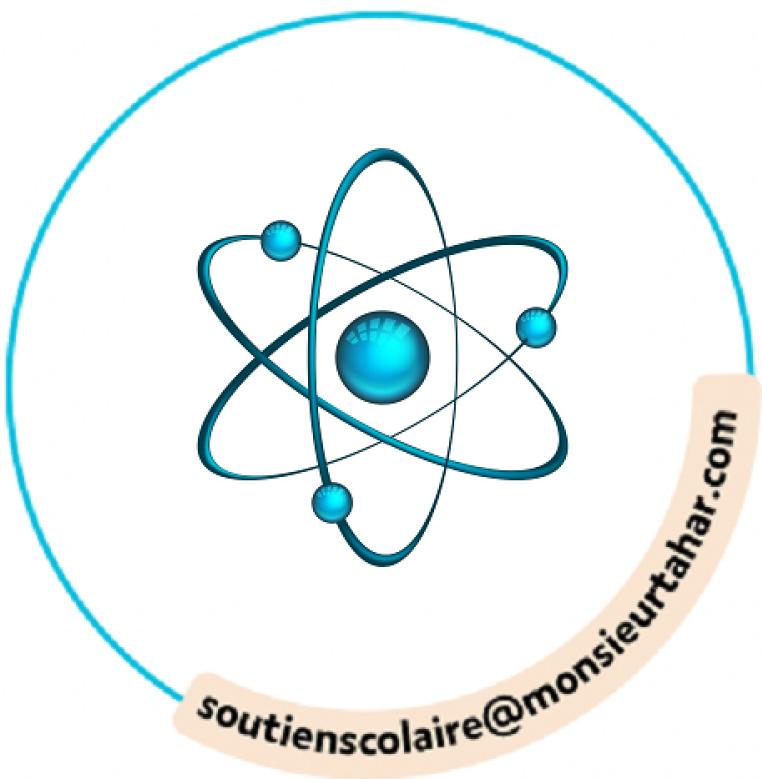


SNT



CHAPITRE 4

Cours • Les données structurées et leur traitement

1 Les données structurées

A La notion de donnée

Une donnée est une valeur décrivant un objet. On parle d'une donnée personnelle lorsqu'elle se rapporte à une personne. Plusieurs descripteurs peuvent être utilisés pour décrire un même objet. Une collection regroupe plusieurs objets partageant les mêmes descripteurs.



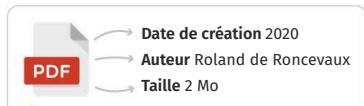
2 Carnet d'adresses avec plusieurs contacts

La structure de table permet de présenter une collection : les objets en ligne, les descripteurs en colonne et les données à l'intersection. Les données sont alors dites structurées.

Pour assurer la persistance des données, ces dernières sont stockées dans des fichiers. Une base de données regroupe plusieurs collections de données reliées entre elles.

B Les métadonnées

Les métadonnées sont des données associées à des fichiers et qui permettent d'en décrire le contenu. Elles sont utilisées pour résumer les informations élémentaires afin d'en faciliter la gestion et les traitements.

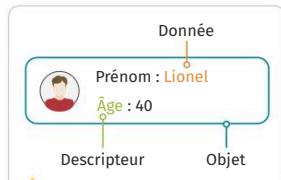


3 Exemple de métadonnées

C Le traitement des données et des métadonnées

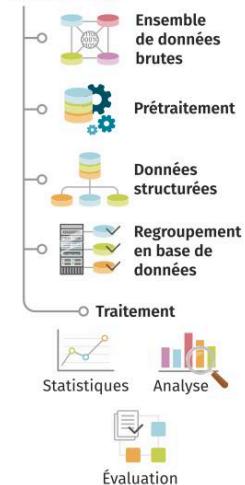
Les données comme les métadonnées peuvent être collectées et enregistrées automatiquement par des algorithmes ou bien renseignées par une personne.

De nombreux traitements, sur une ou plusieurs collections, sont possibles : l'indexation, les tris, la recherche, les filtres, les calculs, la visualisation, etc. Ces opérations permettent notamment de réaliser des statistiques, d'étudier des évolutions, etc.



1 Données liées à un contact

Forage de données



4 Forage de données ou data mining

Vocabulaire

CSV : Comma-Separated Values, ou données séparées par des virgules, désigne un format de fichier simple permettant d'enregistrer une table. Les fichiers .csv peuvent être ouverts par plusieurs logiciels différents (tableur, traitement de texte, etc.).

2 Le stockage des données

A Le stockage à l'aide du matériel personnel

Les fichiers de données sont stockés sur des supports de stockage, papier ou informatiques : ces stockages informatiques peuvent être internes (disque dur ou disque SSD) ou externes (disque dur externe, clé USB, CD-ROM, DVD-ROM, carte SD, etc.). Il est nécessaire de réaliser des sauvegardes régulières et sur différents supports car les données peuvent facilement être corrompues ou détruites.

B Le stockage dans le cloud

Le stockage en ligne, ou dans le *cloud*, consiste à placer des documents et des fichiers dans un espace accessible via Internet. Toutes ces données sont sauvegardées sur des serveurs distants et il est possible d'y avoir accès depuis n'importe quel appareil connecté à Internet.

3 Les évolutions sociétales liées aux métadonnées

A Le Big data

L'évolution des besoins du numérique, de traitement et de diffusion des données contribue à une surabondance d'informations qui nécessite le développement de nouveaux algorithmes capables de les exploiter et de machines de plus en plus performantes pour les stocker.

L'exploitation massive de données est en plein essor dans des domaines aussi variés que les sciences, la santé ou encore l'économie. Les conséquences sociétales sont nombreuses tant en termes de confidentialité, de propriété que de libertés individuelles.

B L'open data

Certaines de ces données sont dites ouvertes (*open data*), mais de nombreuses données à caractère personnel suscitent la convoitise d'entreprises qui les collectent et les revendent. D'où l'importance d'un cadre juridique permettant de protéger les usagers, préoccupation à laquelle répond le Règlement Général sur la Protection des Données (RGPD).

C L'impact environnemental

Les centres de données (*data centers*) nécessitent des ressources en eau pour le refroidissement des machines, en électricité pour leur fonctionnement et en métaux rares pour leur fabrication. Ils sont une source de plus en plus importante de pollution et contribuent au rejet de dioxyde de carbone (CO_2) dans l'atmosphère.

Avec l'évolution des usages du numérique, ce phénomène est en progression : la consommation électrique augmente de 3 % chaque année dans le monde.



Disquette

~ 1 Mo
1967



CD-ROM

~ 1 Go
1982



Clé USB

~ 100 Go
1999

~ 1 To

2010



Disque dur externe

5 Évolution des moyens de stockages physiques