

Séquence Données structurées — Cours

1 Les supports de stockage

1.1 Historique

1930 Utilisation de cartes perforées

1956 Invention du disque dur

1.2 Différents supports de l'information

Les données numériques sont stockées sur des supports très variés.

FIGURE 1 – Premier support numérique : les cartes perforées



Nom	Capacité	Vitesse	technologie
Cartes perforées	1 octet/carte	1 octet par seconde	carton perforé
CD-ROM	700Mo	1 à 100 Mo /s	Laser et feuille métallique
Cartes SD	1-1000 Go	1-100Mo /s	Puces de stockage
Disques durs	1Go-100To	4Mo-1Go /s	Plateau métallique et pointe magnétique
SSD	1Go-100To	100Mo-10Go /s	Puces de stockage
RAM	100Mo-100Go (volatile!)	100Mo-100Go /s	Puces de stockage

FIGURE 2 – Quelques exemples de moyens de stockage

INFO 1

Un support de stockage est défini par les caractéristiques suivantes :

- Sa capacité de stockage (en octet).
- La vitesse à laquelle on peut lire et écrire les informations (on ne peut pas toujours les écrire).
- La technologie utilisée.

2 Traitement des données

2.1 Historique

1979 Création du premier tableur, VisiCalc.

2013 Charte du G8 pour l'ouverture des données publiques.

2.2 Données structurées

DÉFINITION 1

Il est possible de structurer des données sous forme de tableau. Les Noms des colonnes sont appelés « descripteurs », et chaque ligne représente une entité.

	A	B	C
1	nom	prenom	date_naissance
2	Danioko	Aya	10/05/1995
3	Glover	Donald	25/09/1983
4	Simpson	Clifford	16/07/1996
5	Monée	Janelle	01/12/1995
6			

3 Hébergement de données en ligne

DÉFINITION 2

Les services d'hébergement en ligne, appelés « nuages » ou « cloud » servent à stocker, modifier et partager des documents.

Les services les plus connus sont Google Drive, Icloud et DropBox. Il ne sont cependant pas interopérables, c'est à dire qu'ils ne peuvent pas communiquer entre eux (on ne peut pas partager directement un document google à un utilisateur dropbox).

DÉFINITION 3

Un système fédéré est un service pouvant être fourni par plusieurs acteurs différents et toujours offrir des services communs et compatibles. Chaque acteur héberge une **instance** du service.

EXEMPLE 1

Les mails sont fédérés car plusieurs acteurs hébergent un serveur (une instance) mail :

- Gmail
- Outlook
- Gresille
- Yopmail

Pourtant, tous ces serveurs peuvent communiquer entre eux pour s'échanger des messages.

Un système de stockage de fichiers fédéré bien connu est : NextCloud.

4 Consommation électrique d'internet

4.1 Quelques estimations

La consommation mondiale d'électricité est de 23 000 TWh en 2019.

En 2017, Greenpeace estime qu'internet consomme 7% de l'électricité mondiale.

En 2020, on estime à 650 terawatt-heure la consommation électrique des centres de données (environ 3%), et à 10% la consommation totale d'internet (centres, réseaux et terminaux).

La consommation électrique d'internet est difficile à évaluer car les composants sont très hétérogènes et dispersés dans le monde.

DÉFINITION 4

On estime qu'internet consomme entre 10% de la production électrique mondiale.

4.2 Un internet durable

Les grands acteurs d'internet comme Google, Apple, Facebook et beaucoup d'autres, annoncent utiliser de plus en plus d'électricité issue de sources renouvelables.

Les énergies renouvelables étant par nature très intermittentes, il est actuellement impossible de produire une quantité constante d'électricité.

DÉFINITION 5

Un centre de données 100% électricité renouvelable ne pourrait pas être allumé en permanence, donc certains services seraient indisponibles en cas de faible production.

5 Données personnelles

DÉFINITION 6

Une donnée personnelle est une information se rapportant à **une personne physique et pouvant mener à son identification** directe ou indirecte.

Parce qu'elles concernent des personnes, celles-ci doivent conserver la maîtrise de leurs données personnelles. Une personne physique peut être identifiée :

directement exemple : nom et prénom.

indirectement exemple : par un numéro de téléphone ou de plaque d'immatriculation, un identifiant tel que le numéro de sécurité sociale, une adresse postale ou courriel, mais aussi la voix ou l'image

L'identification d'une personne physique peut être réalisée :

- à partir d'une seule donnée (exemple : nom);

— à partir du croisement d'un ensemble de données (exemple : une femme vivant à telle adresse, née tel jour et membre dans telle association) :

Par contre, des coordonnées d'entreprises (par exemple, l'entreprise « Compagnie A » avec son adresse postale, le numéro de téléphone de son standard et un courriel de contact générique « compagnie1@email.fr ») ne sont pas, en principe, des données personnelles.