

Top 5 des cas d'utilisations

Benoit Simard

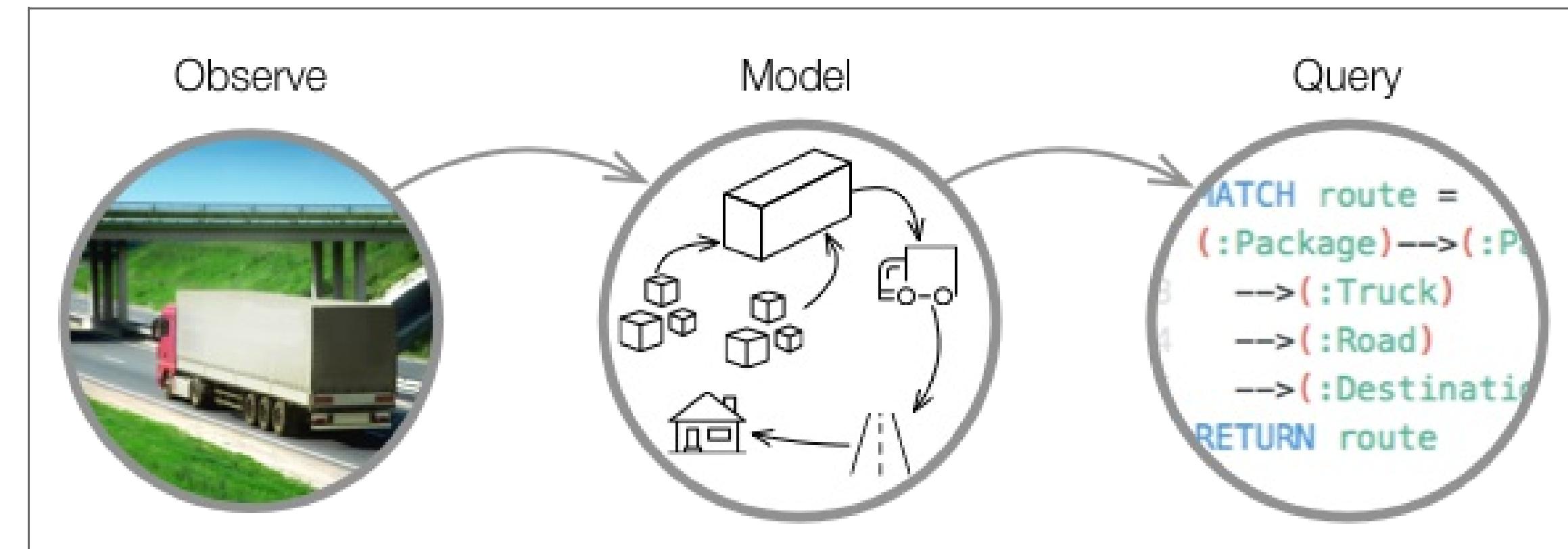
Pourquoi du graphe ?

Le monde est un graphe !



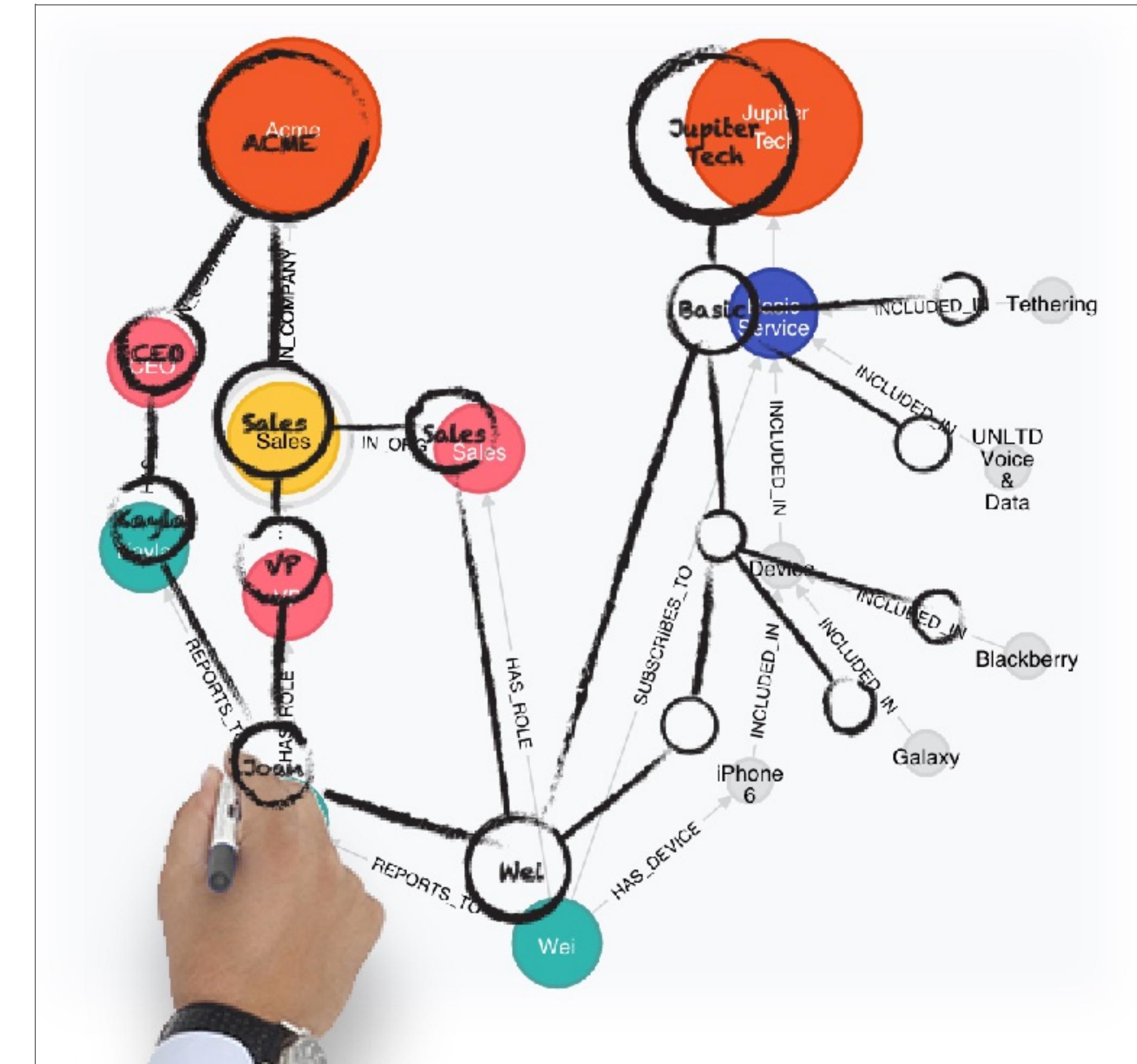
- les gens, les lieux et les évènements
- Les entreprises, les marchés et les clients
- la technologie, les réseaux, les machines, les applications et les utilisateurs
- les applications, le code, les dépendances, les architectures ...
- les criminels, les fraudeurs et leurs comportements

Un langage naturel



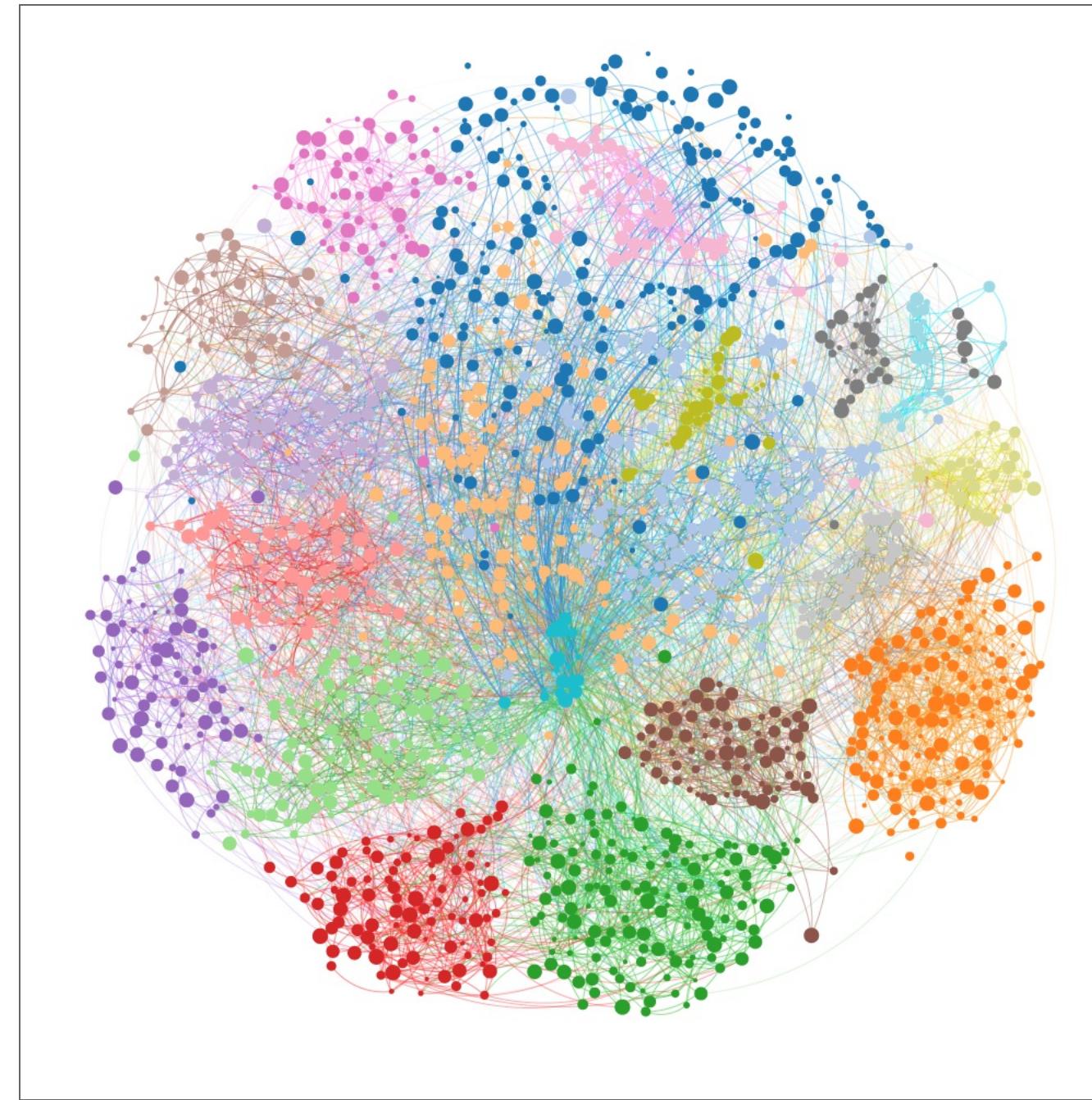
Un modèle unique

La modélisation fonctionnelle et technique sont identiques



Croisez vos données

Les relations entre les données sont plus importantes que la donnée en elle-même.





Quels sont les cas d'utilisations ?

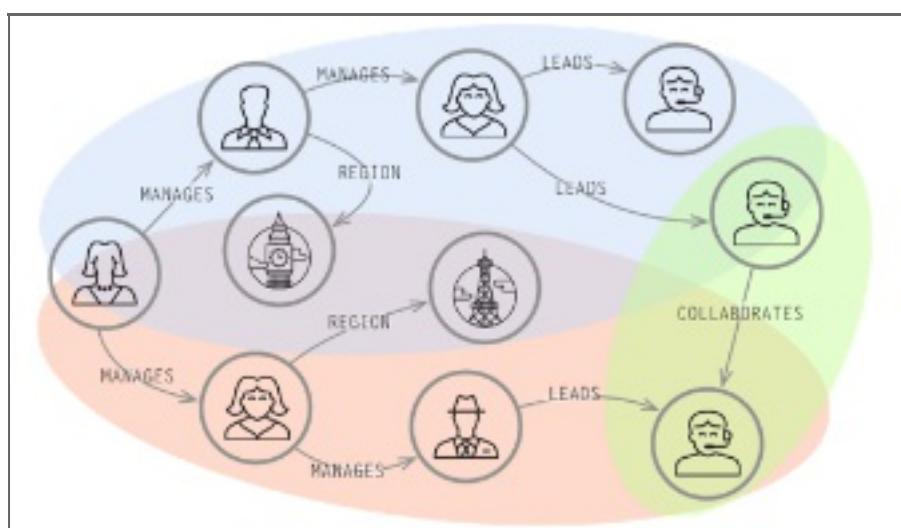
Les cas d'utilisations

En interne

Master Data Management

Analyse des réseaux

Détection de fraude

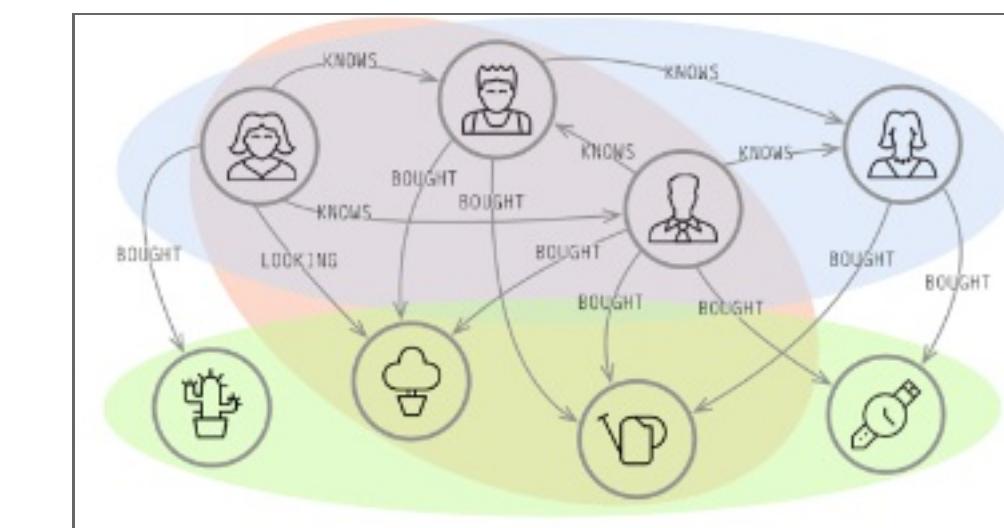


Orienté client

Recommandation

Graph-Based Search

Gestion des accès



La recommandation

La recommandation

Offrir une recommandation en **temps réel, hautement ciblé** et contextuelle
afin quelle soit **pertinente** 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 €

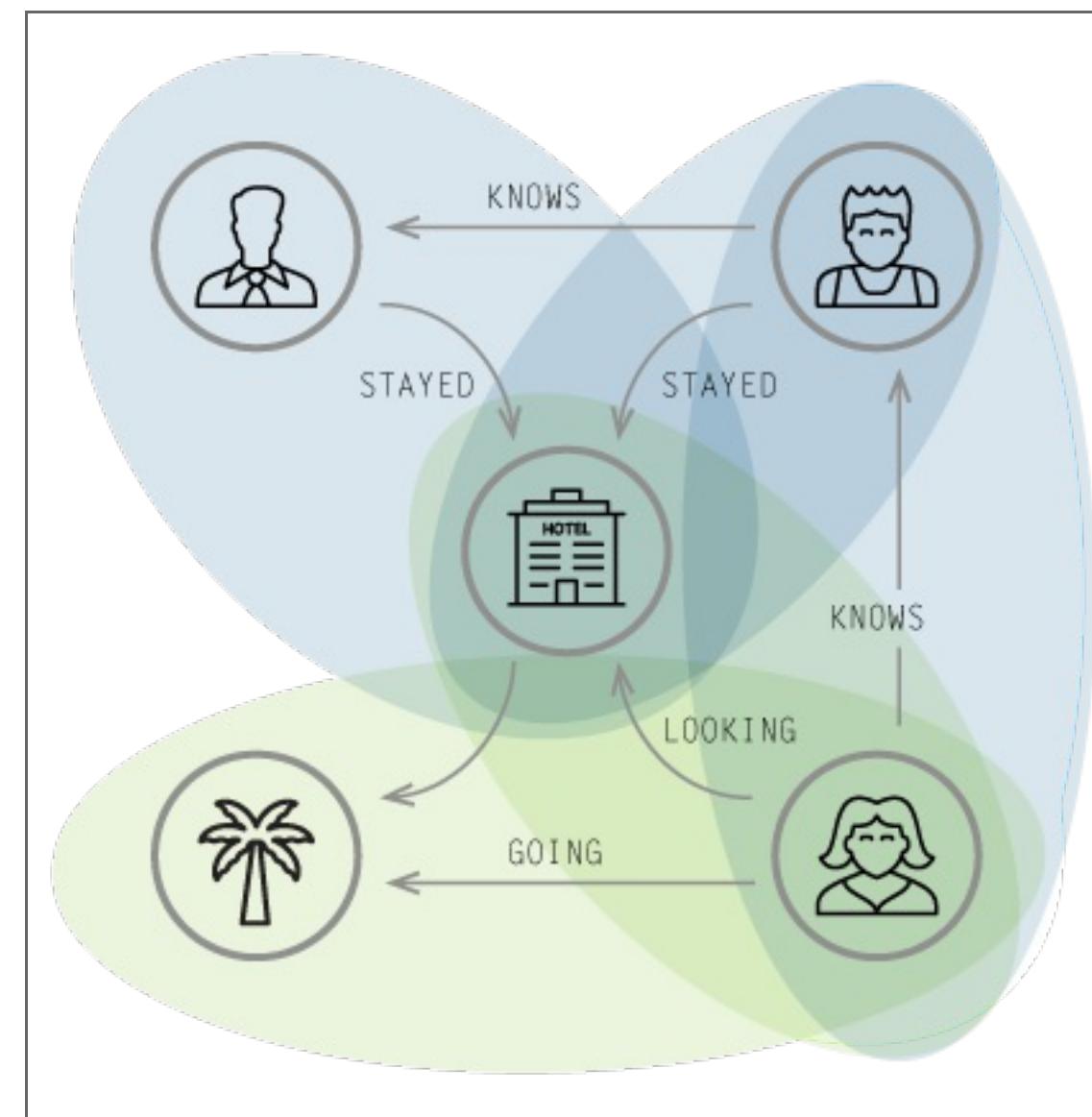
Suggestions · Actualiser · Tout afficher

- linkurious @Linkurious X
Suivre
- Philip Rathle @prathle X
Suivre

Trouver des amis

Pourquoi du graphe ?

Faire des recommandations **efficaces** repose sur une base qui **comprend les relations** existants entre les entités, ainsi que la **qualité et la solidité** de ces connexions.



Les défis

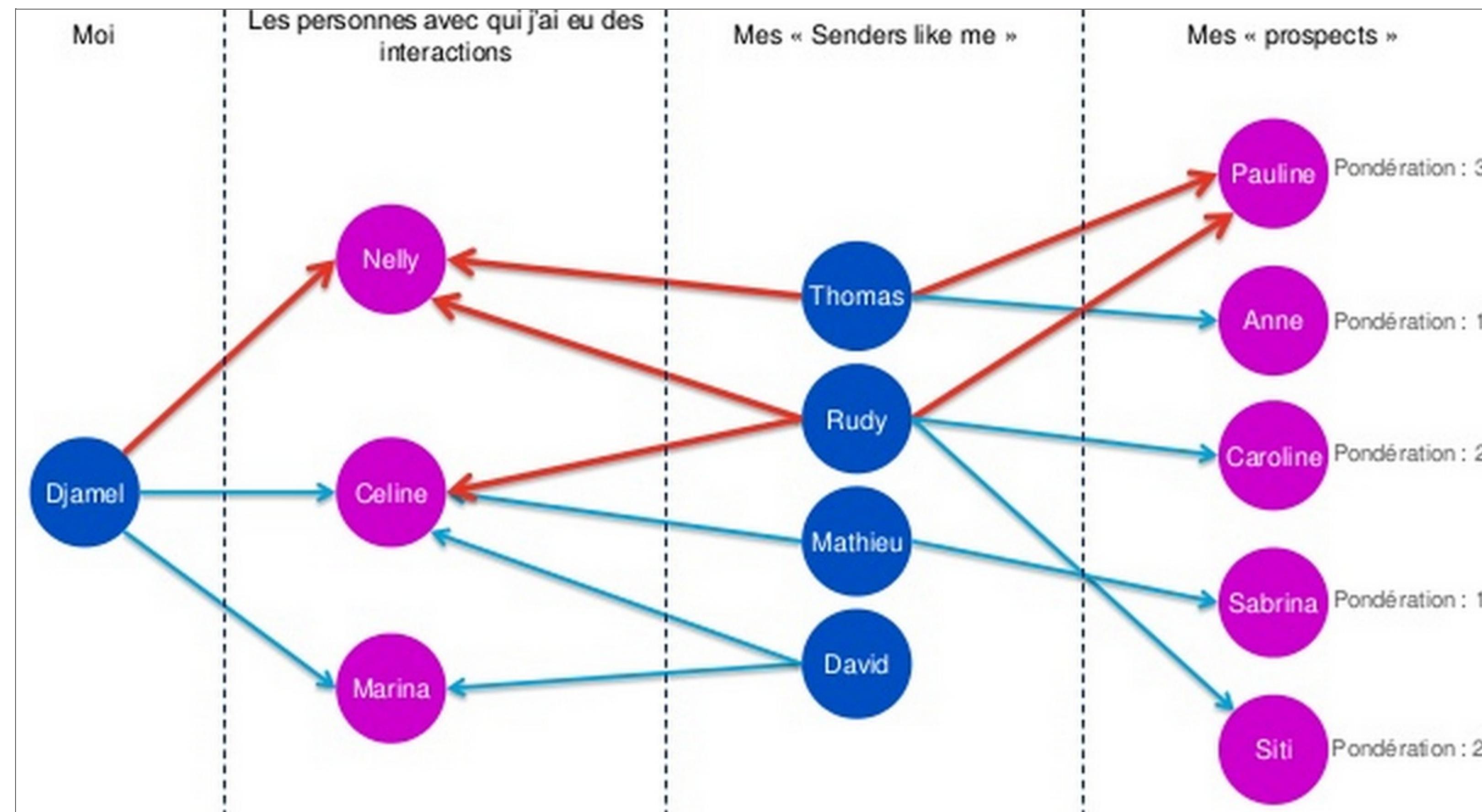


Traiter de **gros volumes** de données et de **relations** pour comprendre le **contexte**

Offrir des recommandations **pertinentes** en **temps réel**, sur des **faits immédiats**

Prendre en compte **continuellement** des **données et relations nouvelles**

L'algorithme du **filtrage collaboratif**



Autres clients



Fournit des recommandations en temps réel sur leur service de vente en ligne, en utilisant les informations relatives aux préférences des clients.



Le géant du commerce électronique eBay a également utilisé avec succès un moteur de suggestion fonctionnant sur un graphe, en l'occurrence pour une solution sophistiquée d'acheminement de paquets/courrier en temps réel (Ebay now).

La détection de fraudes

La détection de fraudes

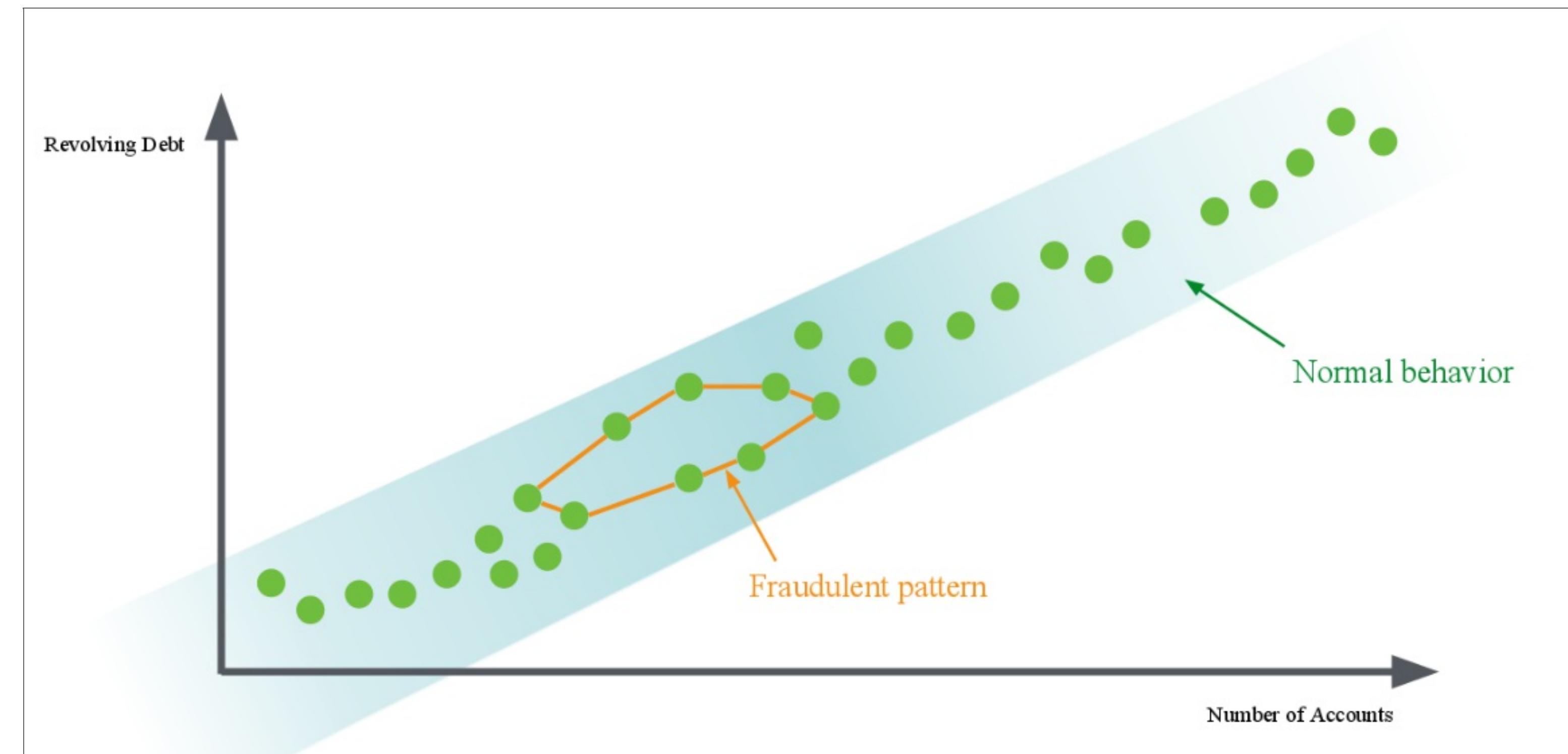
- La fraude coûte des **milliards** chaque année aux banques et assurances
- Les techniques de fraudes sont de plus en plus **sophistiquées**
- Il faut constater la fraude en **temps réel** (ie. avant quelle est lieu)



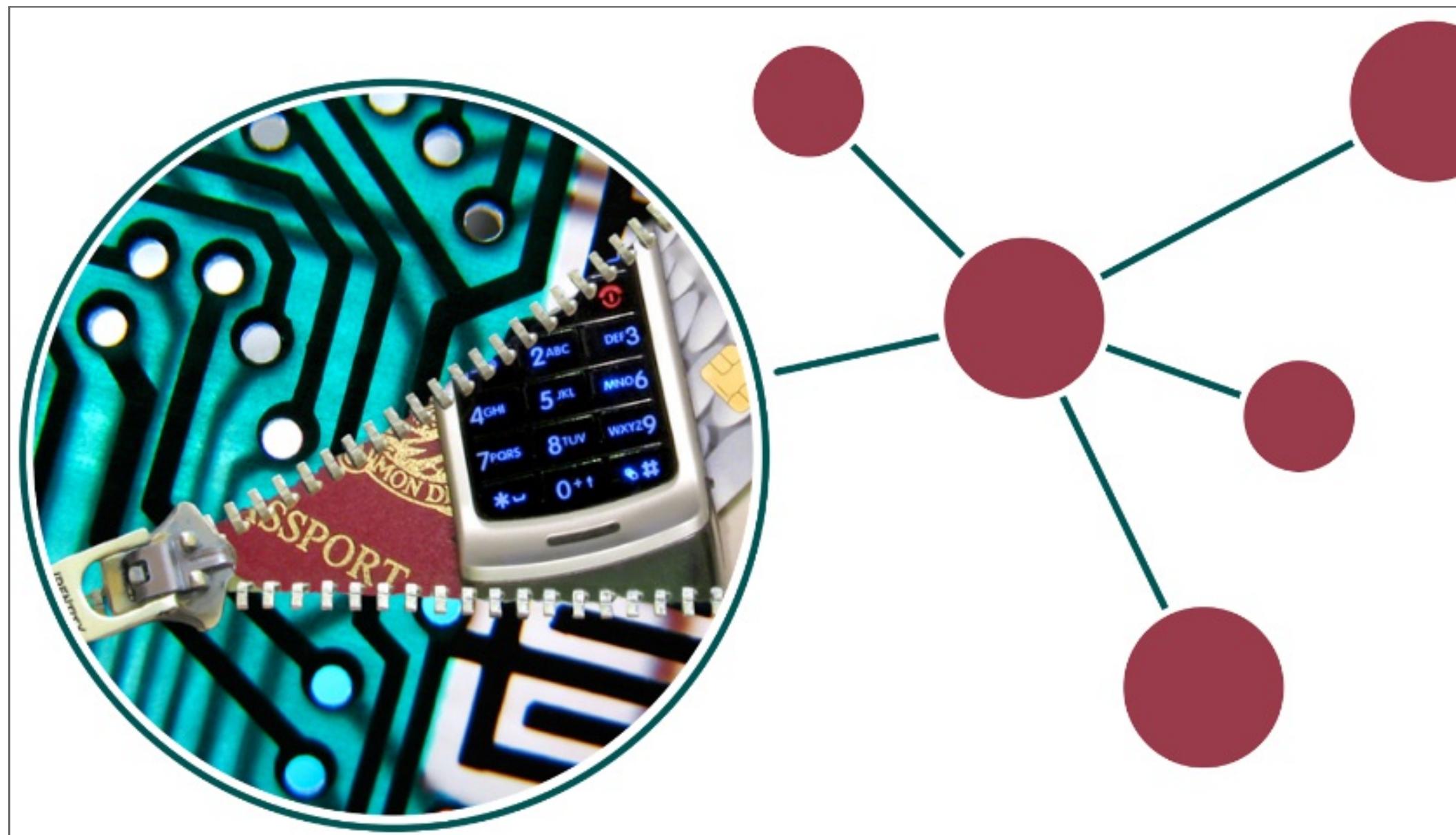
Pourquoi du graphe ?

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.**



Les défis



Effectuer une **analyse complexe des liens**
pour découvrir les **schémas et les réseaux de fraude**

Déetecter et empêcher les fraudes **en temps réel**

Réseaux de fraude en **évolution et dynamiques**

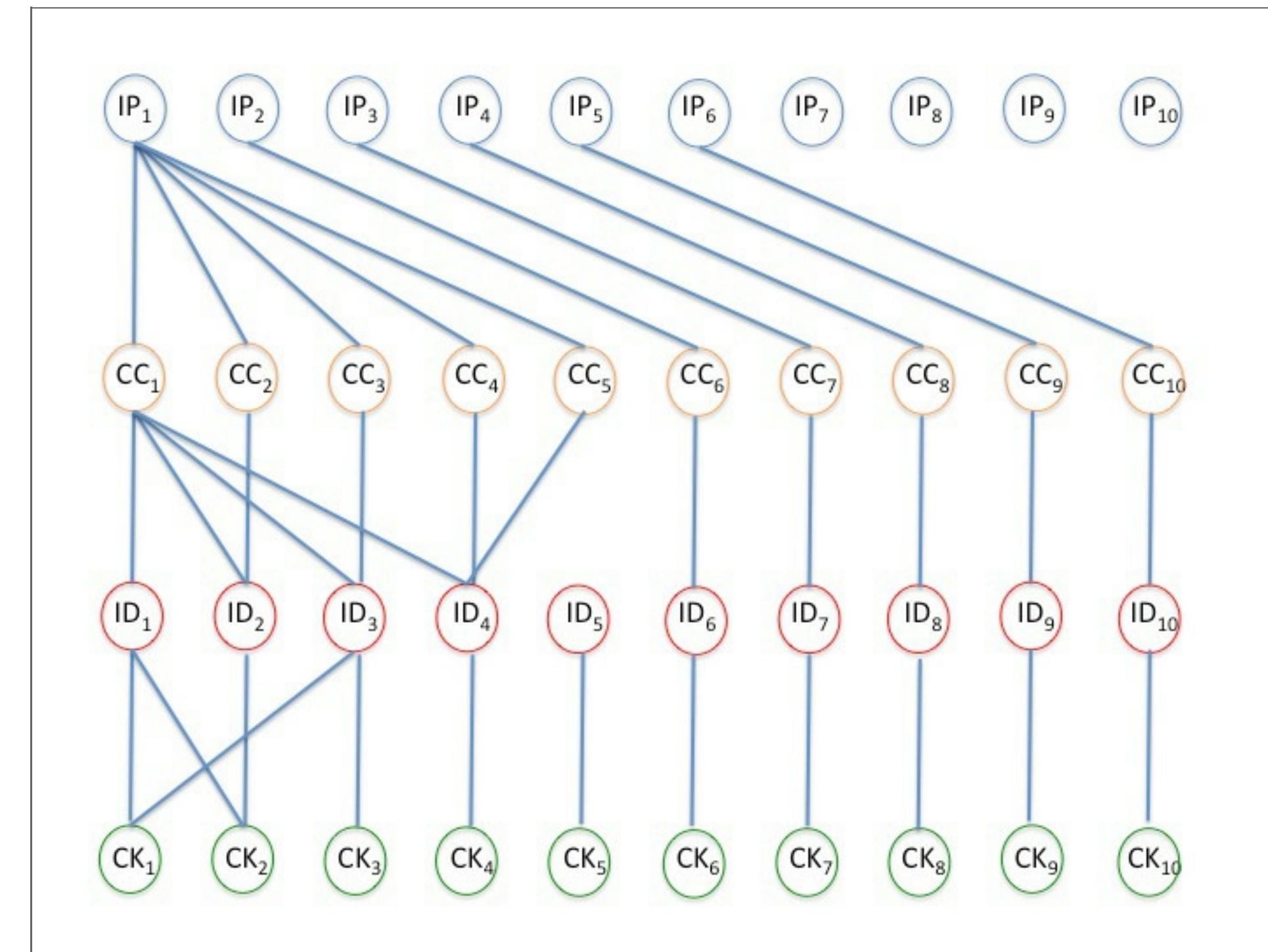
Les Paradise Papers



<https://projects.icij.org/paradise-papers/the-influencers/#/>

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)



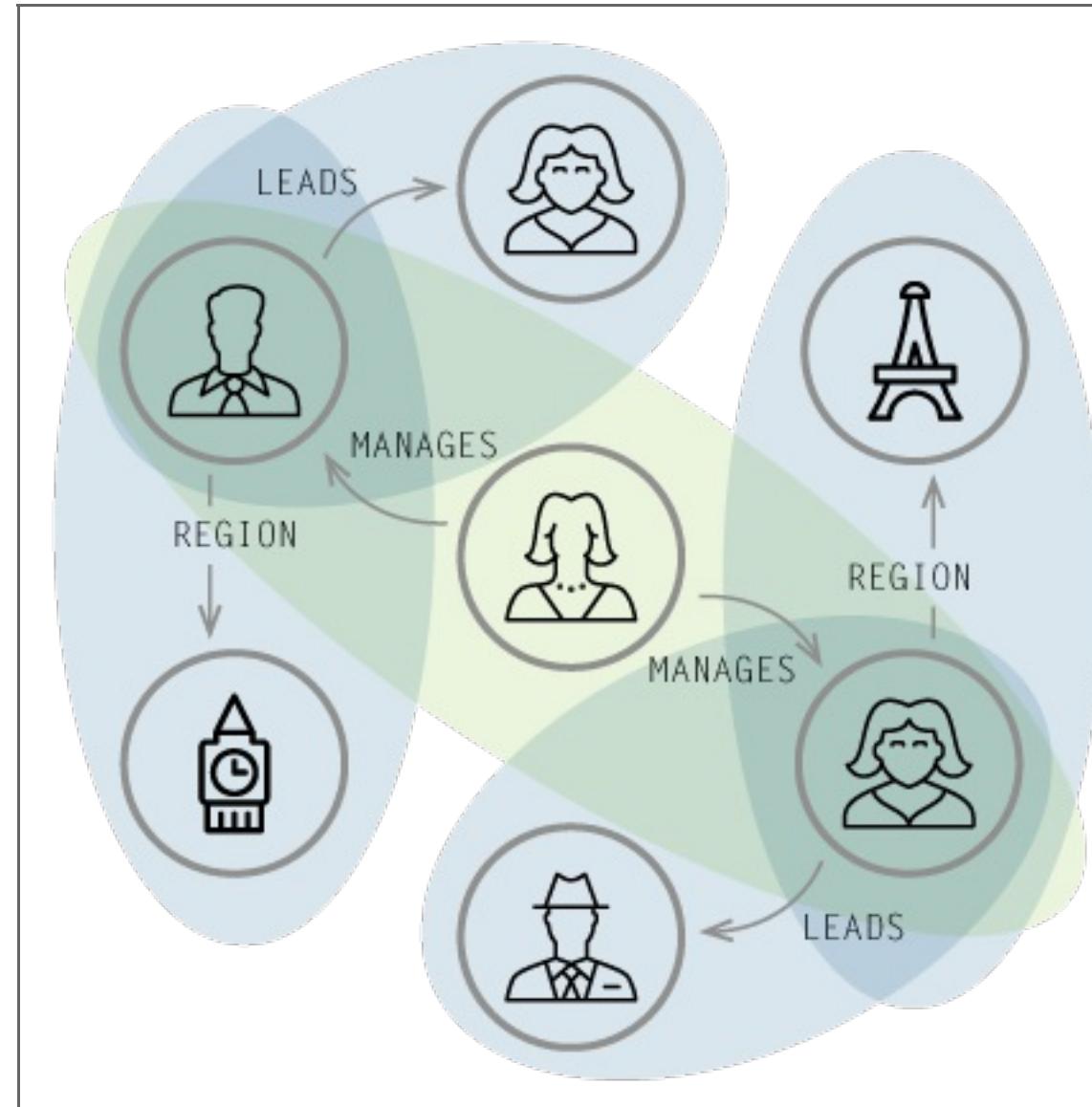
La gestion de données de référence

La gestion de données de référence

Les données de référence constituent le **pilier de votre entreprise** et incluent des données telles que les utilisateurs, les clients, les produits, les comptes, les partenaires, les sites, les unités commerciales, ...

- stockées à différents emplacements (**effet silo**)
- avec des **chevauchements** et des **redondances**
- et des **qualités variables**

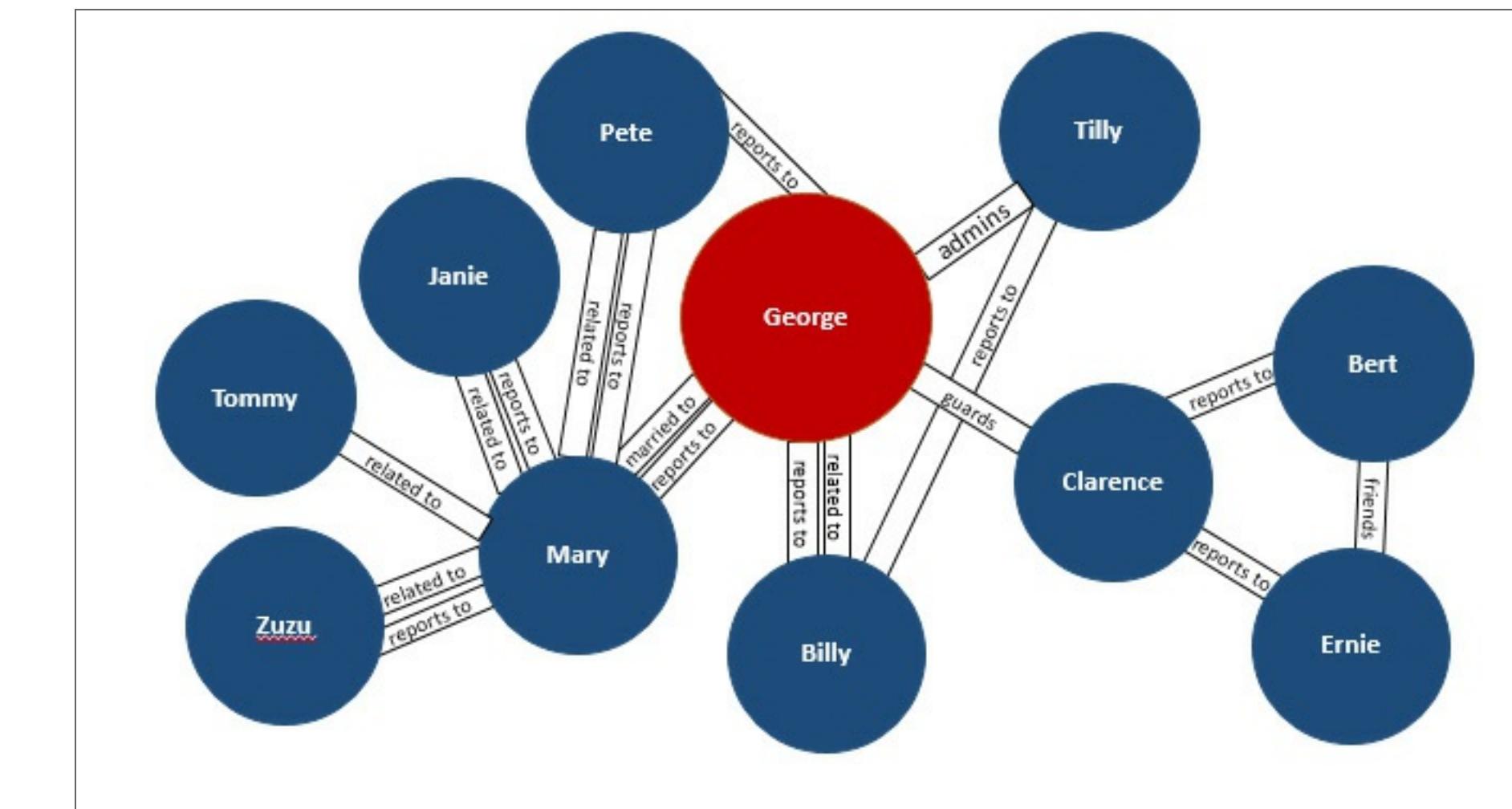
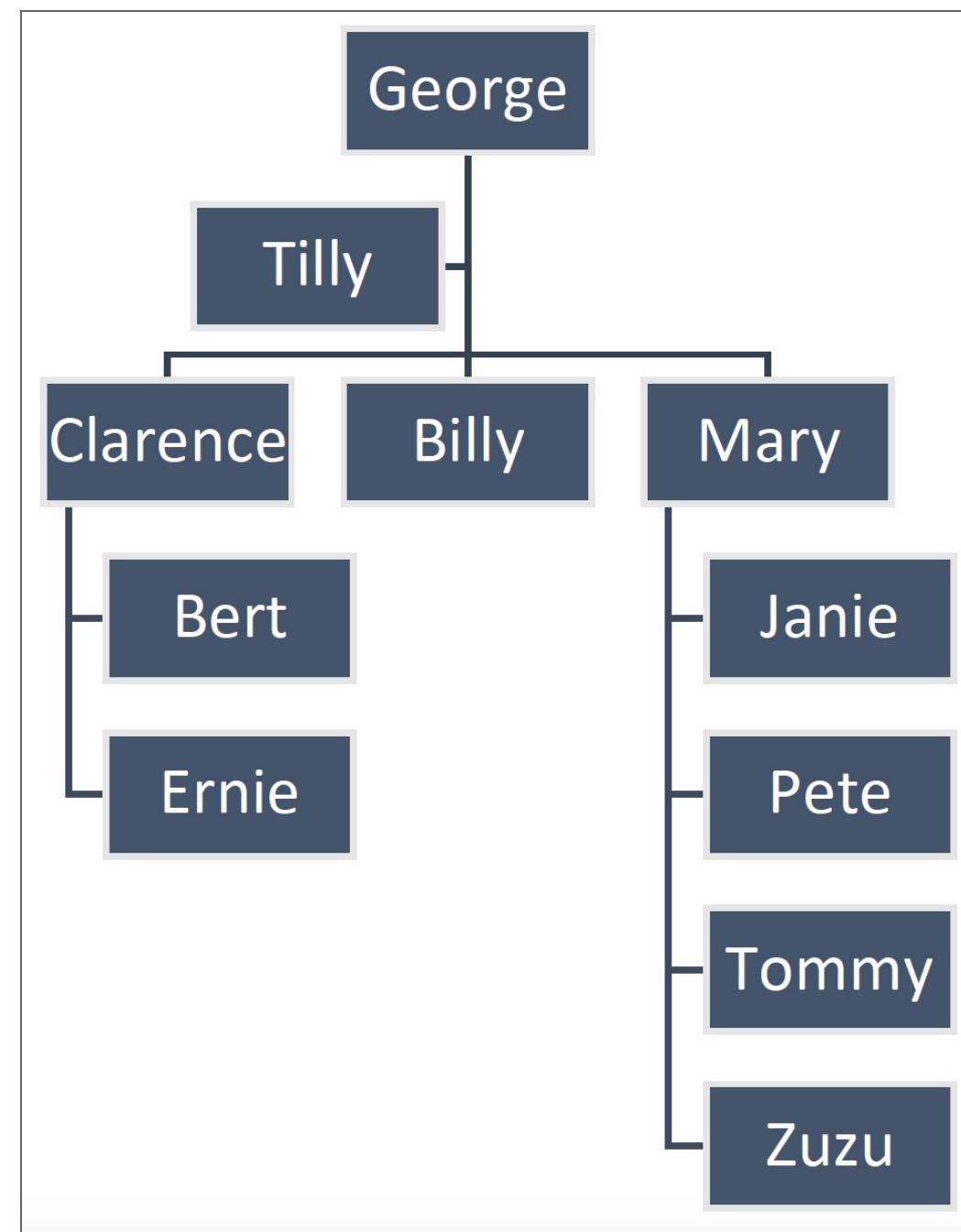
Pourquoi du graphe ?



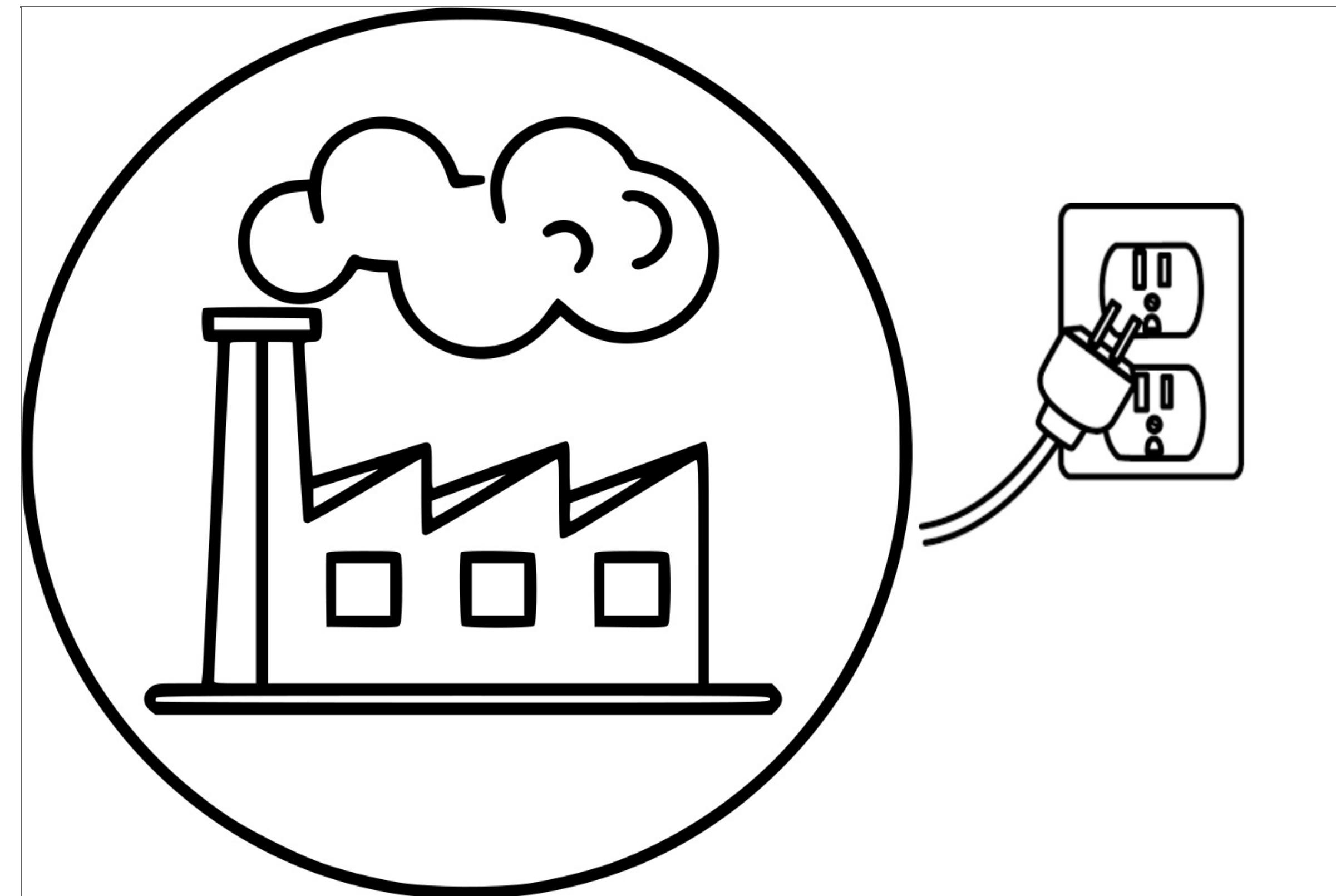
- Les données sont **hautement connectées**
- Les systèmes de gestion traditionnels ne sont pas **flexibles**
- L'**interconnexion des données** de références permettent d'obtenir un **avantage concurrentiel**

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**.



L'entreprise connectée



Client



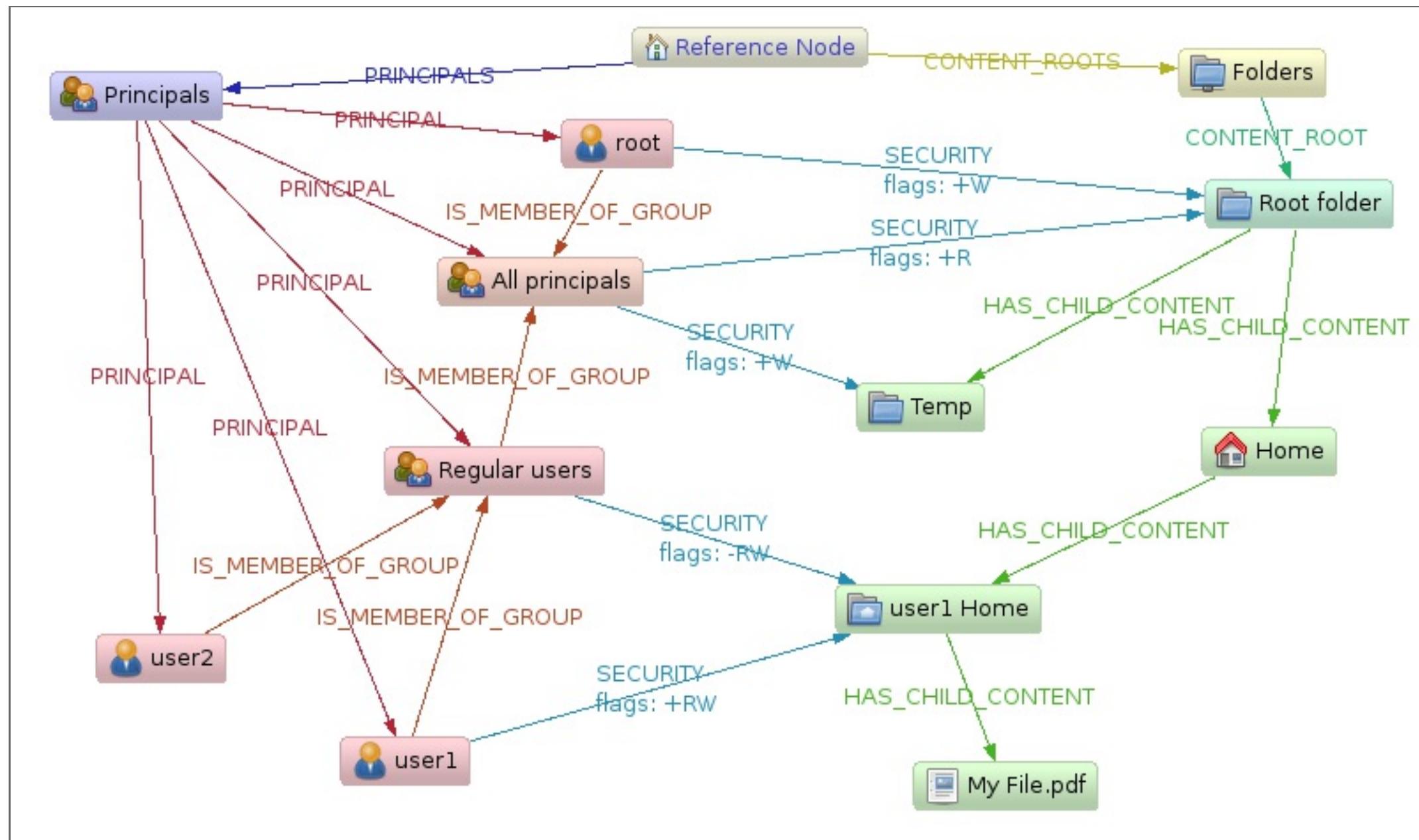
Adidas utilise Neo4j pour créer leur référentiel interne, incluant les données sur les produits, les marchés, les médias sociaux, les actifs numériques, ...



Wine Data System utilise Neo4j pour réconcilier les bases de données des négociants afin de construire un référentiel unique.

La gestion des identités et des accès

C'est un graphe !

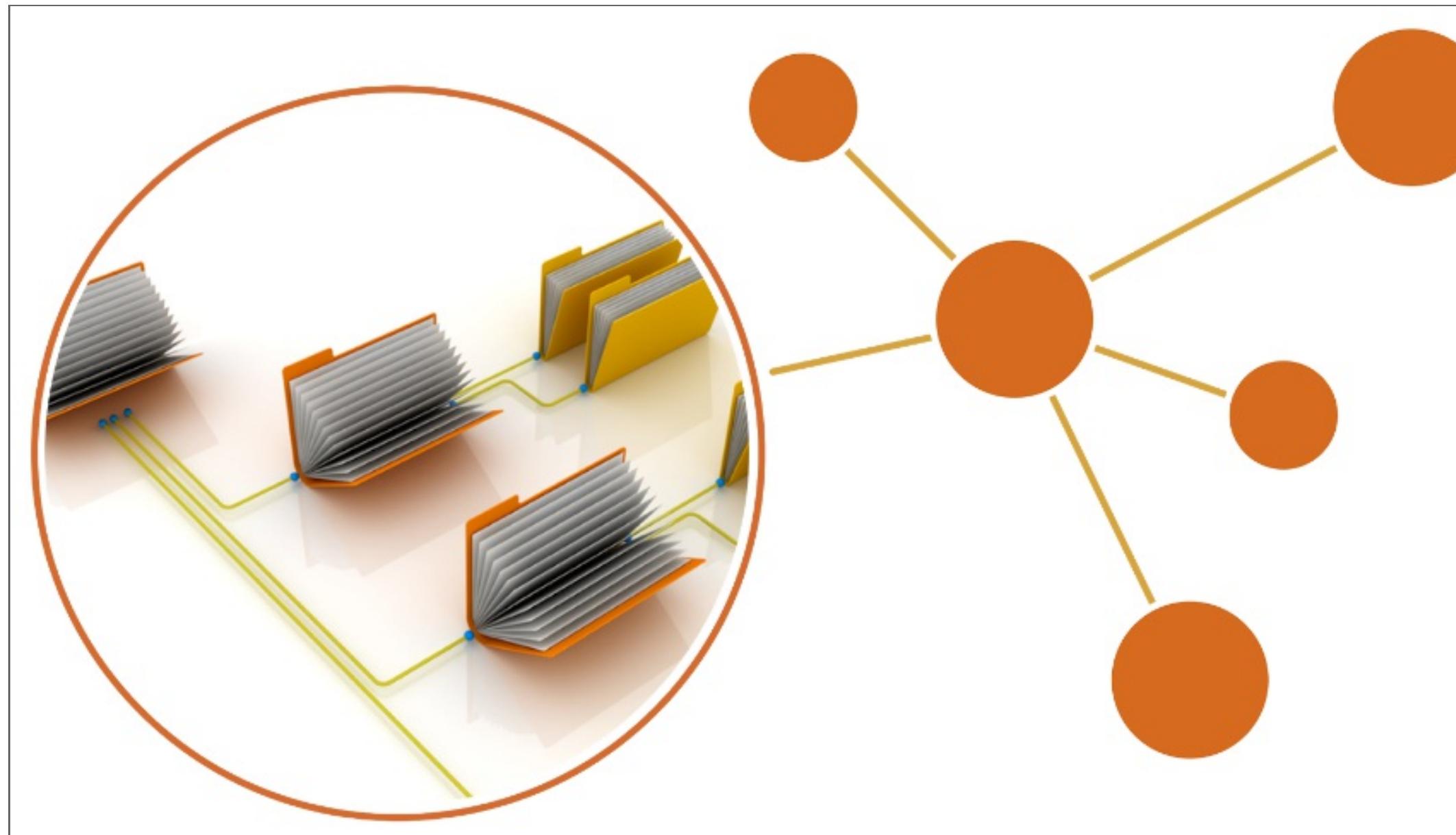


Des **personnes** qui appartiennent à des **groupes**, qui eux-mêmes peuvent appartenir à des **groupes**.

Les groupes et utilisateurs possèdent des **rôles**.

Est-ce que l'utilisateur **Benoit** a accès en écriture au répertoire **\samba\clients** ?

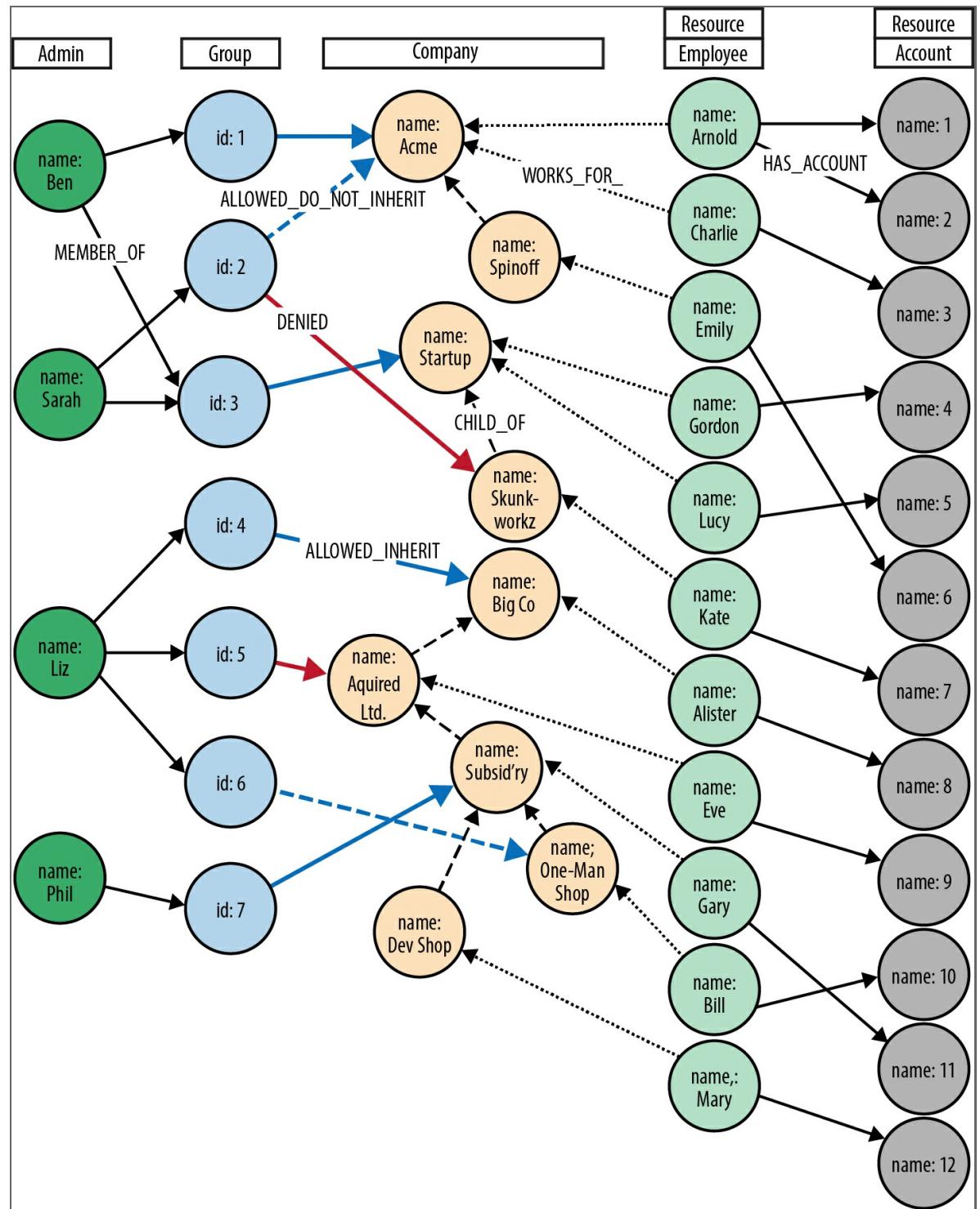
Les défis



Les identités et autorisations d'accès sont
hautement interconnectées

Garder la **cohérence des données** et
faciliter sa **compréhension**

Retrouver un accès en **quelques**
millisecondes pour satisfaire les clients



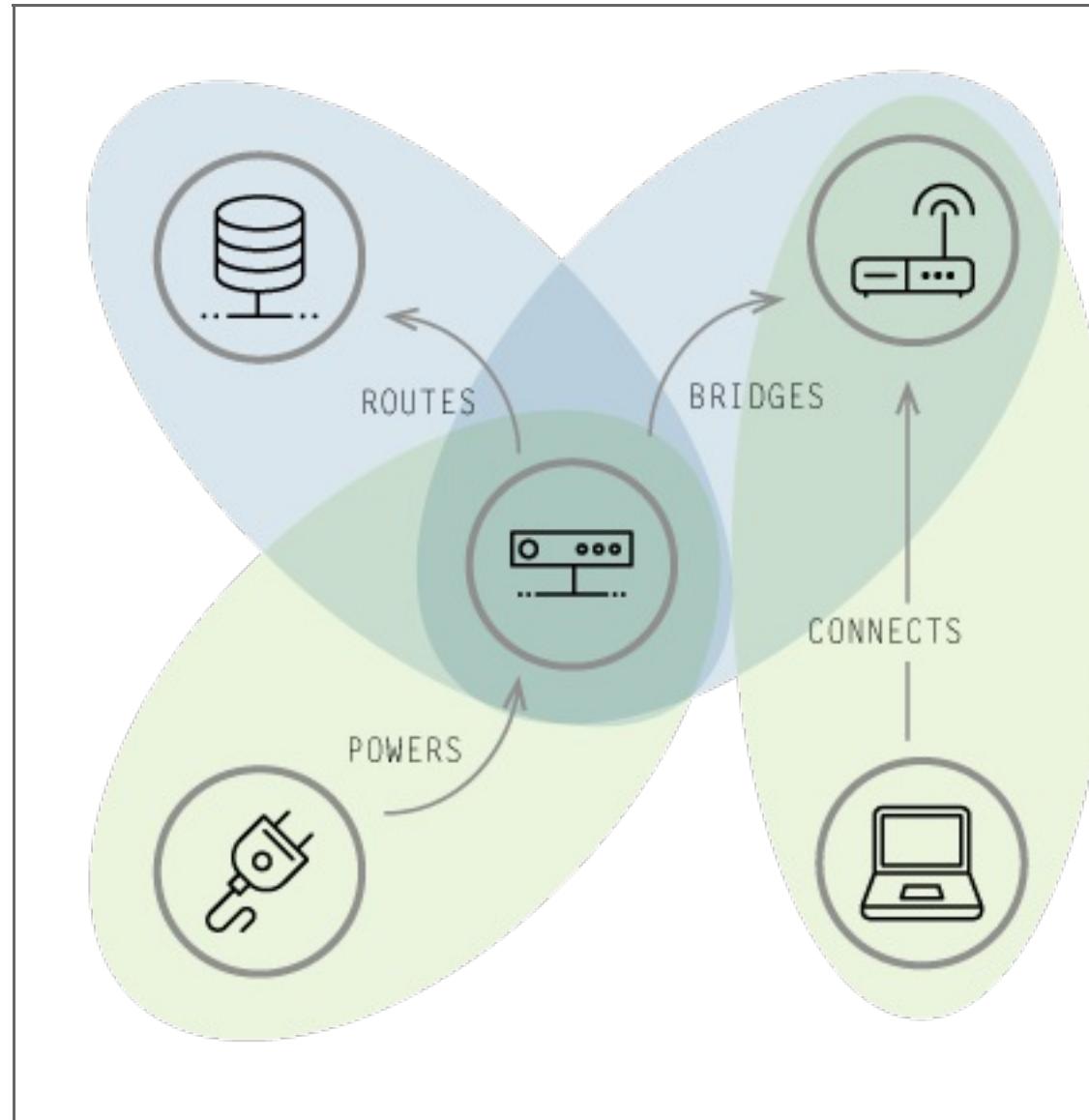
Telenor permet à ses clients de gérer les services de leurs employés.

Les administrateurs ne voient que les parties de l'entreprise et les services qu'ils ont le droit de gérer.

- **5 millions** de clients, souscriptions et accords
- **dépendances complexes** entre les groupes, sociétés, individus, comptes, produits, souscriptions, services et accords

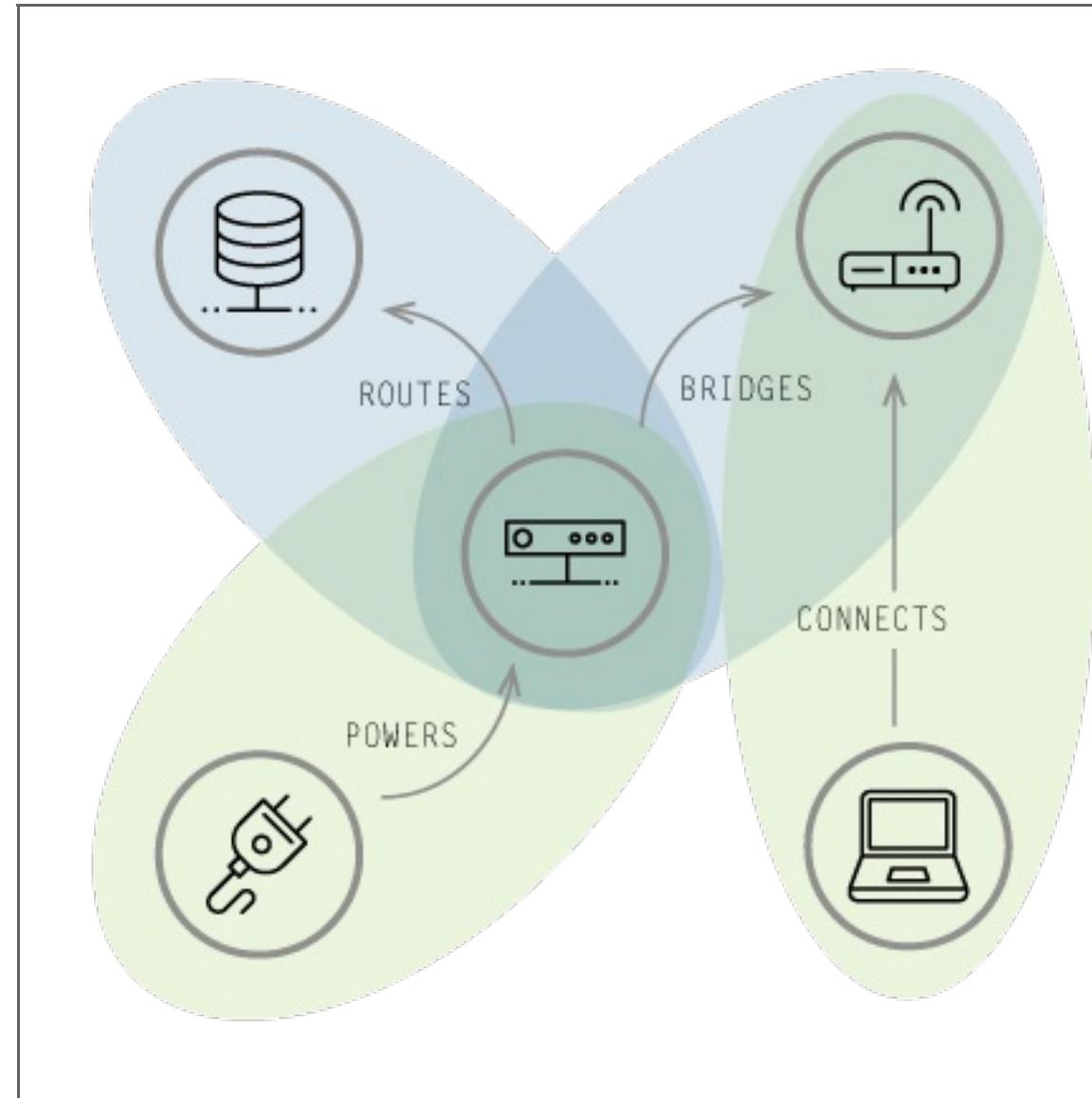
Les réseaux informatiques

Un réseau est un graphe !



- Modélisation et stockage naturel
- Meilleur visualisation et appréhension des données
- Permet de réaliser des analyses d'impacts en cas de panne
- A l'inverse, voir les redondances du réseau

Les défis



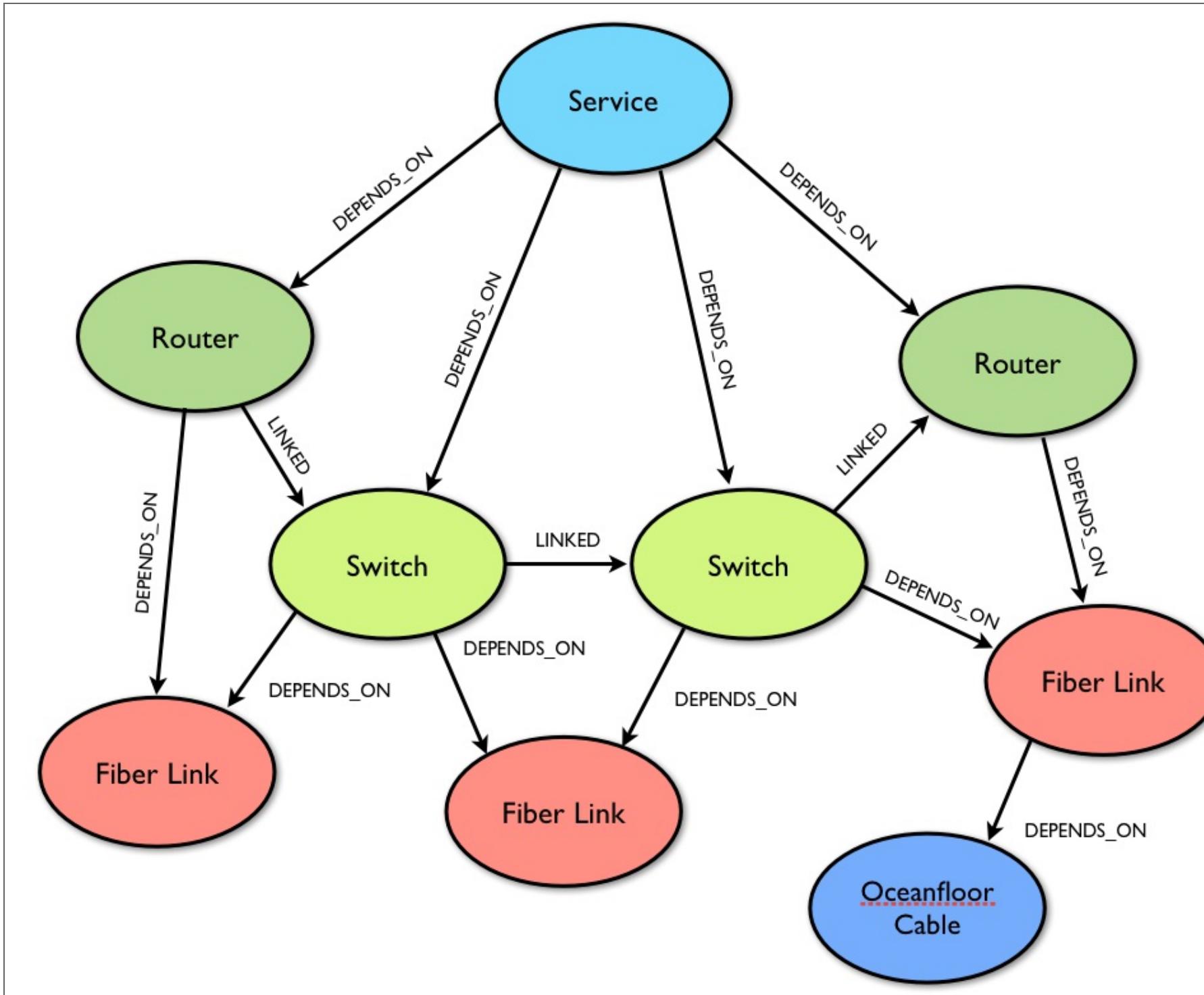
Résolution des problèmes sur un réseau

: trouver la source d'un dysfonctionnement
le plus rapidement possible

Analyse d'impact : vérifier la résilience du
réseau

**Nombre croissant de noeuds physiques
et virtuels**

Analyse d'impact chez SFR



Identification **temps réel** des clients liées à une panne, ou une maintenance.

Gestion d'inventaire réseaux **très flexible**.

Une seule source de vérité (Neo4j) représentant le réseau entier.

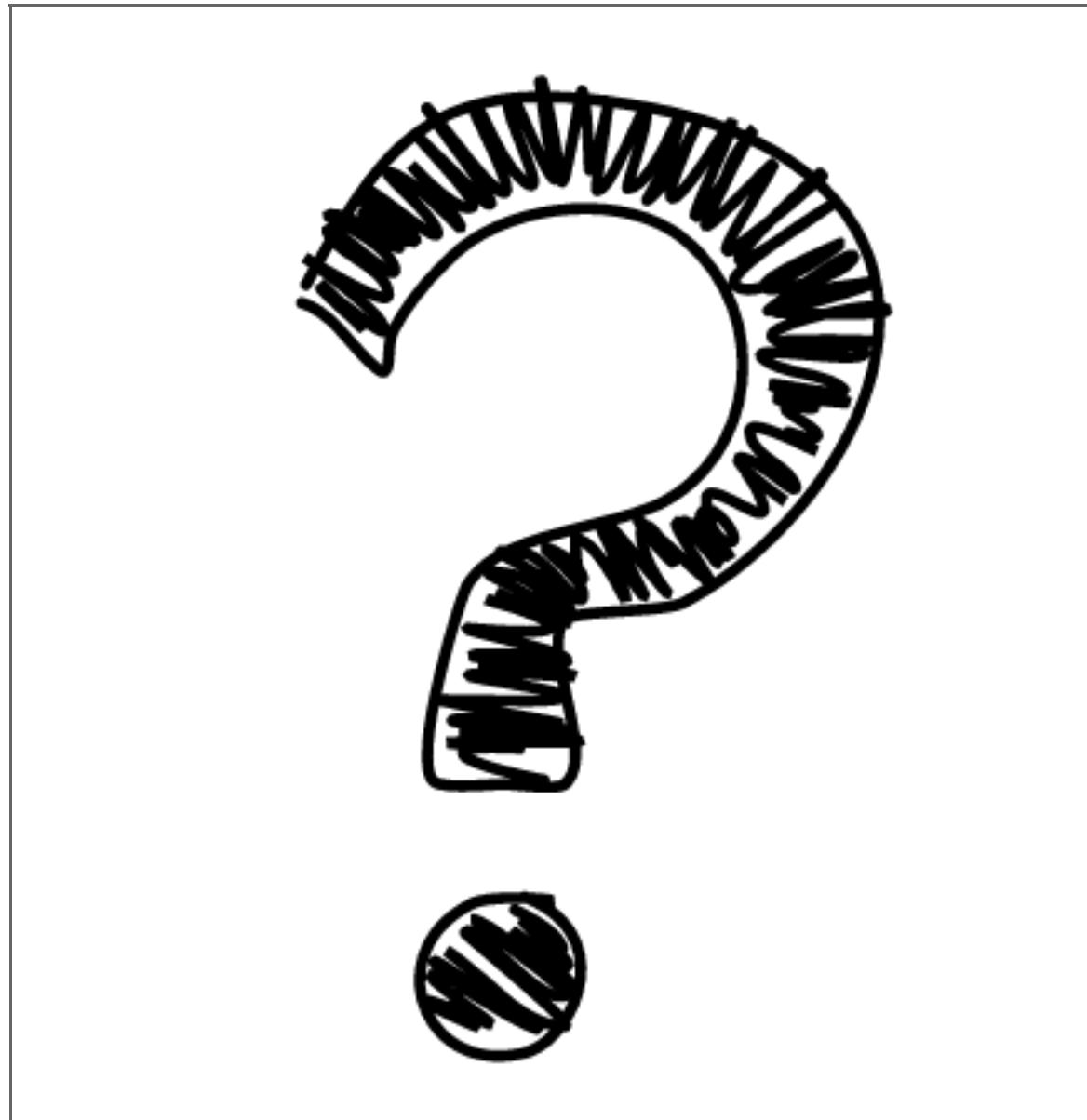
Système dynamique de chargement (ETL) des données issues de **+30 sources variées**.

Identifier les **faiblesses du réseau**



Graphs are everywhere !

Des questions ?



Suivez nous sur *twitter* **@neojfr & @neo4j**
Posez vos questions par email sur les *google groups* **Neo4jfr & Neo4j**, ou sur
stackoverflow avec les tags **neo4j & cypher**

Parlez avec nous sur **Slack**