

QUELLE SOCIÉTÉ NUMÉRIQUE ?

Vidéo

Les enjeux du Big Data - 14,35 min

foucherconnect.fr / 19snt10

Big Data ou mégadonnées

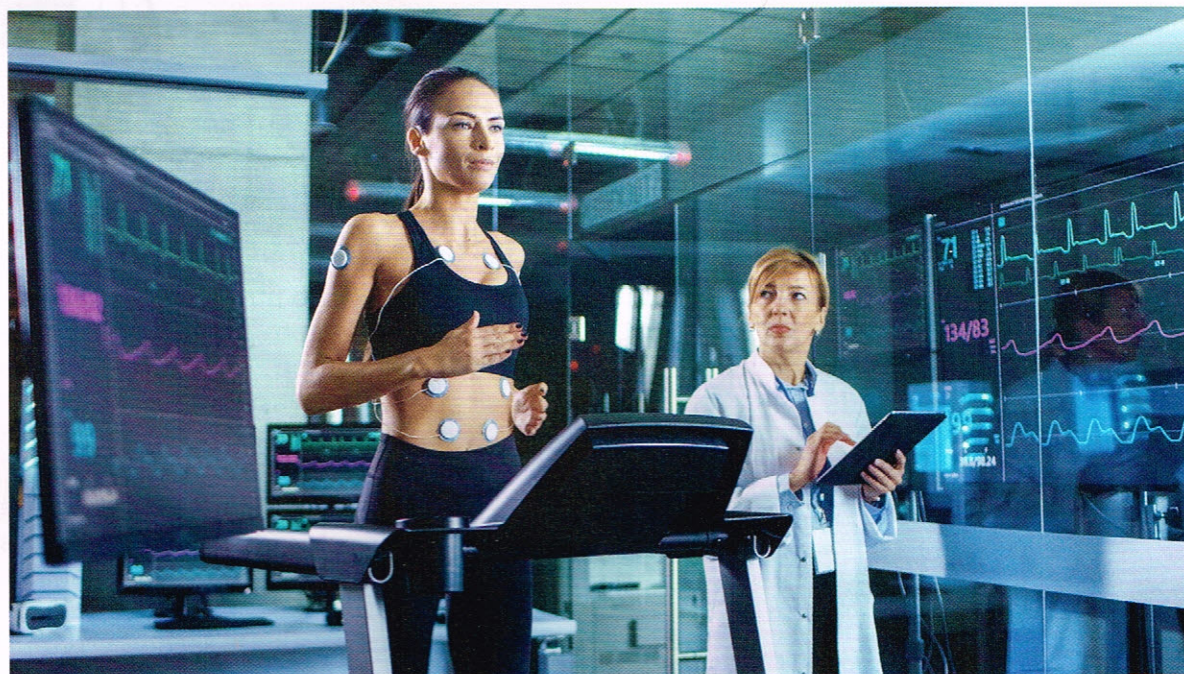
L'explosion quantitative des données numériques a obligé les chercheurs à trouver de nouvelles manières de voir et d'analyser le monde. Il s'agit de découvrir de nouveaux ordres de grandeur concernant la capture, la recherche, le partage, le stockage, l'analyse et la présentation des données. Ainsi est né le « **Big Data** ». Il s'agit d'un concept permettant de stocker un nombre indicible d'informations sur une base numérique. [...]

Le Big Data, c'est quoi ?

Littéralement, ces termes signifient mégadonnées, grosses données ou encore **données massives**. Ils désignent un ensemble très volumineux de données qu'aucun outil classique de gestion de base de données ou de gestion de l'information ne peut vraiment travailler. En effet, nous procréons **environ 2,5 trillions d'octets de données tous les jours**. Ce sont les informations provenant de partout : messages que nous nous envoyons, vidéos que nous publions, informations climatiques, signaux GPS, enregistrements transactionnels d'achats en ligne et bien d'autres encore.

Source www.lebigdata.fr

Le Big Data au service du sport



Le *Sports Analytics* consiste à collecter des données sur les performances sportives et à les analyser afin d'effectuer des prédictions ou d'améliorer les performances en question. Cette pratique est de plus en plus courante dans toutes les disciplines sportives. [...]

Grâce à l'analyse de données, il est désormais possible de prédire les résultats de compétitions sportives ou de trouver des points à améliorer pour permettre aux sportifs de réaliser de meilleures performances. Par exemple, l'équipe d'Allemagne de football a exploité le *Big Data* pendant la Coupe du Monde 2014 et l'Euro 2016 pour développer sa stratégie [...].

Dans tous les sports, la *Data Science* est de plus en plus utilisée. Les données peuvent par exemple être obtenues grâce à des capteurs [...] ou encore en utilisant des technologies d'analyse vidéo.

Source www.lebigdata.fr, 15 juin 2018.

Des données de 29 millions de comptes récupérées par des pirates

Des pirates ont récupéré des informations personnelles issues de 29 millions de comptes Facebook, a annoncé vendredi 12 octobre 2018 le réseau social. [...]

Il reconnaît aujourd'hui que le piratage a bel et bien fait des victimes. Dans le détail, les pirates ont eu accès aux noms et aux informations de contact (numéro de téléphone ou e-mail) de 15 millions de personnes. Pour 14 millions de personnes supplémentaires, ce sont les données telles que « le nom d'utilisateur, genre, langue, statut marital, religion, ville d'origine, ville actuelle, date de naissance, les informations relatives à l'éducation et au travail, les dix derniers endroits où un utilisateur s'est géolocalisé ou a été identifié, les pages qu'il suit, et les quinze recherches les plus récentes » qui ont été récupérées par les pirates.

Source Damien Leloup et Martin Untersinger, www.lemonde.fr, 12 octobre 2018.

Les sondages et le *Big Data*

À seulement quelques mois d'intervalle, les résultats de deux votes ont fait mentir les sondages qui les ont précédés : la sortie du Royaume-Uni de l'Union européenne et l'élection de Donald Trump. Dans le cas de la présidentielle américaine, le *New York Times*, le *Huffington Post* et le site FiveThirtyEight, estimaient respectivement la probabilité qu'Hillary Clinton l'emporte à 85 %, 98 % et 71 %. Les traitements analytiques de grands volumes de données, sur lesquels ces médias se sont reposés pour réaliser leurs projections, ont montré ici leurs limites.

La qualité des données est primordiale

Le cas de cette élection a notamment mis en lumière deux mythes du *Big Data*. Tout d'abord, les *data scientists* émettent des probabilités et non des certitudes. [...] C'est la qualité des données qui va définir l'intérêt et la pertinence des analyses. [...]

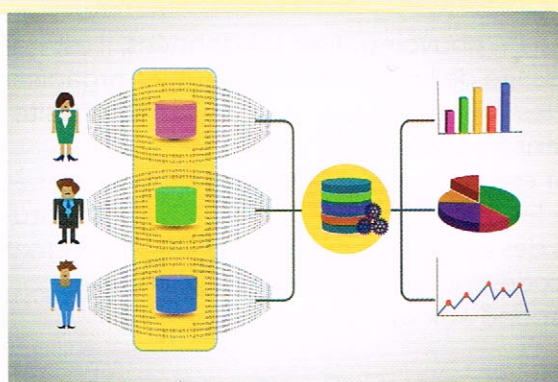
Se méfier de la pertinence des algorithmes

Mais la qualité des données à elle seule ne fait pas tout. Les modèles algorithmiques doivent également rester pertinents. Leur longévité, voire les algorithmes eux-mêmes, doivent être constamment remis en question pour prendre en compte les changements de notre environnement politique, économique, sociétal, technologique... [...]

Le *Big Data* n'anticipe pas les émotions

[...] Il s'agit d'ailleurs de l'un des défis que les entreprises doivent relever en matière de traitement de données. Les outils analytiques ne permettent en effet pas d'anticiper les émotions et les prises de décision qui en découlent.

Source Eddye Dibbar, <https://bfmbusiness.bfmtv.com>, 9 février 2017.



Pistes de recherches



RECHERCHE
D'INFORMATIONS

<1> Listez les données que vous produisez chaque jour, à titre personnel. Constituez ensuite un groupe et classez les données recensées en suivant des critères que vous aurez définis. Commentez les types et la quantité de données produites par jour.



DISCUSSION

<2> À partir de vos recherches, listez les avantages et les inconvénients que peut offrir l'utilisation de l'intelligence artificielle pour le traitement des données.



EXPOSÉ ORAL

<3> Recherchez un sondage qui s'est avéré inexact, puis analysez les éléments qui ont pu en fausser les résultats.