

Des données structurées & leurs traitements

- des données ?
- le BO
- qqs chiffres
- proposition de séquence

Vocabulaire

- une donnée :
 - structurée
 - non structurée

Types de données	Caractéristiques	Métaphore associée
Données structurées	Organisation logique Formats identiques Facilité de recherche Appréhensible par un ordinateur	Couverts rangés dans un tiroir à couvert
Données non structurées	Entassement sans logique Formats différents Difficulté de recherche Nécessite un traitement humain	Couverts entassés dans une caisse à une brocante

Figure 8 – Différence entre données structurées et non structurées

image [<https://www.webmarketing-com.com/2017/04/20/58377-tirer-parti-big-data-reseaux-sociaux>]

- métadonnées : données qui décrivent d'autres données
- connaissance = somme des infos (modèle interprétatif+donnée)
- rq : selon descripteur, données +/- faciles à utiliser

vidéos

- [https://culturenumerique.univ-lille.fr/module3.html#sec_1]
- [https://culturenumerique.univ-lille.fr/module3.html#sec_2]

B.O.

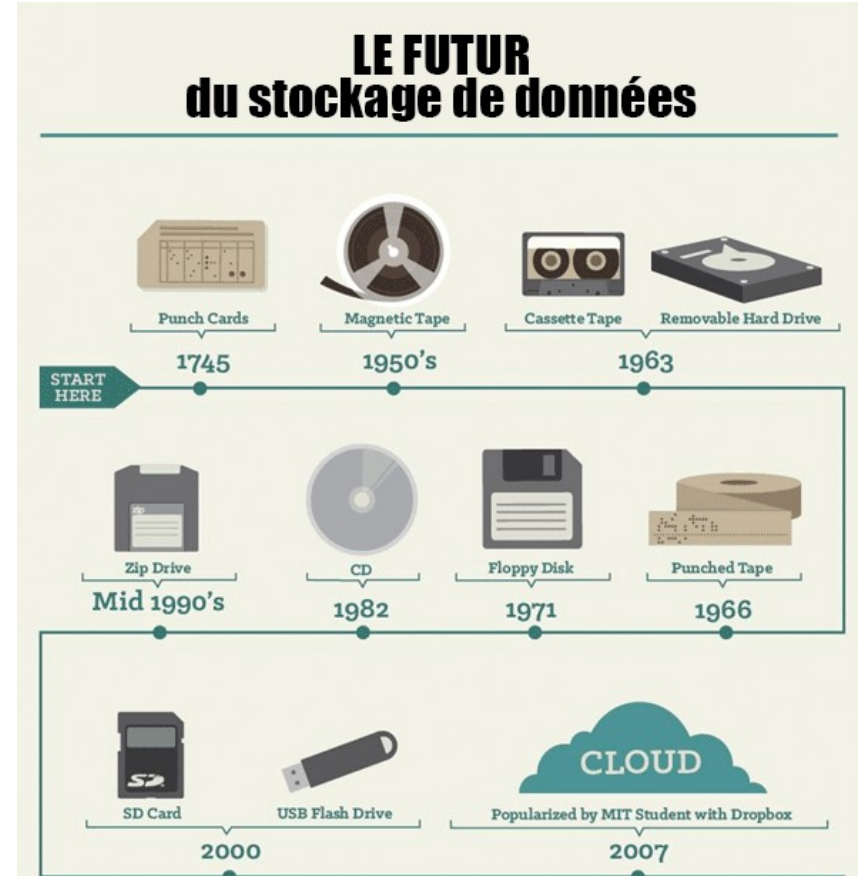
image
[<https://www.education.gouv.fr/cid140434/les-nouveaux-programmes-du-lycee-general-et-technologique-a-la-rentree-2019.html>]

Contenus	Capacités attendues
Données	Identifier les principaux formats et représentations de données.
Données structurées	Identifier les différents descripteurs d'un objet. Distinguer la valeur d'une donnée de son descripteur. Utiliser un site de données ouvertes, pour sélectionner et récupérer des données.
Traitement de données structurées	Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.
Métadonnées	Retrouver les métadonnées d'un fichier personnel.
Données dans le nuage (<i>cloud</i>)	Utiliser un support de stockage dans le nuage. Partager des fichiers, paramétrer des modes de synchronisation. Identifier les principales causes de la consommation énergétique des centres de données ainsi que leur ordre de grandeur.
Exemples d'activités	
<ul style="list-style-type: none">— Consulter les métadonnées de fichiers correspondant à des informations différentes et repérer celles collectées par un dispositif et celles renseignées par l'utilisateur.— Télécharger des données ouvertes (sous forme d'un fichier au format CSV avec les métadonnées associées), observer les différences de traitements possibles selon le logiciel choisi pour lire le fichier : programme Python, tableur, éditeur de textes ou encore outils spécialisés en ligne.— Explorer les données d'un fichier CSV à l'aide d'opérations de tri et de filtre, effectuer des calculs sur ces données, réaliser une visualisation graphique des données.— À partir de deux tables de données ayant en commun un descripteur, montrer l'intérêt des deux tables pour éviter les redondances et les anomalies d'insertion et de suppression, réaliser un croisement des données permettant d'obtenir une nouvelle information.— Illustrer, par des exemples simples, la consommation énergétique induite par le traitement et le stockage des données.	

Quelques chiffres

- nb de données créées en 1 mois : 43.10^{18} bytes (!!!)
[vidéo Big data]
- nécessite de l'énergie : 10-15 %
[cliquer c'est polluer]

image [<http://www.emsi-histoireinformatique.fr/le-stockage-en-informatique/>]



proposition de séquence

- prévoir un TP0 sur l'utilisation de l'OS (sauvegarde, chemin des fichiers, ...)
- après la thématique internet
- vidéos générales + étude de documents sur l'écologie : quizz et/ou présentation par les élèves
- pollution de l'air : [<http://forhan.maths.free.fr/index.html>]
 - xml (possiblement connecté)
 - csv
 - python (nécessite les boucles)
- villes de France : csv [<https://www.angelique-renaud.com>]