



Données structurées

1.1 Les données

★ Vidéo introductory d'Allô la Hotline.

Une **donnée** est un élément se rapportant à une personne, un objet ou un événement. On parle de **données personnelles** si celles-ci permettent d'identifier directement ou indirectement une personne physique. Par exemple, une photo, un numéro de téléphone, de sécurité sociale, une adresse, etc.

Les **métadonnées** sont des données secondaires s'ajoutant à une donnée principale. Par exemple, pour une photo, cela peut-être la date, la localisation, les personnes ou choses présentes dessus, etc ; pour un morceau de musique, cela sera le nom de l'artiste, de l'album, le genre, etc.

1.2 Données structurées

Les données récoltées sont organisées sous forme de tableau – on parle de **table de données** – organisé à l'aide de **descripteurs** prenant diverses **valeurs**.

Exemple :

Prénom	Nom	Sexe	Profession	Planète d'origine
Luke	Skywalker	M	Jedi	Tatooine
Anakin	Slywalker	M	Sith	Tatooine
Leia	Organa	F	Rebelle	Alderaan
Obi-Wan	Kenobi	M	Jedi	Inconnue
	R2D2		Astrodroïde	Inconnue

Ici les descripteurs sont : prénom, nom, sexe et les valeurs prises sont par exemple : R2D2, astrodroïde...

Ces tables de données structurées sont généralement contenues dans des fichiers aux formats CSV, XML ou JSON mais il en existe d'autres. Leur traitement peut s'effectuer de diverses façons : tri, recherche, calcul, etc.

1.3 Stockage de données

Bien que les données soient stockées sur nos appareils et ordinateurs, elles le sont de plus en plus sur des serveurs informatiques auxquels on se connecte via Internet : le **Cloud**. Le Cloud permet aujourd’hui à de multiples appareils d'accéder et de partager des données, par exemple une photo prise sur téléphone est accessible depuis notre ordinateur ; ces données sont **synchronisées**. Le Cloud ne permet pas seulement le stockage de données mais aussi leur utilisation via des logiciels eux-mêmes stockés sur le Cloud.

Le Cloud a un immense impact écologique de part les ressources nécessaires à la fabrication des serveurs qui le composent et aux infrastructures qui le relient à Internet ainsi que de part sa consommation énergétique qui représente aujourd’hui 10% de l'électricité mondiale.

1.4 Big Data, profilage et marketing politique

★ Vidéo de Datagueule, Le Big Data.

★ Vidéo de Fouloscopie, Peut-on prédire le crime ?

★ Vidéo de Datagueule sur le marketing politique

Quelques questions pour alimenter la réflexion et la discussion.

- Qu'est-ce qu'une donnée personnelle ? Comment sont-elles protégées ?
- Que permet l'exploitation de ces données personnelles ?
- Que permet l'exploitation de ces données en grand volume, le Big Data ?
- Le Big Data est-il toujours pertinent ? Éthique ?
- Comment le profilage d'utilisateurs permet de cibler des individus à des fins politiques ? Ce ciblage peut-il aller au delà du simple marketing politique ?
- Notre navigation sur le web influence-t-elle nos opinions ?
- Quelles sont les données que l'on récolte sur les utilisateurs ? Sont-elles toutes pertinentes ? Cette récolte est-elle souhaitable ?
- Des puissances étrangères peuvent-elles utiliser le marketing politique afin d'influencer la vie politique d'un autre pays ?
- Le profilage des utilisateurs doit-il être régulé ?
- Le marketing politique sculpte-t-il notre vision du monde ?