

<http://www.coincoin.fr.eu.org/?Notes-Conference-Big-Data-INRIA>



# Notes Conférence Big Data INRIA (Euratechnologie)

- 1- Blog-Notes - Au boulot -

Date de mise en ligne : jeudi 26 novembre 2015

---

Copyright © L'Imp'Rock Scénette (by @\_daffyduke\_) - Tous droits réservés

---

Big Data & Data Scientists

Sylvain Carfe, INRIA

Collaboration INRIA avec es entreprises

René Lefébure, Wana Group R&D Conexance, Lille2

- Notion de capteur et de fréquence de données
- difficulté de trouver un algo (rien de clé en main)
- bruit -> audible -> mélodieux (avantage concurrentiel)
- besoin de montrer "élasticité de données" sur impact économique

Compétences attendues d'un data scientist

- mathématiques & statistiques
- informatique et nouveaux langages : R, python, matlab, sql => monde opensource cité pour des problématiques de coût essentiellement
- sens des affaires

outils bases de données : hadoop, sparks, storm, hive => s'entraîner avec caggle

Digital Ethics, Paul Olivier Gilberts : qui veut un café

Comparaison GAFA => création d'un monde datafié

L'université s'intéresse à tout ce qui est neuf, si ça marche pas, ça fait une publication,

L'industriel est plus réservé, si ça ne marche pas, c'est une perte sèche.

Le Big data devient le Smart Data

Amazon : théorie de la longue traîne (profondeur de produits) + besoin 0 pb client car l'insatisfaction cliente est coûteuse

=> obtenir satisfaction client donc des clients

Nouveaux outils : agrégateur de compte bancaire : Linkso

buzzwords : ubérisation (concurrence de travail artisanal) => disruption (on se fait piquer des parts de marché)

C. Biernacki, prof de math à Lille1, MODAL INRIA => super présentation !

<https://modal.lille.inria.fr/>

Big Stat en mode SAS, <https://modal-research.lille.inria.fr/BigStat/>

données x14000, coût divisé par 70000, phénomène de société pour utiliser ce nouveau joujou  
volume, velocity, variety, veracity

nouvelles données, mais les questions sont les mêmes (ex : scoring bancaire)

Notez que le traitement humain peut faire évoluer le résultat malgré le choix technique.

On a plusieurs modèles, tous sont faux, mais certains serviront

Présentation d'un outil INRIA de traitement des données

Pixeo vu par Deloitte

<http://www.lecomparateurdeassurances.com>

Modèle économique = lead generator, clients = compagnies d'assurance

Collecte de beaucoup de données (questions assurance) + données comportementales (produitsn garantie, )

Clarusse Dhaenens INRIA Dolphin + David Delrue / Alicante

Médecine personnalisée

Recherche CRISTAL

Médecine : personnalisée, préventif, prédictif, participatif

Problème de stockage des données : exemple du papier, cliché, imagerie numérique parfois énormes,  
difficulté de relier les documents

évocation du problèmes du voyageur de commerces (nombre de villes à visiter)

En général, besoin d'optimisation multi-objectif (exemple du trajet : km / prix : selon critère de choix)

## Notes Conférence Big Data INRIA (Euratechnologie)

---

=> recherche du front Pareto

A chercher un AWSEC2 agréé données de santé

David Saponi, directeur INRIA

Le Big Data, c'est comme la fricandelle, tout le monde en parle, personne ne sait ce que c'est

Pourtant activité forte à Lille