

# Les outils du Big Data

Leleux Laurent

2017 - 2018

# Data warehouse

- Début des années 1990
- Informatique décisionnelle
- Données rassemblées
- Collecte, nettoyage, normalisation...
- Gestion, MAJ difficile
- Historique

# Data warehouse

- Business intelligence
- Data mining
- Cubes multi-dimensionnels

# Hadoop

- Années 2000
- Grand nombre de données à traiter
- Stockage de fichiers
- HDFS
- MapReduce
- Hbase
- Hive

# Les outils de prédiction

- Trafic routier instantané + météo + historique  
= prédiction de guidage
- Modèles basé sur l'historique
- Succès / échecs => amélioration
- Machine learning

# Les outils de prédiction

- Détection de fraudes (banque, assurance...)
- Aide au diagnostic médical
- Tendances sur les marchés financiers
- Google – Epidémie de grippe
  - 50 millions de requêtes sur plusieurs milliards
  - Alertes en temps réel
  - Flutrends, Trend, Correlate

# L'analyse temps réel

- Stream analysis
- Données GPS
- Réseaux sociaux
- Données télécoms

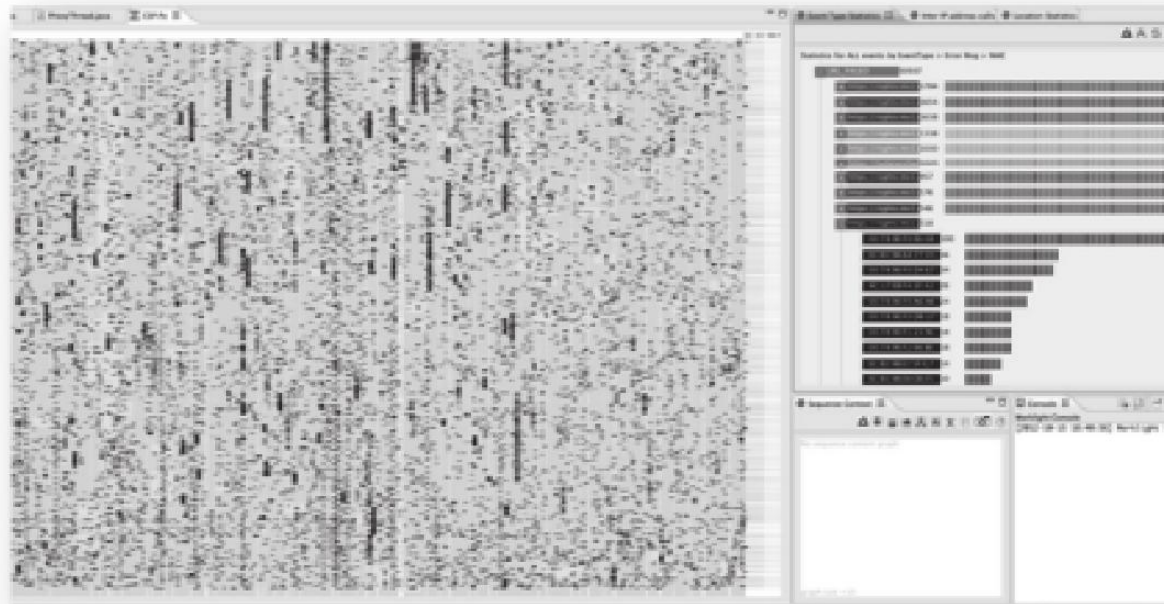
# Les outils de visualisation

- Pas d'approche causaliste
- Détecter les relations
- Data mining
- Graphes traditionnels
- Graphes, heatmaps...



# Les outils de visualisation

- Décodeurs instables
- Recherche de continuum



Analyse basée sur la visualisation de continuum<sup>1</sup>