

Les 3V : Volume, Vitesse et Variété

Leleux Laurent

2017 - 2018

Chiffres

- 2,5 milliards d'internautes
- 1,9 milliards sur les réseaux sociaux
- 6,5 milliards de mobiles
- 10 milliards d'objets connectés

Encore des chiffres

- Chaque jour :
 - Google traite 24 pétaoctets de données
 - 350 millions de photos sur Facebook
 - 400 millions de tweets sont envoyés
- Chaque minute :
 - 100 heures de vidéos sur YouTube

=> Nouveaux outils

Remarque

- Analyser de gros volumes traditionnels
!= Big Data
- Combiner des données volumineuses ou non,
structurées ou non pour décider d'une action
= Big Data

Volume

- Evolution en fonction des progrès scientifiques
- Exemple : être humain

« Plus on a de données sur un phénomène et son environnement, mieux on l'appréhende. »

Volume

- Corrélations
 - Grippe
 - Mots clefs et marchés boursiers
 - Analyse du climat économique
- Prédiction plutôt que causalité
- Échantillonnage (sondage) vs. Big Data

Vélocité

- Eurovision 2012
 - 100 millions de téléspectateurs
 - 9 millions de votes
 - 4h d'émission
-
- Plusieurs millions de tweets
 - 1,5 millions significatifs

Vélocité

- Amazon
- Geo-fencing (géorepérage)
- Real time bidding (RTB)
- Changement de fournisseur
- Instantanéité, « fraîcheur »

Variété

- Tout devient donnée
- Texte (OCR, analyse sémantique)
 - Relations
- Son (reconnaissance vocale)
 - Help desk VOD
- Image (reconnaissance de formes)
 - Vidéo-surveillance
- Odorat, toucher...

Variété

- Données
 - Structurées
 - Semi-structurées
 - Non structurées
- Intelligence artificielle