

Thème 4 – Les données structurées et leur traitement

Je fais le bilan

Les notions à retenir

Traitement des données

Le numérique a permis de collecter, d'agréger et d'analyser automatiquement de grandes quantités d'informations. Alors que notre monde est de plus en plus rempli de données, il est nécessaire d'avoir la capacité d'évaluer de manière critique les informations. Ainsi, la datavisualisation est utile pour aider les humains à comprendre rapidement de grandes quantités de données et constitue un outil de communication utile lors de la présentation de résultats. Cependant, toutes les visualisations ne sont pas égales et dans de nombreux cas, le type de visualisation utilisé peut même induire le lecteur en erreur.

Le cloud

Le cloud permet de stocker les données en ligne plutôt que sur son propre disque dur. Au quotidien l'un des plus gros avantages est de pouvoir accéder à ses fichiers partout dans le monde, tant qu'il y a une connexion internet. Il existe de multiples utilisations du cloud, depuis le stockage de grandes quantités de données jusqu'à l'exécution de programmes informatiques sur plusieurs appareils à la fois. Il ne faut pas oublier que cela nécessite un énorme volume de centres de données.

Big Data

« Big Data » est une expression utilisée pour désigner un volume énorme de données, structurées et non structurées. Ces données sont si nombreuses et variées qu'il est difficile de les traiter à l'aide de techniques de base de données et de logiciels traditionnels. On nécessite des algorithmes puissants, permettant de les traiter rapidement pour un usage en temps réel.

Repères historiques

1930 : Utilisation des cartes perforées, premier support de stockage de données.

1956 : Invention du disque dur permettant de stocker de plus grandes quantités de données.

1970 : Invention du modèle relationnel (E. L. Codd) pour la structuration et l'indexation des bases de données.

1979 : Création du premier tableur, VisiCalc.

2009 : Open Government Initiative du président Obama pour l'ouverture des données publiques.

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

Définitions

Descripteur : élément qui décrit et caractérise les données d'un fichier.

Métadonnées : données qui permettent de décrire des données.

Open Data : données publiques brutes, librement accessibles et réutilisables.

DataViz : désigne les techniques de visualisation des données afin d'en faciliter la compréhension et/ou l'analyse.

Intelligence artificielle : une machine ou un algorithme capable de simuler le comportement et l'intelligence des humains.