

SNT, Données structurées

Exercice 1

En vous servant de la 1ere vidéo de la page: http://www.monlyceenumerique.fr/snt_seconde/donnees_structurees_traitement/donnees_structurees.php, répondez aux questions suivantes.

1. Quelle invention datant du XVIIIe siècle a été mis à profit par les informaticiens pour stocker des données ?
2. En quelle année a été créé le premier disque dur ?
3. En quelle année a été développé le premier modèle permettant de structurer et d'indexer des bases de données ?
4. Comment s'appelle le premier tableur mis à la disposition du grand public ?
5. En quelle année l'*Open Government Initiative* a t-elle été mise en place ?
6. Quelle est le nom du texte de loi européen qui protège les citoyens européens quant à leurs données personnelles ?

Notions de données structurées / non structurées

Une **donnée** est une valeur (numérique ou textuelle) représentant une information.

Pour pouvoir utiliser cette donnée d'un point de vue informatique, il va être utile de la "**structurer**".

On peut pour cela utiliser par exemple des **tables de données**.

- La première ligne décrit la forme des lignes suivantes ; les éléments de cette ligne donnant la signification portent plusieurs noms possibles : ce sont les **noms de champ** ou **descripteurs** ou **critères**.
- Les lignes suivantes sont appelées des **objets** : elles sont la liste des valeurs de chacun des descripteurs.

prenom	nom	age	classe
Jean	Céplusse	16	Seconde 3
Leïla	Mina	14	Seconde 2
Arjuna	Mahatman	15	Seconde 3
Mathilda	Vantage	15	Seconde 3

Entourer en bleu sur la table de données ci-dessus : les descripteurs.

Entourer en vert sur la table de données ci-dessus : un objet.

Entourer en rouge une valeur.

De quel type sont les valeurs du champs "âge" ?

De quel type sont les valeurs du champs "prénom" ?

Différents formats pour représenter des données structurées.

Pour pouvoir stocker, transmettre et exploiter ce type de tableau de valeurs, il peut être représenté sous différentes formes appelées **format**.

Les principaux formats utilisés pour représenter un ensemble de données sont le CSV , le JSON et le XML.

- L'acronyme **CSV** signifie "Comma Separated Values" c'est-à-dire "valeurs séparées par une virgule".
- L'acronyme JSON signifie "JavaScript Object Notation" c'est-à-dire "notation d'objets en Javascript".
- L'acronyme XML signifie "eXtensible Markup Language " c'est-à-dire "langage de balisage extensible".

Les données du tableau précédent en .csv seront représentées sous la forme :

```
1 | prenom;nom;age;classe
2 | Jean;Céplusse;16;2nde
3 | Leïla;Mina;14;2nde
4 | Arjuna;Mahatman;17;1ere
5 | Paul;Vantage;15;2nde
```

en .json elles seront représentées sous la forme :

```
1 | [{
2 |   "prenom":"Jean",
3 |   "nom":"Céplusse",
4 |   "age":16,
5 |   "classe":"2ndé"
6 | },
7 | {
8 |   "prenom":"Leïla",
9 |   "nom":"Mina",
10 |  "age":14,
11 |  "classe":"2ndé"
12 | }
13 | .....
14 | ]
```

en .xml

```

1 <eleve>
2   <prenom>Jean</prenom>
3   <nom>CepLusse</nom>
4   <age>16</age>
5   <classe>2nde</classe>
6 </eleve>
7 <eleve>
8   <prenom>Leila</prenom>
9   <nom>Mina</nom>
10  <age>14</age>
11  <classe>2nde</classe>
12 </eleve>
13 .....

```

A quel langage peut-on penser en voyant du .xml ?

Pour étudier, modifier et traiter des tables de données , on peut utiliser des logiciels spécifiques appelés **tableurs**.

Donner le nom de 3 tableurs différents :

.....

Open data ou données libres

L'**open data** correspond à une volonté politique d'ouvrir les données jugées d'intérêt public afin d'encourager leur réutilisation par tout un chacun. Cela permet d'encourager la transparence démocratique, de bénéficier de services au quotidien ou de prendre des décisions plus éclairées.

RGPD

Qu'est-ce que le RGPD ? Le sigle RGPD signifie « **Règlement Général sur la Protection des Données** » (en anglais « General Data Protection Regulation » ou GDPR). Le RGPD encadre le traitement des données personnelles sur le territoire de l'Union européenne.

Quels sont les objectifs principaux du RGPD ?

L'objectif du RGPD est de permettre aux citoyens européens de vérifier et faire valoir la mise en application de leurs droits par les organismes qui sont :

1. Le droit à l'information : droit de savoir comment sont collectées les données et dans quel but
2. Le droit d'accès : droit de savoir quelles données les entreprises collectent sur un individu
3. Le droit d'opposition : droit de s'opposer à la collecte des données personnelles (sauf obligation légale)
4. Le droit de rectification : droit de modifier des données transmises à une entreprise
5. Le droit à l'oubli : droit de suppression des données ainsi que celles des sous-traitants
6. Le droit à la portabilité : droit de récupérer ses données de façon lisible

Data Center

Après avoir regardé la vidéo : <https://www.youtube.com/watch?v=jvwiMNCg9sw>

répondre aux questions suivantes :

1. Qu'est-ce qu'un **data center** ou **centre de données** ?
2. Quelles sont les deux principales raisons de la consommation électrique de ces centre de données ?
3. Quelle est la consommation électrique, en térawattheure par an, estimée en 2018 pour l'ensemble des centre de données mondiaux ? Comparer à la consommation de la France en 2018 qui fut de 475 térawattheure selon [EDF](#).
En 2020, la Commission Européenne a publié un rapport (cf. [page 99 de ce rapport](#)) estimant que la consommation énergétique des data centers mondiaux pourrait dépasser les 800 térawattheures par an en 2030.
4. Que pensez-vous de la phrase *"Si Internet était un pays, il serait le troisième plus gros consommateur mondial d'électricité derrière la Chine et les États-Unis"* ?
5. Quelles mesures, à votre niveau, permettent de réduire votre empreinte carbone ?