

Des datas & des graphes

Benoit Simard (@logisima)

Presentation

Benoit Simard



- **Graphe addicte**
- Libérateur de données
- benoit@neo4j.com
- @logisima

L'IMA et moi

Mon parcours

- Ancien de l'IMA (2007)
- Vacataire
 - En technologies web de 2010 à 2013
 - En algorithmique et Java de 2016 à 2017

Ce que j'ai appris

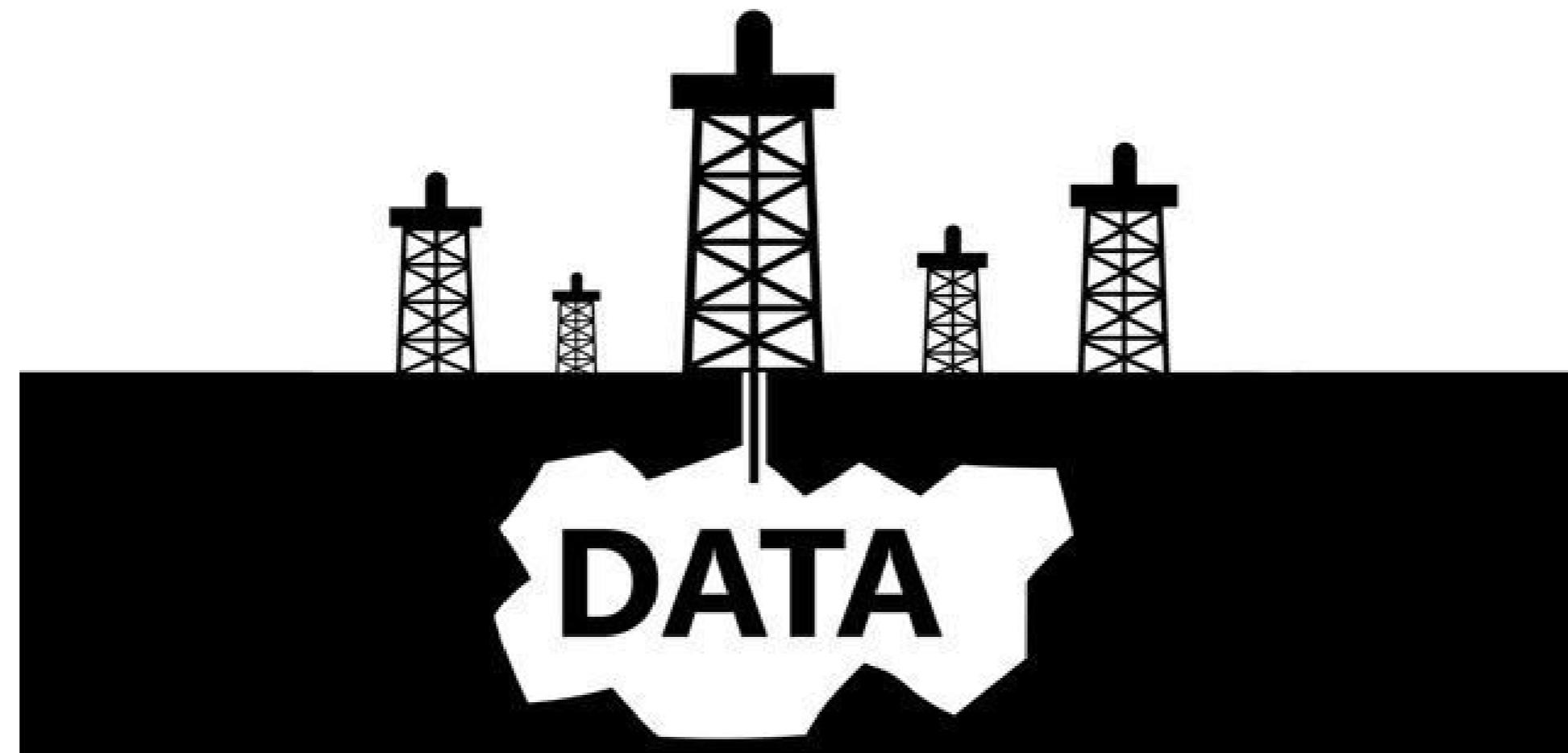
- Statistiques
- L'informatique
- Les graphes
- La comptabilité



The background of the slide features a network graph composed of numerous small, semi-transparent grey circles of varying sizes, connected by thin grey lines. This visual metaphor represents data points and their relationships.

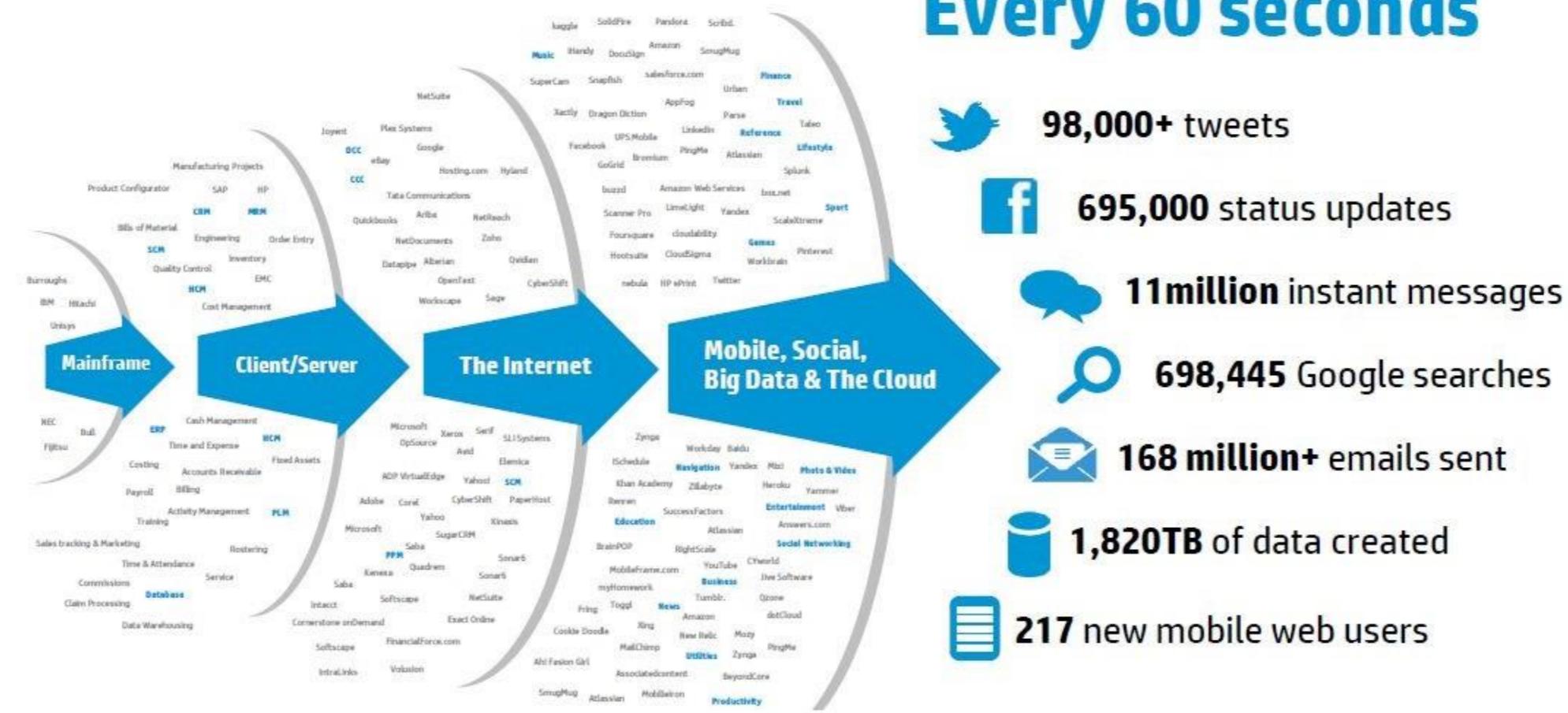
La data

Le nouveau pétrole



Une question de taille ?

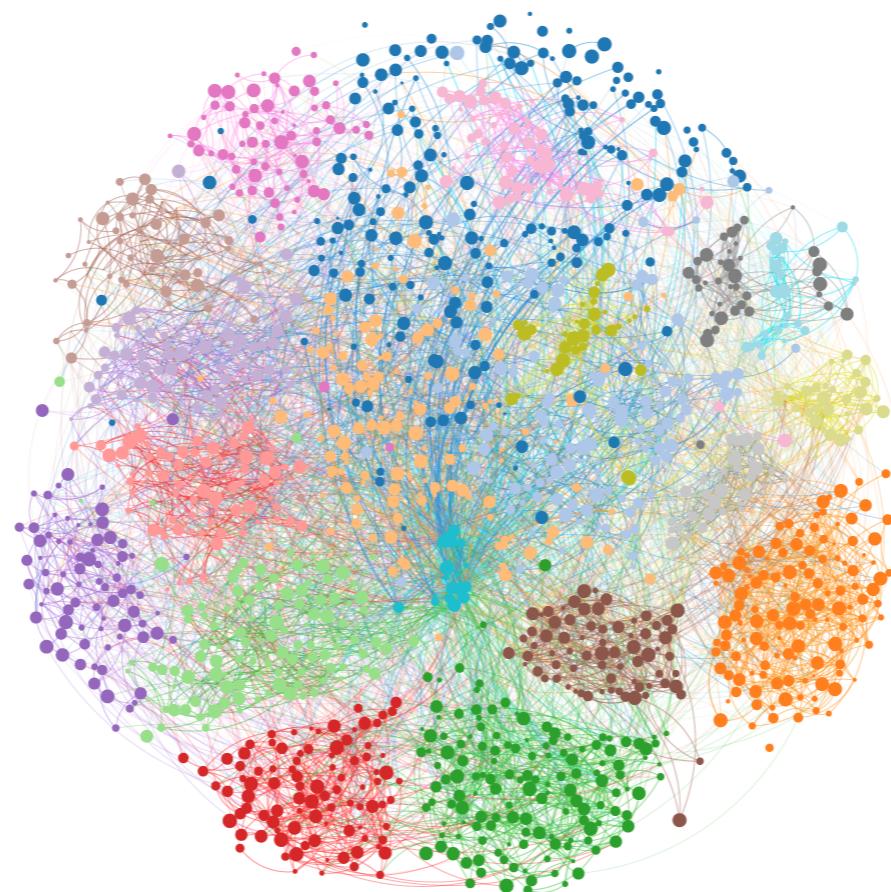
Tous les deux ans, on double la quantité de données produites



Le Big Data

Linked data

La valeur n'est pas dans la donnée, mais dans ses relations

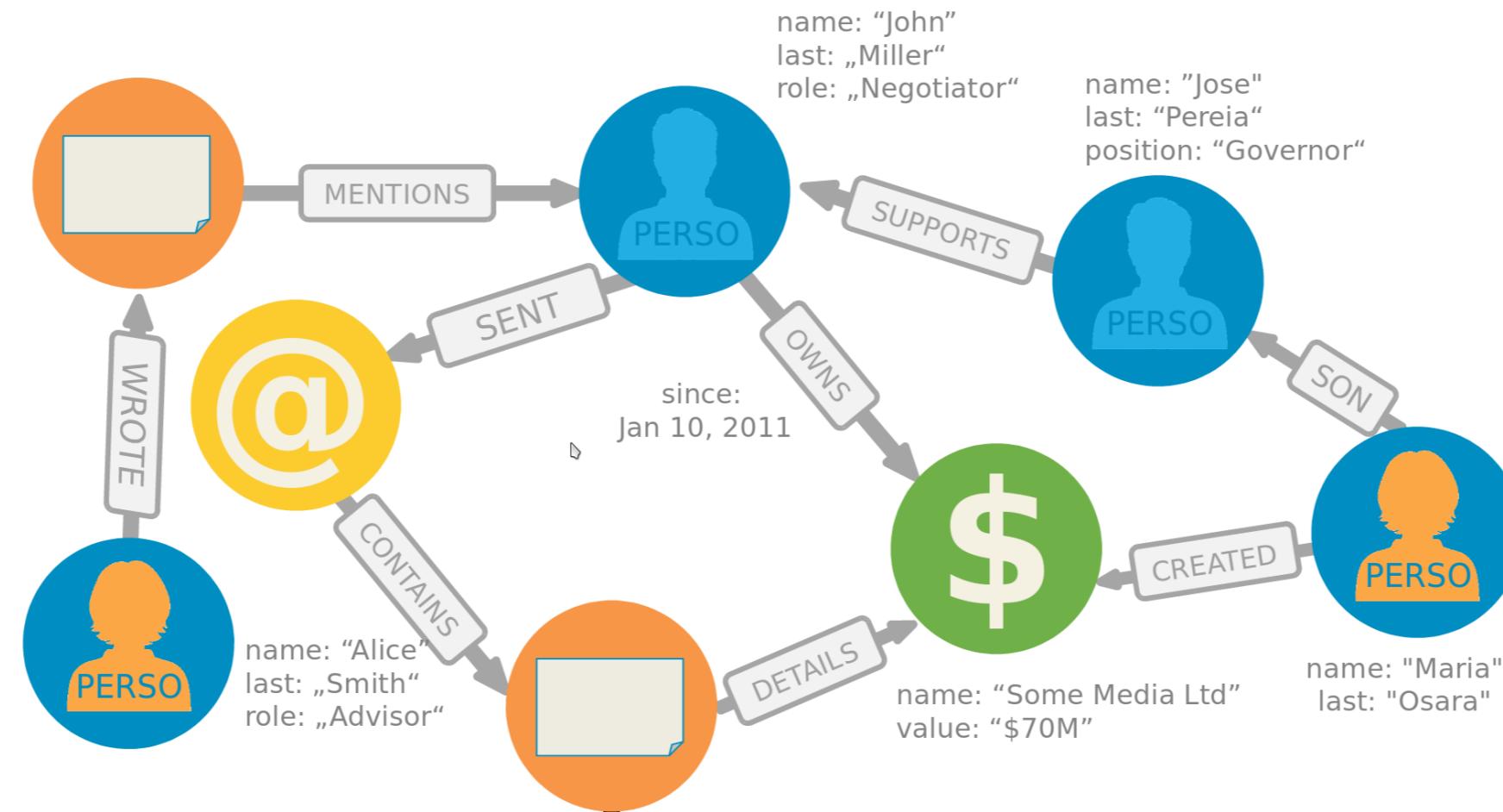


Le processing

complexity = $f(\text{size}, \text{connectivité})$



Les bases graphes



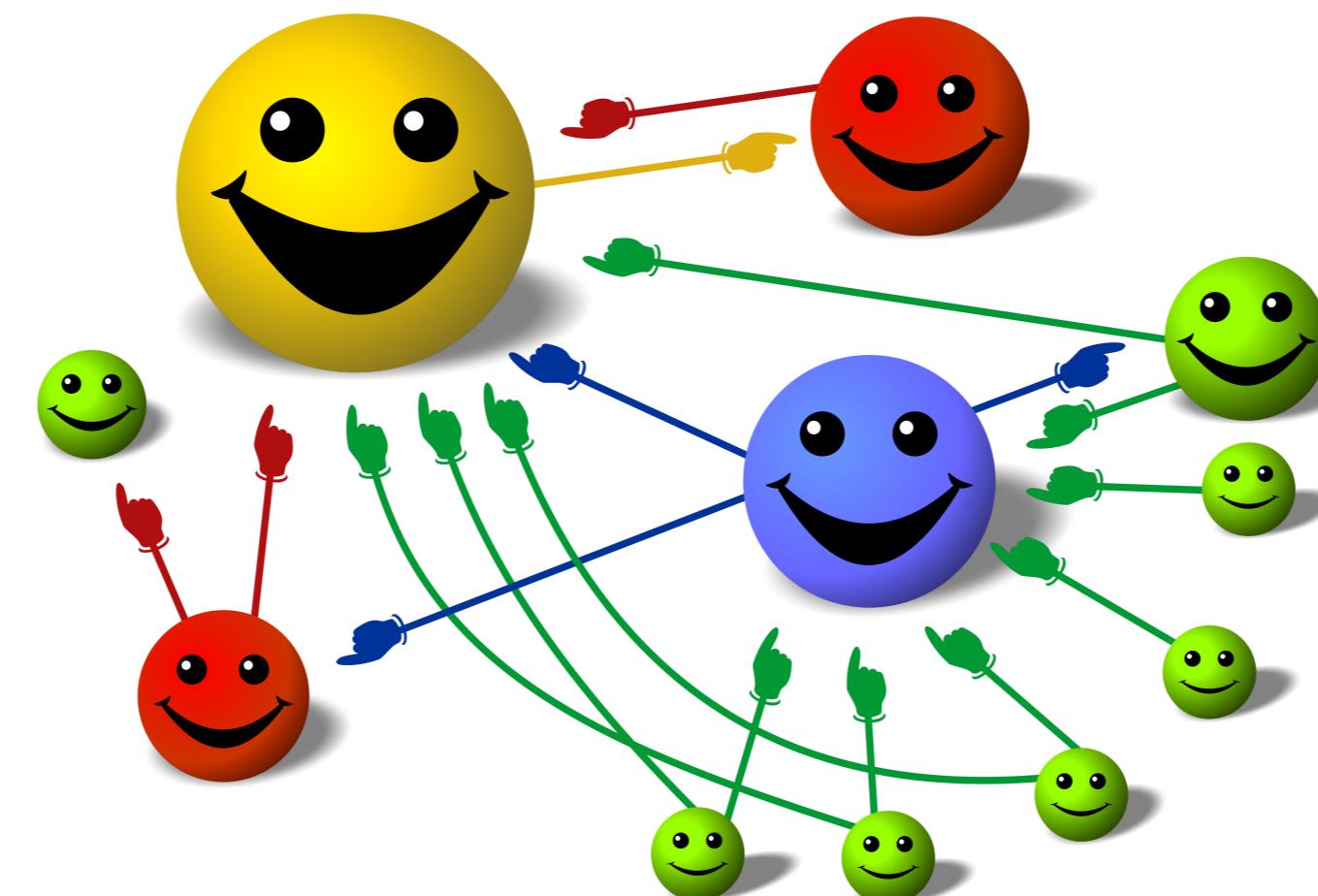


A quoi cela peut servir ?



Google est devenu le premier moteur de recherche grâce à son algorithme basé sur les graphes :
le page rank

Cartographie thématique du web



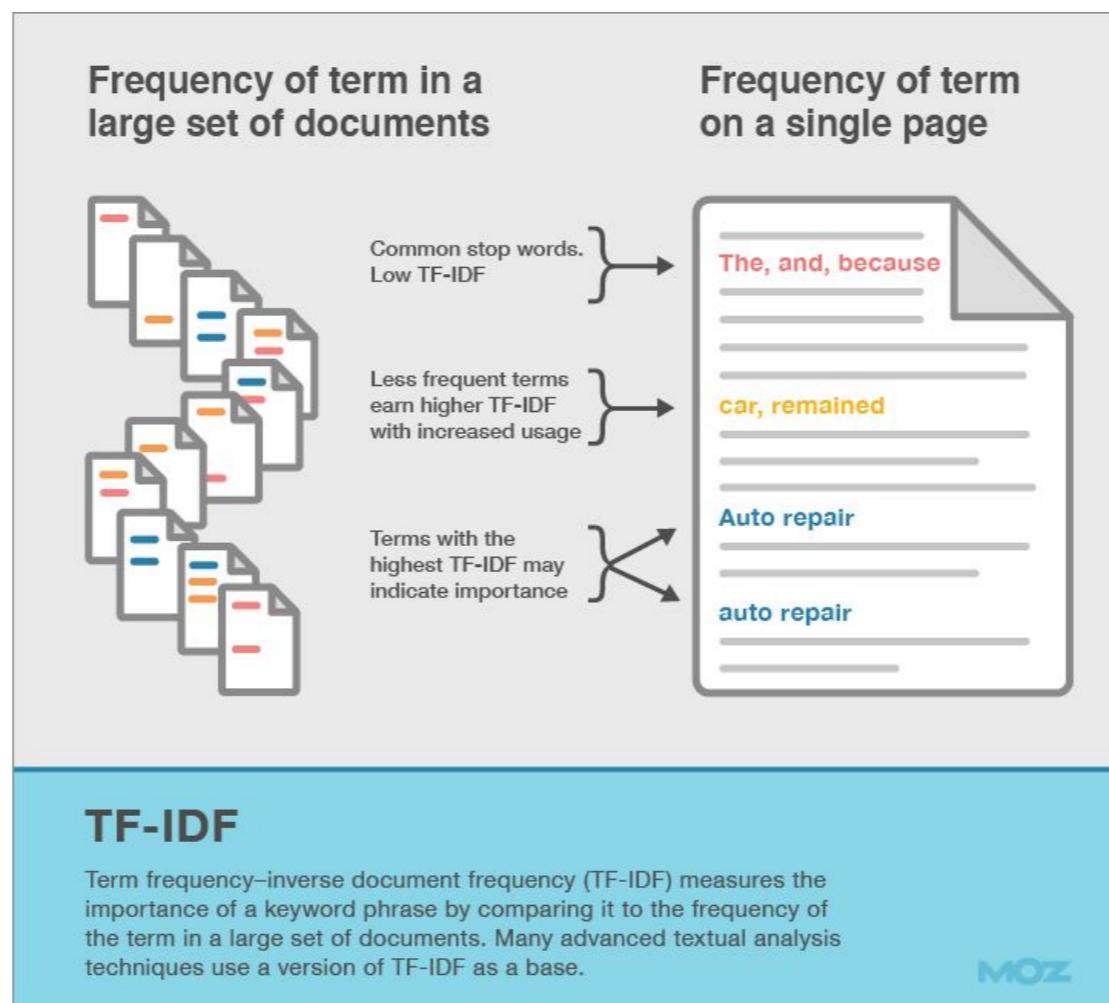
Google™
YAHOO!®
LYCOS



TF / IDF

Terme frequency : Nombre de fois que le mot apparaît dans un document

Inverse Document Frequency : Nombre de fois que le mot apparaît dans un ensemble de documents



Invisible touch



Recommendation

Moteur de recommandation en temps réel

La recommandation est partout : des réseaux sociaux, à la vente, la publicité, les médias, ...

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

LES CLIENTS ONT ÉGALEMENT ACHETÉ...



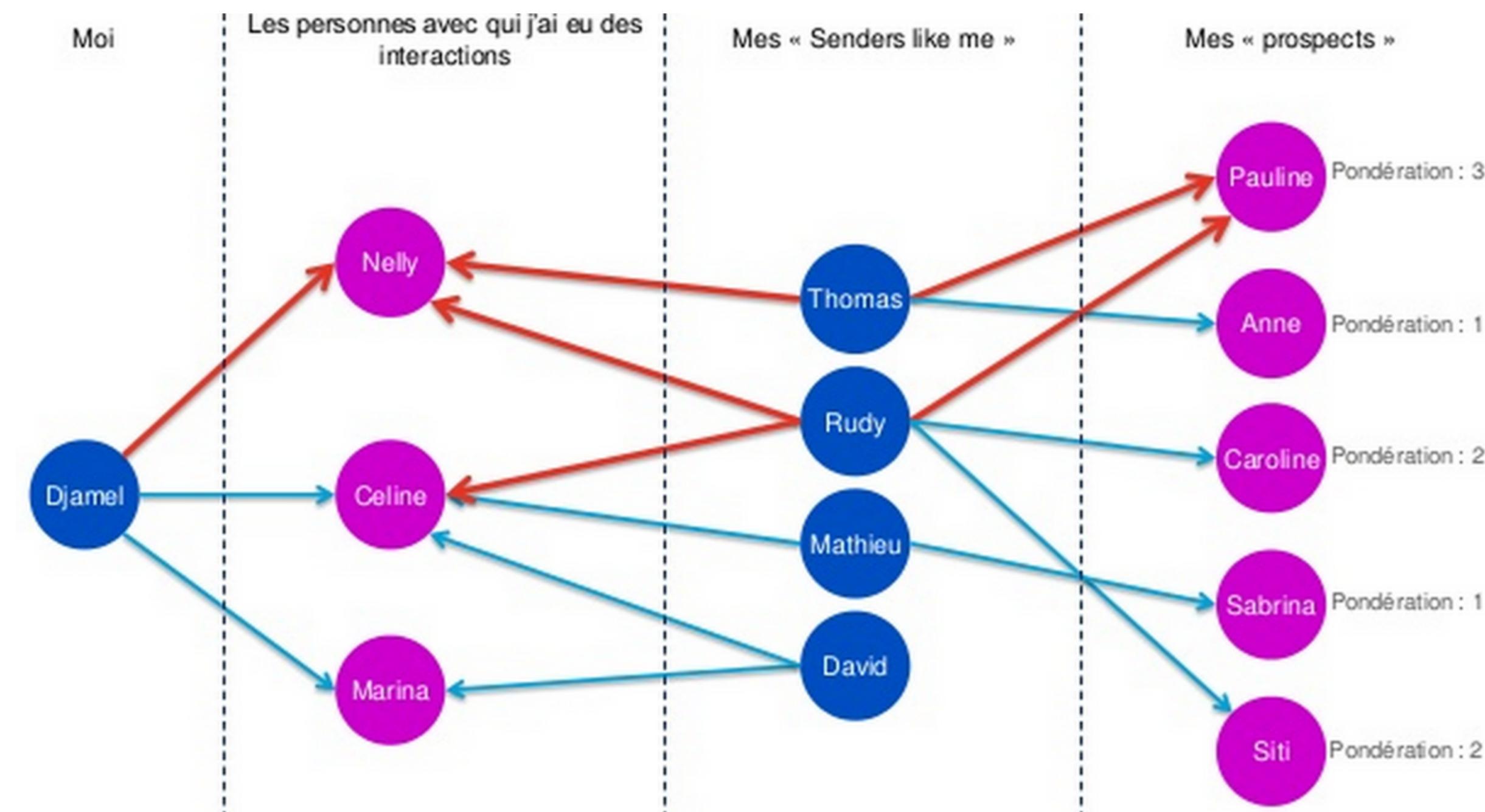
Suggestions · Actualiser · Tout afficher

Filtrage collaboratif

Les amis de mes amis avec lesquels je ne suis pas ami

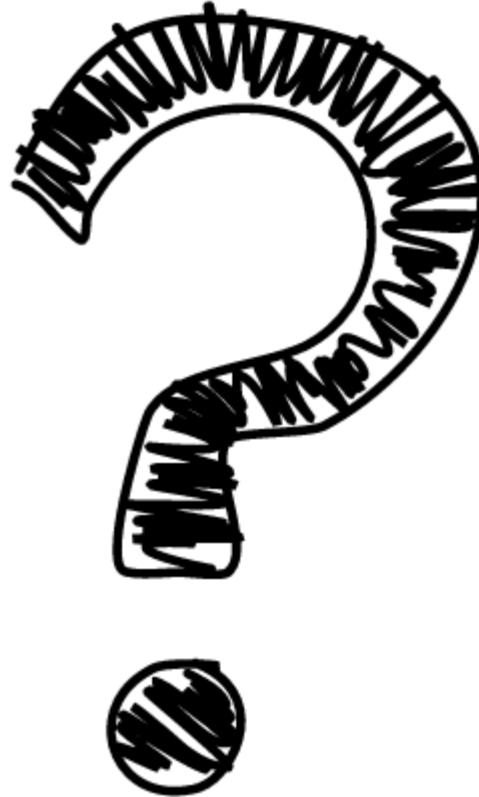
- trouver des personnes qui me ressemblent,
- voir ce qu'elles ont fait et que je n'ai pas fait

Meetic



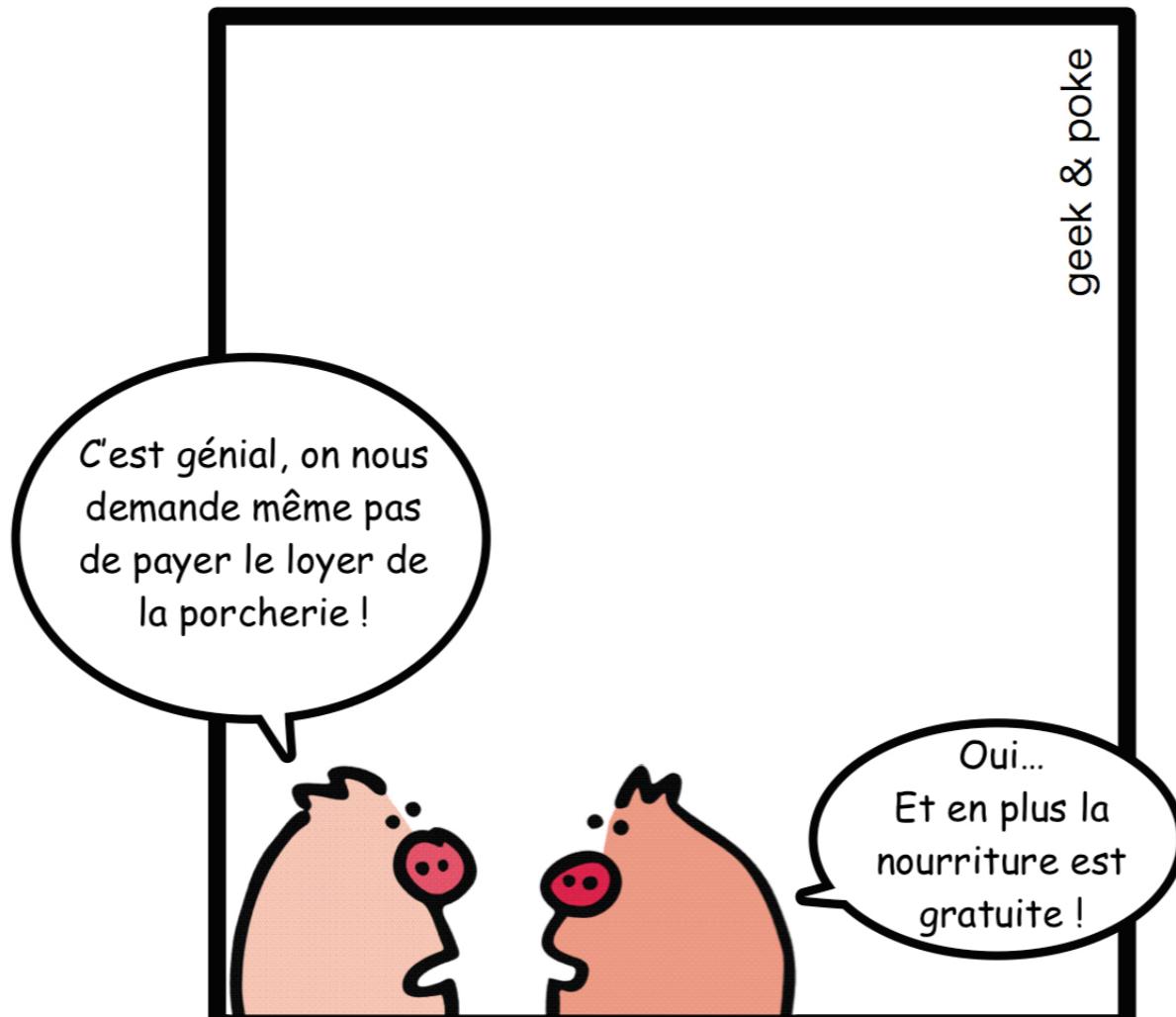
Quel est le modèle économique de

- Facebook
- Google
- Twitter
- ...



La publicité

Vendre de la publicité ultra-ciblée



Deux cochons discutant
du modèle « gratuit »

La détection de fraude

Les paradise papers

C'est la plus grande fuite de données de l'histoire, et de loin !

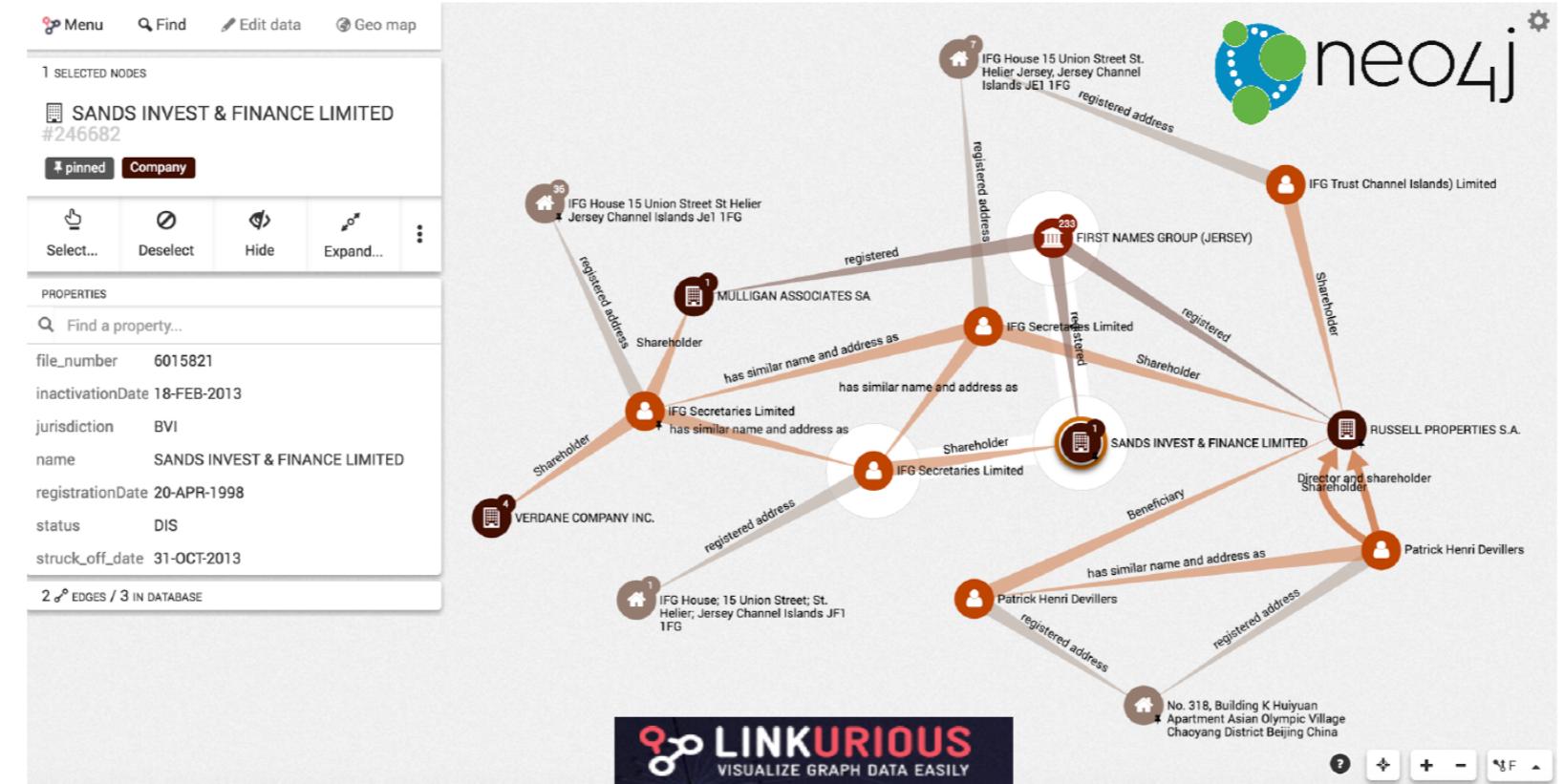


L'opacité

Ajouter un maximum d'intermédiaire pour rendre le plus opaque possible (holding + offshore).



C'est du graph !



Il faut suivre les petits cailloux

Le routing

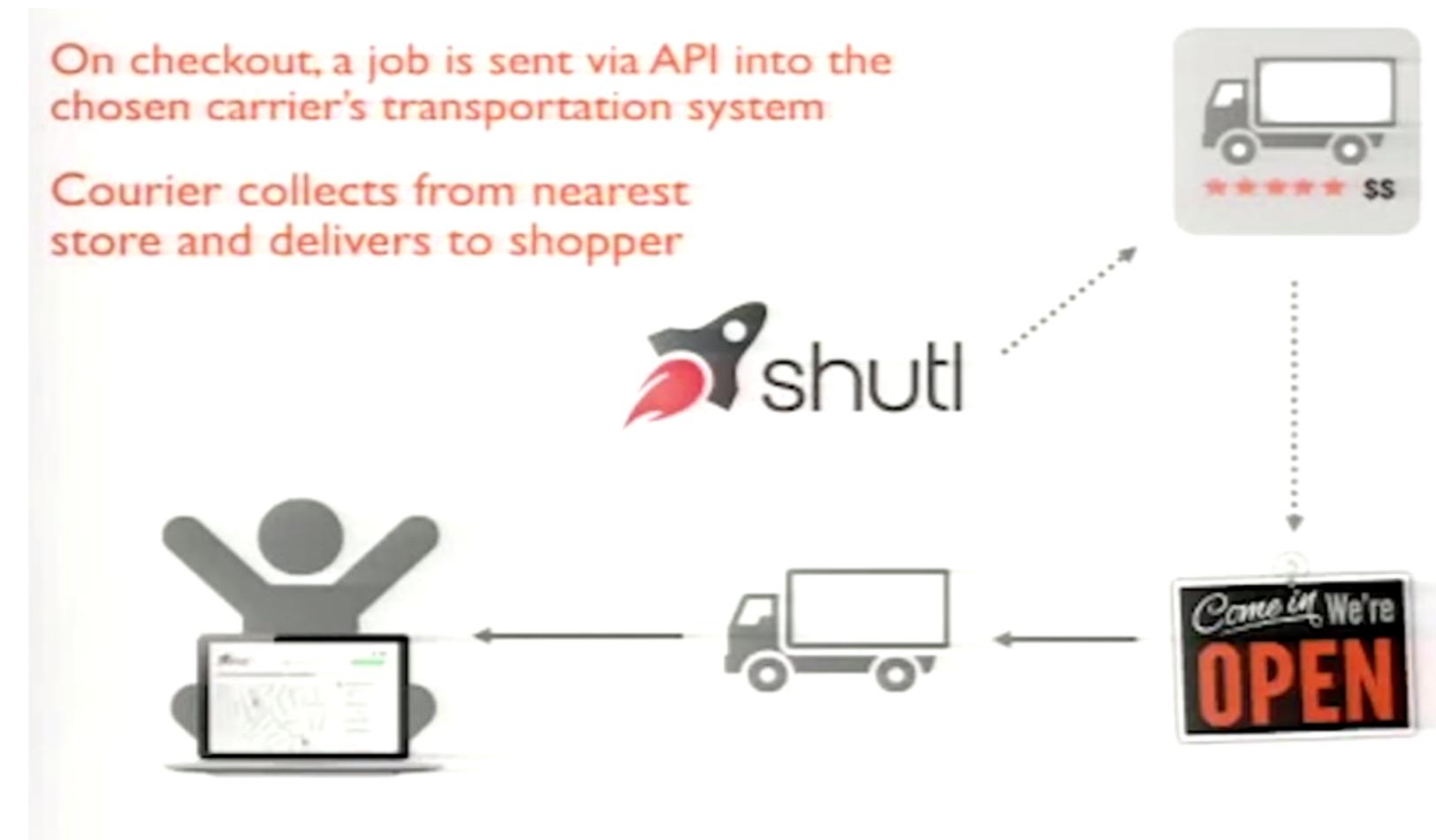
GPS

Calcul du plus court chemin



Ebay Now !

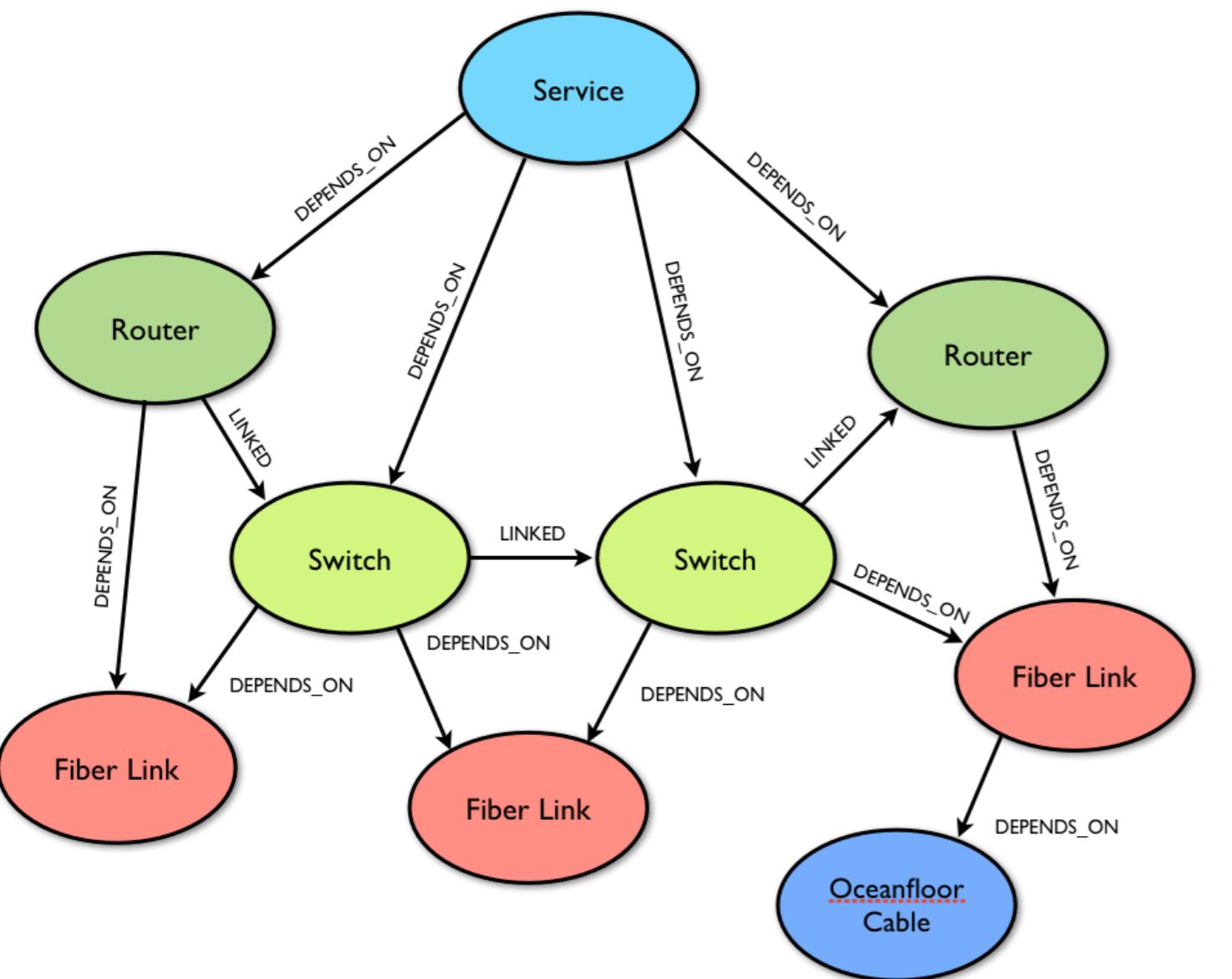
La livraison en moins de deux heures !



<https://vimeo.com/110675062>

Même principe qu'Uber pool

L'analyse des dépendances



Etudes d'impacts

- Pouvoir planifier une maintenance rapidement
- Identifier les clients impactés par une maintenance ou une panne.
- Identifier les faiblesses du réseau



Maintenance des avions

- Manuel de la maintenance
- Analyse des risque d'une modification

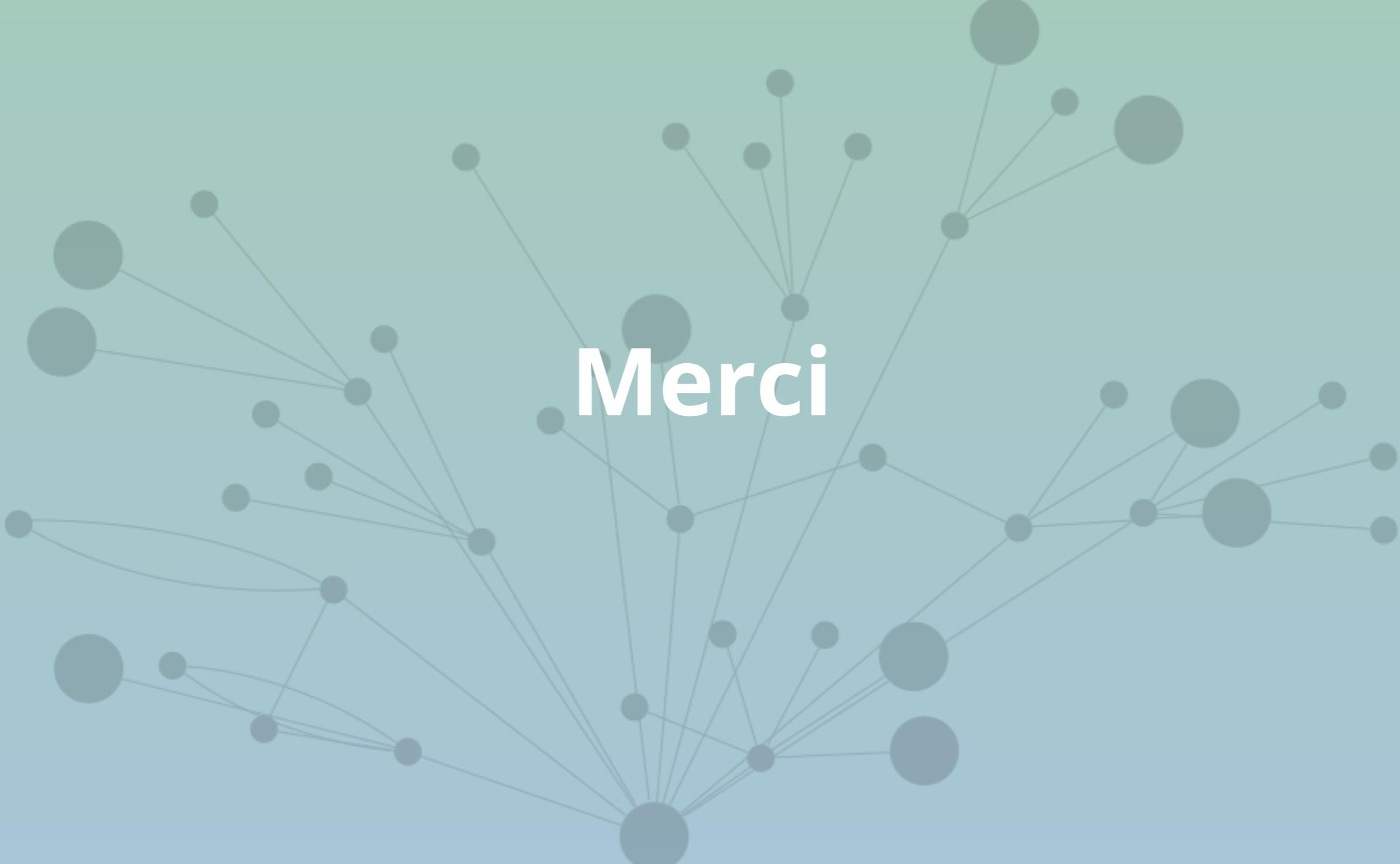


Graphs are everywhere!

Un est tout et tout est un

Le monde est connecté





Merci

Des questions ?

