ESIEE Paris : 2 ème année d'école de management, dans le cadre d'un cours de BioStatistiques animé par Bertrand Roudier .

Les activités humaines, la vie en société, génèrent des données et des informations. Il en a toujours été ainsi. Plus les sociétés se développent, plus le volume créé est important.

Depuis une trentaine d'années, l'arrivée de l'informatique a bouleversé notre rapport à l'information, tant dans notre quotidien que dans le monde de l'entreprise.

Le mot informatique, issu d'une contraction d' « information » et d' « automatique », désigne la technique permettant de traiter cette information de manière automatique. D'abord confiné à un public d'experts et de scientifiques, qui y voyait un fabuleux outil pour calculer plus rapidement, stocker et pour échanger des informations, l'informatique, aujourd'hui, se démocratise.

Trois facteurs convergent et participent à rapprocher données et utilisateurs :

- l'augmentation des capacités de calcul et de stockage,
- l'accessibilité accrue de l'information, en particulier grâce aux réseaux haut-débits,
- l'augmentation de l'offre (de production d'information) et la démocratisation des outils (logiciels, de traitements) facilitée par le monde de l'open source

Mais, l'information n'est pas connaissance. Et plus le volume de données auquel nous avons accès augmente, plus il est difficile de les trier, de les organiser, et finalement de les analyser.

En considérant la distance qui sépare les données brutes de la production de connaissances, ce cours examine quelles sont les grandes étapes, les outils méthodologiques et techniques, qui jalonnent ce chemin.

Cours 1 - Introduction, objectifs et principales définitions

Cours 2 - Du questionnement à la délinéation d'un champ (Aurélie Delemarle , Lionel Villard)

Cours 3 - Introduction aux bases de données relationnelles (part 1)

Cours 4 - Introduction aux bases de donnees relationnelles (part 2)

Cours 5 - De l'extraction à la structuration des données

Cours 6 - Homogénéisation de données et projet