

Introduction

- Depuis le début des années 2000 nous utilisons des volumes de données de plus en plus importants:
 - Réseaux sociaux
 - Logs
 - Objets connectés
- Définition du Big Data
 - Le concept, tel qu'il est défini actuellement, englobe un ensemble de technologies et de pratiques destinées à stocker de très grandes masses de données et à les analyser très rapidement.

Les 3 V (5 V)

- **Volume**

La quantité de données générée est en pleine expansion et suit une loi quasi exponentielle. Le commerce électronique et les réseaux sociaux sont les grands contributeurs de cette profusion de données.

- **Vélocité (rapidité)**

La rapidité de renouvellement des données dans un monde connecté n'est plus à démontrer. Toutes les nouveautés et mises à jour sont stockées en respect de la devise du big data : "On garde tout!".

- **variété**

Les bases de données spécialisées sont en mesure de gérer la multiplicité des formats des données : numérique, texte, image... Voir les bases de données [NoSql](#) ou NewSql.

Valeur

- **Véracité (ou fiabilité)**

la fiabilité des données collectées est essentielle pour les transformer en informations utilisables dans l'entreprise, les données doivent être nettoyées (data cleansing) avant qu'elles ne soient exploitées dans le processus décisionnel

- **Valeur**

Elle correspond au profit qu'on peut tirer des données

- **Vertu (6^e V)**

La vertu fait référence aux réglementations en matière de confidentialité et de conformité des données telle que le RGPD en europe.

Unités de mesure dans le domaine du big data

Nom	Symbole	Valeur
kilo-octet	ko	10^3
méga-octet	Mo	10^6
giga-octet	Go	10^9
téra-octet	To	10^{12}
péta-octet	Po	10^{15}
exa-octet	Eo	10^{18}
zetta-octet	Zo	10^{21}
yotta-octet	Yo	10^{24}

- Les unités de mesure utilisées dans le monde du Big Data (en gras)

Exploitation des données du Big Data

- Data Science
- Data Mining
- Machine Learning
- Deep Learning