



manuel utilisé et mentionné dans cette fiche →

Les données constituent la matière première de toute activité numérique. Afin de permettre leur réutilisation, il est nécessaire de les conserver de manière persistante. Les structurer correctement garantit que l'on puisse les exploiter facilement pour produire de l'information.

À regarder (< 3 min) →
https://youtu.be/79t_bG6sRvA?t=21



<https://youtu.be/XfIMhukDH-s>



Avec les données, on peut tout faire. Par exemple :
 Avec le "crédit social", la Chine classe les "bons" et les "mauvais" citoyens

À regarder (< 4 min) →
<https://youtu.be/5otaBKsz7k4>



p.76 :

DOCUMENT 3 Data center



Les *Data Centers* (centres de données) sont des lieux de stockage des informations sous formes de données structurées (base de données) dans des pièces ou des bâtiments. Les quantités de données stockées sont de l'ordre du zettaoctet (10^{21}) voire du yottaoctet (10^{24}). L'équipement informatique de ces centres de données comporte des ordinateurs, des serveurs par milliers et des baies de stockage immenses. Ces réseaux informatiques sont en permanence en fonctionnement et connectés sur le Web.

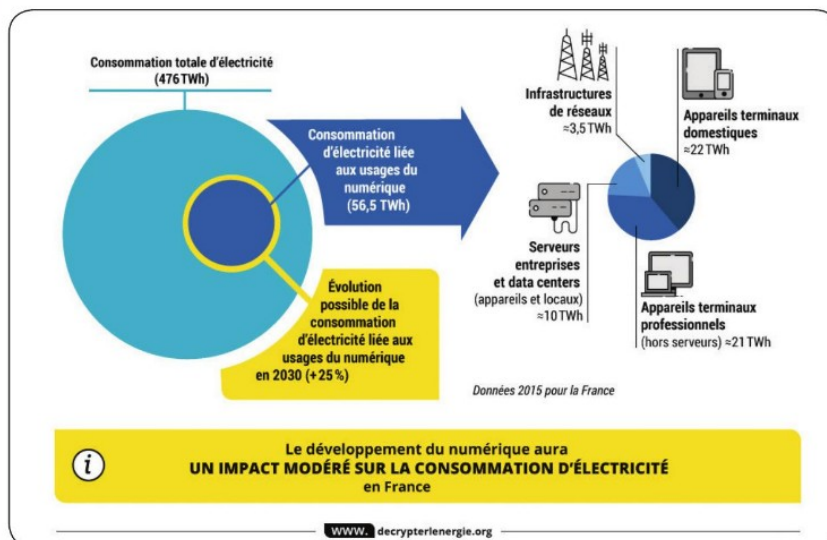
Leur consommation énergétique est importante, le coût de l'électricité représente près de 30 % des coûts d'exploitation des centres dont 30 à 60 % utilisés pour le refroidissement des serveurs.

À regarder (< 10 min) →
<https://youtu.be/rO6bXt7d2L8>



DOCUMENT 4

Consommation énergétique



En France, afin de tenir compte de la forte consommation énergétique due à l'usage des systèmes informatiques, un coup de pouce fiscal sur le coût de l'électricité a été donné aux *Data Centers*. Cet amendement, voté dans le cadre du projet de loi de finances pour 2019, consiste à réduire de moitié la taxe intérieure de consommation finale d'électricité : de 22 euros par mégawattheure (MWh) elle passe à 12 euros/MWh.

L'usage du numérique a souvent pour objectif de réaliser des économies d'énergie, par exemple dans les transports, les installations domotiques, le traitement des déchets...

Cependant, la chaleur générée par les *Data Centers* (et donc leur consommation en énergie) est si importante (4 % de la consommation énergétique mondiale en 2015) qu'ils nécessitent d'être refroidis par des systèmes de climatisation qui représentent 40 à 50 % de leur consommation.

Ce sont donc aussi de véritables gouffres énergétiques.

UTILISER DES DONNÉES STRUCTURÉES

p.75 :

DOCUMENT 4 Base de données

Une base de données est un moyen de recueillir et présenter des informations de façon **structurée**, généralement sous forme d'un tableau à plusieurs colonnes. Le nom des colonnes correspond aux **descripteurs** et les **valeurs** se trouvent dans le tableau. On peut y effectuer des recherches en sélectionnant un ou plusieurs descripteur(s).

Effectifs d'élèves des établissements du second degré public et privé sous tutelle du ministère en charge de l'Éducation nationale (2017-2018)

| | Année scolaire | Numéro d'établissement | Académie | Type d'établissement | Secteur d'enseignement | Nombre d'élèves | Libellé région |
|----|----------------|------------------------|----------|----------------------|------------------------|-----------------|----------------|
| 14 | 2017-2018 | 0290023T | RENNES | LYCÉE | PUBLIC | 535 | Bretagne |
| 15 | 2017-2018 | 0290022S | RENNES | LYCÉE | PUBLIC | 772 | Bretagne |
| 16 | 2017-2018 | 0561607T | RENNES | LYCÉE | PUBLIC | 943 | Bretagne |
| 17 | 2017-2018 | 0220060W | RENNES | LYCÉE | PUBLIC | 965 | Bretagne |
| 18 | 2017-2018 | 0292047T | RENNES | LYCÉE | PUBLIC | 396 | Bretagne |
| 19 | 2017-2018 | 0290062K | RENNES | LYCÉE | PUBLIC | 679 | Bretagne |
| 20 | 2017-2018 | 0290071V | RENNES | LYCÉE | PUBLIC | 959 | Bretagne |
| 21 | 2017-2018 | 0290076A | RENNES | LYCÉE | PUBLIC | 987 | Bretagne |
| 22 | 2017-2018 | 0290063L | RENNES | LYCÉE | PUBLIC | 160 | Bretagne |
| 23 | 2017-2018 | 0290008B | RENNES | LYCÉE | PUBLIC | 1 153 | Bretagne |
| 24 | 2017-2018 | 0290012F | RENNES | LYCÉE | PUBLIC | 1 697 | Bretagne |

<https://data.education.gouv.fr>

Dans l'exemple ci-dessus, on relève des informations sur différents lycées. Par exemple, en 2017-2018, le lycée n°0290012F de l'académie de Rennes avait un effectif de 1 697 élèves.

p.75 :

DOCUMENT 5 Le format CSV

Les données, avant d'être exploitées, peuvent être contenues dans des **fichiers CSV** (*Comma-separated values*). Dans ces fichiers, elles sont sous un format texte simple et séparées les unes des autres par un caractère (une virgule, un point-virgule, une tabulation). Il est ainsi possible d'exporter ou d'importer des données, par exemple un carnet d'adresses.

Ces fichiers peuvent être convertis sous forme d'un tableau nommé table de données. Chaque ligne du fichier CSV correspond à une ligne du tableau et les séparateurs à ses colonnes. Il devient alors possible d'organiser et de trier les données.

| | A | B | C | D | E |
|----|--------|-------------|------------|----------------------------------|----------------|
| 1 | NUM | INVENTAIRE# | NUM | ENREGISTREMENT# | SYMBOLE DANGER |
| 2 | 2# | Aucun | symbole | | |
| 3 | 115# | 189#C | - Corrosif | | |
| 4 | 47# | 63#X | - Irritant | | |
| 5 | 41# | 55#X | - Irritant | | |
| 6 | 410# | 391# | Aucun | symbole | |
| 7 | 13188# | 10160# | Xn | - Nocif | |
| 8 | 13188# | 10160# | N | - Dangereux pour l'environnement | |
| 9 | 419# | 471# | X | - Irritant | |
| 10 | 13173# | 10127# | F+ | - Extrêmement inflammable | |
| 11 | 505# | 571# | Aucun | symbole | |
| 12 | 513# | 579# | Aucun | symbole | |
| 13 | 497# | 561# | Aucun | symbole | |
| 14 | 569# | 659# | C | - Corrosif | |
| 15 | 915# | 1098# | C | - Corrosif | |
| 16 | 647# | 743# | C | - Corrosif | |
| 17 | 936# | 1129# | Aucun | symbole | |
| 18 | 767# | 882# | N | - Dangereux pour l'environnement | |
| 19 | 760# | 873# | X | - Irritant | |
| 20 | 760# | 873# | N | - Dangereux pour l'environnement | |
| 21 | 920# | 1105# | O | - Comburant | |
| 22 | 920# | 1105# | C | - Corrosif | |
| 23 | 847# | 1005# | C | - Corrosif | |
| 24 | 847# | 1005# | N | - Dangereux pour l'environnement | |
| 25 | 974# | 2180# | N | - Dangereux pour l'environnement | |

Trois formats de données très utilisés

- **JSON** : (prononcer *ji-son* ou *djay-son*) JavaScript Object Notation

Fonctionne avec un système de paire clé / valeur.

```
{
  "menu": {
    "id": "file",
    "value": "File",
    "popup": {
      "menuitem": [
        { "value": "New", "onclick": "CreateNewDoc()" },
        { "value": "Open", "onclick": "OpenDoc()" },
        { "value": "Close", "onclick": "CloseDoc()" }
      ]
    }
  }
}
```

- **XML** : Extensible Markup Language

Format utilisé pour l'échange automatisé entre système d'information

```
<menu id="file" value="File">
  <popup>
    <menuitem value="New" onclick="CreateNewDoc()" />
    <menuitem value="Open" onclick="OpenDoc()" />
    <menuitem value="Close" onclick="CloseDoc()" />
  </popup>
</menu>
```

- **CSV** : Comma Separated Values

Format de texte où chaque valeur est généralement séparée par une virgule (plus souvent un point-virgule en France, pour pouvoir différencier les virgules des décimaux et les virgules de séparation).

```
email,prenom,tel,creation_date,genre
johan@example.com,Johan,0682145672,2019-01-15,1
marie@example.com,Marie,0745632109,2019-03-10,2
amelie@example.com,Amelie,0675489125,2019-08-22,1
```

→ En TP, nous manipulerons un tel fichier qui contient près d'un million et demi de données !