

Data, données, informations, connaissance

A quoi servent les data ?

Véronique Pelletier,
membre du Comité d'ADELI

Les data, données en français, sont des éléments numériques, représentant une partie du monde réel. L'« afflux » de données, ces dernières années, a contraint les professionnels, à cause des réseaux sociaux, des courriels, des tweets, des articles publiés, etc., à inventer de nouveaux concepts pour traiter ces masses gigantesques de données. On parle aussi, de gouvernance des données.

A quoi servent-elles ? Les acteurs comme Facebook, Google ne s'y sont pas trompés ; ils stockent, classent, relient, croisent nos données entre elles, pour savoir qui nous sommes, comment nous vendre de nouveaux produits et/ou services. La connaissance de leurs clients est fondamentale, et les autres entreprises commencent à prendre la mesure de l'importance des données pertinentes.

Le Big Data est né, difficilement traduisible. Les ontologies permettent de relier les concepts entre eux et sont aux données ce que la grammaire est au langage.

Big Data

Peut-on dire que le Big Data est un terme fourre-tout ?

A-t-on, en réalité, toujours eu les mêmes problématiques ?

La difficulté est-elle liée à l'accélération du temps et de l'espace, lié aux nouvelles technologies ?

L'article de Kim Estivalet nous fournit beaucoup d'informations, de références autour du Big Data.

Est-ce un nouveau concept marketing ?

Peut-on le démystifier pour le rendre accessible au plus grand nombre ?

Ne cherche-t-on pas à créer un système mondial, unique qui permettrait toute interrogation, toute création de valeur, toute collecte d'informations issues de systèmes et de protocoles différents ?

L'amélioration d'Internet aujourd'hui passe par la compréhension du monde, par le Web sémantique. Les acteurs de ce secteur seront des acteurs majeurs.

Ontologies

Aristote a décrit dans une branche de la philosophie que l'on appelle ontologie, l'étude de l'être. Le monde des objets vient de là. Les objets liés entre eux dans un réseau sémantique apportent de la connaissance.

L'article de Pierre Jourdan nous présente un cheminement entre data, information, connaissance en s'appuyant sur le référentiel ITIL.

Les datacenters

Pour traiter avec rapidité et en toute sécurité les gigantesques volumes de données de leurs clients, des entreprises construisent de puissantes usines appelées « datacenters ».

L'article rédigé par Alain Coulon brosse ce nouveau paysage informatique et indique les bases sur lesquelles repose le fonctionnement durable de ces nouveaux monstres.

Bases de données

On a structuré l'information dans des fichiers, puis des bases de données, on a créé des méta modèles permettant de gérer les informations structurées, on a créé des systèmes de gestion de base de données gérant leur évolution dans le temps.

Puis on a créé des bases de données multidimensionnelles permettant d'agréger, de consolider les données selon des axes.

Mais l'information pertinente, il faut la trouver au milieu des articles publiés, des sites Web, des courriels, des fichiers,...

C'est pourquoi, il faut au-dessus des bases de données un système de systèmes capable de se « plugger » à n'importe quelle base de données, n'importe quel protocole (HTTP, SOAP, SMTP...) pour collecter l'information pertinente.

Pour quoi en faire ?

Pour alimenter les services Web, pour faire des statistiques, pour faire du prédictif (en poursuivant la droite de non-régression), pour aider à la décision, pour aligner les services sur la stratégie de l'entreprise... ▲

veronique.pelletier@adeli.org