

Chapitre I : Les données structurées

Les données constituent la matière première de toute activité numérique. Il faut alors pouvoir les conserver de manière persistante et organisée afin de les exploiter facilement pour produire de l'information. Le stockage des données numériques représente ainsi des enjeux sociétaux, individuels, écologiques et économiques.

I. Repères historiques

Histoire des données

	Événement	Inventeur	Année
1	Invention du ruban perforé	Basile Bouchon	1725
2	Invention carte perforée	Jean-Baptiste Falcon	1728
3	Premier disque dur IBM 350	Reynold Johnson	1956
4	Modèle relationnel	Edgar Codd	1970
5	Création du premier tableur VisiCalc		1979
6	Invention de la mémoire flash	Fujio Masuoka	1984
7	Apparition du terme Big Data		1997
8	Open data : Charte du G8 pour l'ouverture des données publiques	Barack Obama	2009

II. Les données

a) Les données personnelles

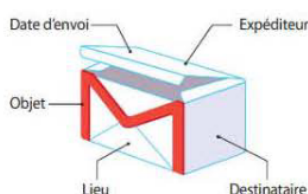
Définition : Une donnée est une valeur décrivant un objet. Elle peut être de nature diverse comme un texte (nom), une valeur numérique (date, numéro de téléphone, ...), une photographie, une empreinte digitale, ...

b) Les métadonnées

Définition : Une métadonnée est une donnée qui accompagne tout fichier numérique pour en décrire le contenu.

Exemple : Pour un fichier de musique, la donnée principale est la chanson. Le nom de l'artiste ou le nom de l'album sont des métadonnées.

Exemple : Les métadonnées d'un courrier



III. Les données structurées

a) Les tables de données

On organise les données sous la forme de tableau appelé "table de données".

Une liste de **descripteurs** permet de caractériser les données. Un **objet** est un élément d'une table. Une collection regroupe des objets partageant les mêmes descripteurs. Les **valeurs** de tous les descripteurs d'un objet sont précisées. Les données sont dites "structurées".

Exemple :

Prénom	Nom	Âge
Antoine	Ledoux	15
Pauline	Darcis	16
Karim	Chanhou	15
Sarah	Goldberg	14

b) Formats et représentations des données

Les données sont organisées de manière à s'adapter à des traitements spécifiques. La manière dont elles sont organisées est le format. Les formats CSV, JSON et XML sont les plus utilisés.

Connaître la syntaxe des trois formats

Données au format CSV	Données au format JSON	Données au format XML
<pre> 1 prenom;nom;classe 2 Antoine;Ledoux;seconde 3 Pauline;Darcis;premiere </pre>	<pre> 1 { 2 { 3 "prenom": "Antoine", 4 "nom": "Ledoux", 5 "classe": "seconde" 6 } 7 } 8 { 9 "prenom": "Pauline", 10 "nom": "Darcis", 11 "classe": "premiere" 12 } </pre>	<pre> 1 <eleve> 2 <prenom>Antoine</prenom> 3 <nom>Ledoux</nom> 4 <classe>seconde</classe> 5 </eleve> 6 <eleve> 7 <prenom>Pauline</prenom> 8 <nom>Darcis</nom> 9 <classe>premiere</classe> 10 </eleve> </pre>

IV. Le traitement des données structurées

Le traitement des données peut être réalisé de diverses manières : recherche, tri, estimation, calcul, ...

Pour réaliser ces opérations, des phrases logiques appelées "requêtes" sont exprimées dans un langage informatique. Le plus utilisé est le SQL.

Une requête est composée de mots-clés dans un ordre précis.

Commandes à connaître :

.tables	pour connaître le nom des tables
select	sélectionner (les descripteurs suivantes)
from	de la table
where	avec la condition suivante
order by	valeurs ordonnées selon

V. Le cloud

a) Le cloud :

Définition : Le cloud désigne l'accès à des ressources informatiques (stockage, logiciels, données, ...) situées sur des serveurs informatiques distants par l'intermédiaire d'un réseau.

Il est possible d'automatiser le stockage de ses données sur le cloud en paramétrant la synchronisation des fichiers sur son ordinateur ou son téléphone. Le partage des données y est également facilité.

Exemple : Les outils collaboratifs de Google



Avantages du cloud :

- Pouvoir accéder à ses données depuis n'importe où
- Sécurité des données
- Protection meilleure contre les intrusions
- Economie de gestion pour les entreprises

Inconvénients du cloud :

- Perte de contrôle des actions
- Problèmes écologiques
- Hyper centralisation des données
- Données utilisées pour le ciblage publicitaire
- Données stockées à l'étranger sous législation étrangère
- Etre connecté à internet

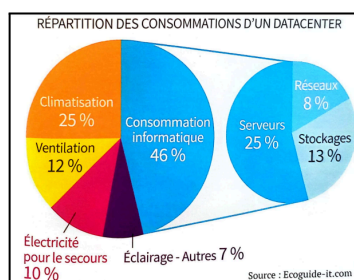
b) L'impact sur la consommation énergétique

Avec l'augmentation de la quantité de données stockées et traitées dans les centres de données ou data centers, le cloud est devenu l'un des premiers consommateurs d'électricité dans le monde. Les entreprises doivent adapter leurs technologies pour réduire leur impact écologique.

Exemple : Les centres de données consomment 10% de l'électricité mondiale

Un centre de données consomme autant d'électricité que 30 000 habitants européens.

Un datacenter green à Grenoble



Quelques conseils pour réduire la consommation énergétique :

- ne pas utiliser google mais écrire directement l'adresse
- effacer les mails, photos, vidéos non regardées
- prendre conscience de sa consommation en calculant son empreinte carbone numérique
- utiliser le wifi plutôt que la 4G/5G
- modifier ses paramètres