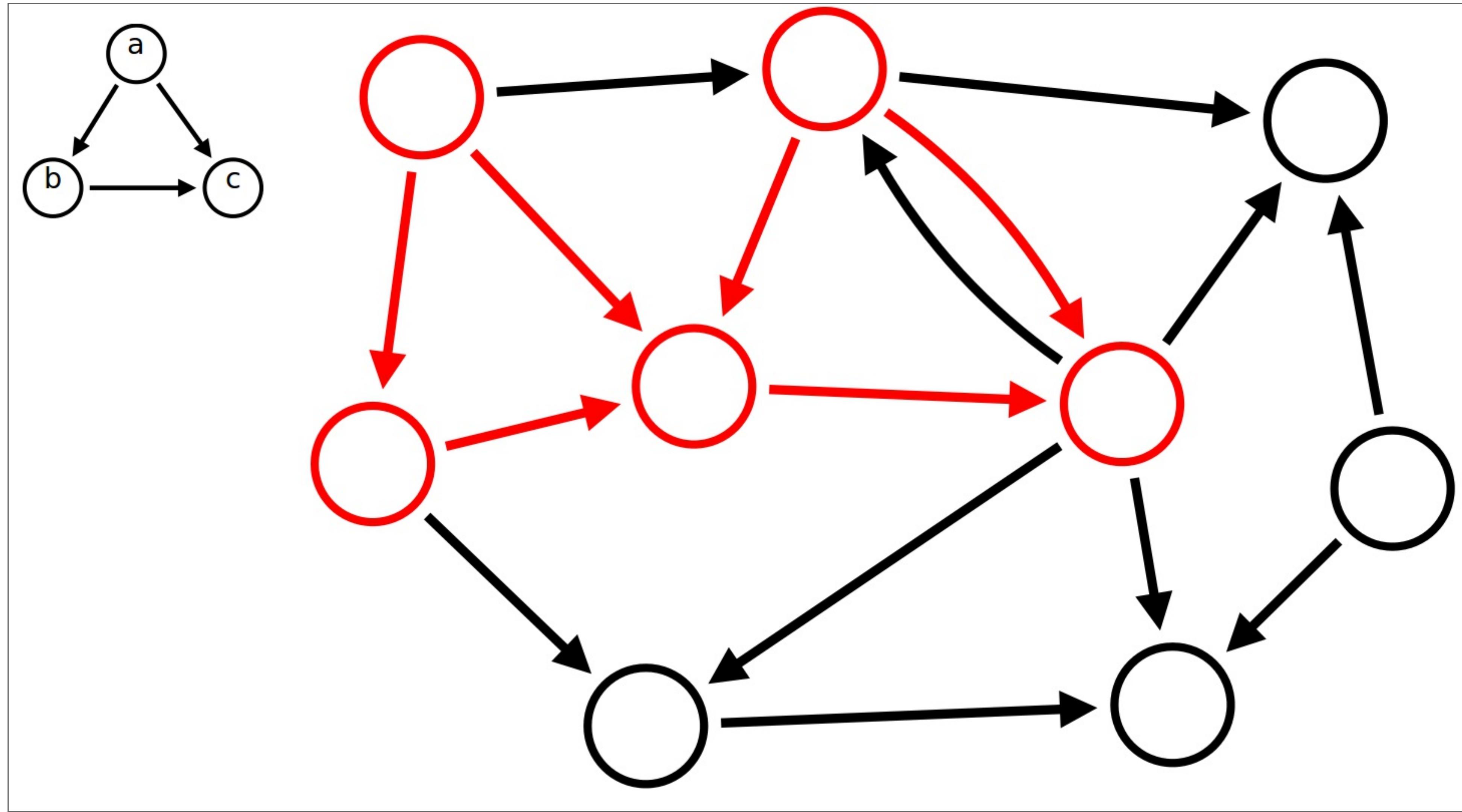
How to query ?

Cypher

Cypher is Neo4j's graph query language

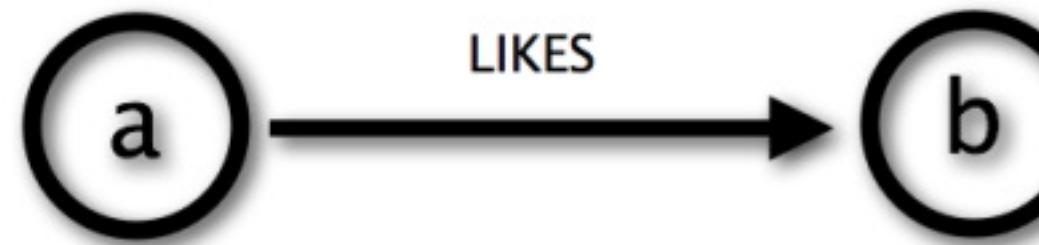
- Declarative Pattern-Matching language
- SQL-like syntax
- Designed for graphs

All is pattern



Ascii Art

Cypher using relationship 'likes'



Cypher

(a) -[:LIKES]-> (b)

© All Rights Reserved 2013 | Neo Technology, Inc.

Let see it in action

Indexes

Query

List all the characters in the movie “The Matrix”

Query

Query

Query

Query

Who are the five busiest actors?

Recommend 3 actors that Keanu Reeves should work with (but hasn't).

Add KNOWS relationships between all actors who acted in the same movie

Friends-of-Friends

Query

Query

Return the names of the people joining Charlize to Kevin.

QUESTION ?

PROCHAIN ÉVÈNEMENTS

- **Graph Day Paris** : mercredi 21 Septembre
- **Training Graph Modeling** : mardi 25 Octobre à Paris