

Data... c'est quoi ça ?

Une définition d'abord

*"Les données sont les matériaux bruts produits dans l'**abstraction du monde** en catégories, mesures et toute autre forme de représentation qui constituent **les fondations sur lesquelles l'information et le savoir sont créés.**"*

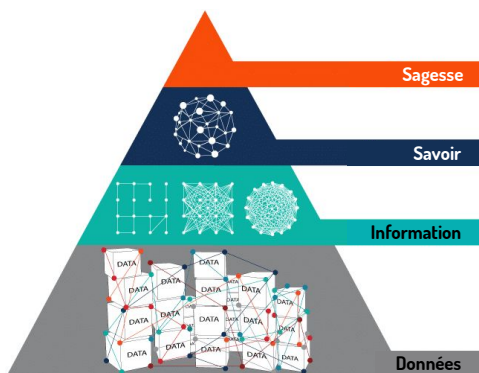
Source : Rob Kitchin, *The Data Revolution. Big Data, Open Data, Data Infrastructures and Their Consequences*, Londres, Sage, 2014.

Une abstraction du monde

Les données permettent d'inscrire la complexité du monde dans des objets que l'on peut transformer et faire circuler.



Le fondement de l'information et du savoir



Les **données** sont la matière "brute" de l'information conçues plutôt pour des machines.

Avec l'**information**, les données ont été interprétées pour dégager du sens pour des humains.

En donnant du sens à de l'information, on obtient du **savoir**. En donnant du sens à la connaissance, on obtient de la **sagesse**.

La base de l'informatique

Unité	Signification
Bit (b)	C'est l'unité de base de l'informatique
Kilobyte (KB)	Une page de texte équivaut à une dizaine de KB
Megabyte (MB)	Un morceau au format MP3 occupe environ 4MB sur un disque dur.
Gigabyte (GB)	Un film de 2h environ peut être compressé en 1 ou 2 GB
Terabyte (TB)	La Bibliothèque Nationale de France archive plus de 600 TB de pages web
Zettabyte (ZB)	On estime que l'humanité a produit plusieurs dizaines de ZB de données dans son histoire.

Les données ne sont pas seulement le fondement du savoir, elles sont aussi la base de l'informatique. **Tout ce qui circule dans un ordinateur, ce sont des données.**

Pensez à votre abonnement téléphonique, chaque mois, vous payez pour consommer un certain volume de données quantifié en octet ou en bit.

Le volume des données créées et traitées ne cesse de croître en même temps que les capacités de calcul des ordinateurs.

