

Données numériques : les enjeux du Big Data

- Comprendre -

Date de mise en ligne : lundi 12 juin 2017

Académie de Versailles - Education aux Médias - Licence « Creative Commons BY-SA »

Big data, Data mining, Open data, data center, autant de termes derrière lesquels se cache une même entité : la donnée numérique. La société actuelle est de plus en plus connectée, et entraîne une explosion du volume des données produites. Pouvoir les collecter, les stocker et les analyser constitue aujourd'hui un enjeu économique et sociétal majeur.

Cet article a été initialement diffusé dans le numéro de mars du journal Epicure (version collègue en page 12 et version lycée en page 12)

Big data, Data mining, Open data, data center, autant de termes derrière lesquels se cache une même entité : **la donnée numérique**. La société actuelle est de plus en plus connectée, et entraîne une explosion du volume des données produites. Pouvoir les **collecter**, les **stocker** et les **analyser** constitue aujourd'hui un enjeu économique et sociétal majeur.

Le vocabulaire des données numériques

- **Big data** : ensembles de données si volumineux (mégadonnées) que des outils de base de données classiques ne sont plus en mesure de les traiter et nécessitent d'autres technologies.
- **Data mining** : exploration des données pour en déduire un savoir et construire des modèles.
- **Open Data** : pratique de publication sous licence ouverte qui garantit un accès libre aux données numériques et autorise leur réutilisation sans conditions techniques, juridiques ou financières.
- **Data center** : un site physique sur lequel se trouvent regroupés des équipements constituant le système d'information

De la description à la prédiction

Le nombre d'objets connectés et d'applications est en forte augmentation. Ils sont le canal privilégié du recueil des données. et du traçage des activités humaines. C'est devenu une véritable industrie qui intéresse tous les secteurs d'activités et dont la matière première est la « data ». La valeur ajoutée créée repose sur la capacité à **analyser ces données pour les exploiter**.

Aujourd'hui, des outils de manipulation de données, permettent de traiter, en temps réel et à grande échelle, le « big data ». On peut maintenant interpréter les données pour **prévoir avec une grande probabilité de futurs comportements**. Demain, les Intelligences Artificielles développées par Google, Facebook ainsi que de nombreuses start-up offriront encore davantage de possibilités d'analyse et de valorisation des données. L'exploitation intelligente d'information à des fins prédictives est devenue un véritable enjeu économique et de compétitivité pour les entreprises.

Un défi à relever : la sécurisation des données

Le terme « donnée » laisse croire que les informations sont **transmises volontairement** par les utilisateurs. Ce n'est que partiellement vrai : certaines données, grâce au système de cookie par exemple, sont récoltées et traitées à **l'insu de l'utilisateur** et, parmi elles, des informations personnelles que l'on ne souhaite pas partager. La multiplication des **cyberattaques** ayant pour conséquence la **divulgaration d'informations** censées rester privées pose la question de la sécurité des données, notamment leur stockage et leur accès.

Bien que des solutions techniques apparaissent régulièrement, ce type de réponse ne suffit pas. Un cadre juridique universel est à définir pour garantir les mêmes droits quel que soit l'hébergeur par lequel nos données transitent. Les législateurs européens se sont d'ores et déjà emparés de ces questions.

En attendant, il est important que chacun soit vigilant lors de ses navigations sur le web en prenant soin :

- de consulter les chartes et mentions légales des sites visités pour évaluer leur niveau de confiance ;
- de naviguer en fenêtre privée et de supprimer ses historiques de navigation ;
- de paramétrer son ordinateur pour régulièrement supprimer les cookies et/ou limiter leur utilisation ;
- de choisir le type d'hébergement de ses données.

Test de positionnement : Big Data

Question 1 : Pour toi, le « big data » est une source de :

- c Progrès.
- `) Méfiance.
- f Danger.

Question 2 : Que penses-tu de l'expression, « Big Data peut devenir Big Brother » prononcée en 2013 par Fleur Pèlerin, alors ministre déléguée à l'économie numérique ?

- f Pourquoi en parler au futur ? Toutes nos activités sont déjà tracées.
- c On ne parle que des dangers potentiels du « big data ». On en met pas suffisamment en avant ses bénéfices.
- ` Il faut effectivement rester vigilant pour ne pas tomber dans la surveillance de masse.

Question 3 : On te propose de participer à un programme donnant accès à l'ensemble de tes données médicales à des fins de diagnostic précoce. Quelle est ta réaction ?

- f Hors de question de donner accès à des données personnelles.
- ` Tu vérifies que les données sont anonymisées avant de donner ton accord.
- c Tu acceptes immédiatement, c'est une chance.

Question 4 : Selon toi, quel film (ou série) exprime le mieux ta représentation de l'intelligence artificielle ?

- ` « Person of Interest », avec les bonnes intentions, l'intelligence artificielle permet de sauver des vies.
- f « Matrix » : L'intelligence artificielle prend le pouvoir sur l'Homme.
- c « AI » : L'intelligence artificielle rend service à l'Homme et peut éprouver les mêmes sentiments.

Question 5 : Amazon vient de lancer un nouveau type de magasin sans caisse dans lequel on entre, on se sert et on ressort sans payer. Tout est géré par une intelligence artificielle, la facturation et le paiement sont automatiques. Que

penses-tu du concept ?

c Tu trouves ça super. Plus de perte de temps pour faire ses courses.

f Nous ne sommes malheureusement plus dans de la science-fiction.

` L'idée est séduisante, mais quelles sont les données à renseigner et utilisées pour un tel service ?

Réponses

Tu as un maximum de c. « big data » = progrès

Pour toi, le « big data » est source de progrès. Tu apprécies les services individualisés qui te facilitent la vie au quotidien, même si ça signifie transmettre tes données personnelles.

Tu as un maximum de `. « big data » = méfiance

Selon toi, les technologies liées au « big data » offrent des perspectives intéressantes à condition de conserver la maîtrise de ses données personnelles.

Tu as un maximum de f. « big data » = danger

Pour toi, le « big data » nuit aux libertés personnelles. Tu redoutes l'utilisation par des tiers de tes données personnelles à des fins de surveillance ou commerciales.

Post-scriptum :

Crédit logo : [luckey_sun](#) via Flickr