Melvyn MUTHS

Benjamin HOHN

Les risques du Big Data :

Le Big Data apporte de nombreux bénéfices aux entreprises, mais représente également un risque de sécurité pour la confidentialité. Les milliers de données partagées par les individus exposent leur vie privée plus que jamais auparavant. Les données personnelles valent de l'or pour les marketeurs, les institutions financières, les employeurs ou les gouvernements. Les conséquences peuvent être désastreuses pour les individus, qui peuvent par exemple se voir refuser un emploi ou un crédit à cause de leurs données.

La collecte de données :

Pour récupérer des données en grande quantité on utilise un système automatisé. Ainsi, les outils pour récupérer les données des réseaux sociaux fleurissent, les outils de récupération de logs sur le web aussi et finalement tous les capteurs permettent aujourd'hui de récupérer des masses de données de manière extrêmement rapide.

Source : https://www.stat4decision.com/fr/big-data-cest-quoi/

Analyse des données :

Une fois les données préparées (ce qui est le plus gros du boulot), un autre terme très utilisé apparaît, il s'agit du machine learning (apprentissage automatique). Il s'agit en fait d'algorithmes permettant d'explorer les données massives en « apprenant » avec les données.

Source : https://www.stat4decision.com/fr/big-data-cest-quoi/

Exploitation des données :

L'exploitation de cette nouvelle masse d'informations et l'accélération de leur flux nécessite des capacités de calcul importantes et un espace de stockage adapté dont peu d'entreprises peuvent disposer en raison des équipements nécessaires mais surtout en raison de leur coût de possession. Le Cloud Computing a permis de rendre accessible à ces entreprises une technologie d'exploitation des données efficace et à moindre coût grâce à la mutualisation des moyens dans le Cloud.

Source : https://www.journaldunet.com/solutions/expert/54620/big-data---l-exploitation-des-donnees-au-c-ur-des-preoccupations-des-dsi.shtml

Les différents risques :

Source : https://www.lebigdata.fr/confidentialite-risques-big-data

- **L'automatisation de la discrimination :** Il y a 5 ans, l'EPIC estimait que l'utilisation des analyses prédictives dans le secteur public et le secteur privé peuvent permettre aux gouvernements et aux entreprises d'évaluer la capacité d'une personne à obtenir un emploi ou un crédit. Or, cette utilisation nuit directement à la liberté d'association. Depuis lors, les choses n'ont fait qu'empirer. La discrimination est illégale, mais l'automatisation des prises de décision la rend difficile à prouver. En somme, le Big Data tend à automatiser la discrimination. Malgré les avancées réalisées dans le

domaine du Big Data, les lois sur la protection n’ont pas évolué.

- **Une augmentation des fuites de données :** Suite aux nombreuses fuites de données de géants du commerce comme Target, Home Depot, ou de sites comme eBay, ayant impacté des dizaines de millions d’individus, le public est très alerte quant aux fraudes de cartes de crédit et aux usurpations d’identité. Hélas, le risque reste toujours très élevé, notamment à cause de l’essor de l’internet des objets. Désormais, de très nombreux meubles sont connectés dans nos foyers, au même titre que les voitures et autres accessoires que nous portons au quotidien. Le nombre cibles potentielles pour les hackers est décuplé.

- **La fin de l’anonymat et de la confidentialité :** De nos jours, il est de plus en plus difficile de faire quoi que ce soit sans que notre identité soit associée à nos actions. Même les données « désidentifiées » représentent un risque pour la confidentialité. Les standards de sécurité utilisés il y a encore un an ou deux ne sont plus suffisants. Les entreprises qui souhaitent rendre les données anonymes sont confrontées à une difficulté croissante. Il sera bientôt impossible d’empêcher les données de pouvoir être à nouveau associées aux individus. En plus d’être vulnérables aux fuites, les objets connectés sont de véritables machines à collecter les données les plus personnelles des utilisateurs. Les constructeurs peuvent changer les conditions de confidentialité à tout moment, et il est difficile de convaincre un utilisateur d’arrêter d’utiliser sa télévision connectée ou sa voiture connectée suite à un tel changement.

- **Les abus gouvernementaux :** Selon l’EPIC, le nombre de bases de données gouvernementales américaines est plus élevé que jamais auparavant. Le FBI par exemple collecte des données personnelles comme le nom, les pseudonymes, l’origine ethnique, le genre, le lieu et la date de naissance, le numéro de sécurité sociale, le numéro de passeport, l’adresse, les numéros de téléphone, les photos, les empreintes digitales, les numéros de compte bancaire, et l’emploi des citoyens. En effet, l’agence s’est elle-même émancipée du Privacy Act de 1974.

- **La vente aux enchères de données :** De nombreuses entreprises collectent et vendent les données des utilisateurs, permettant d’établir des profils d’individus. Les entreprises peuvent désormais savoir si une femme est enceinte, si une personne est homosexuelle ou si elle est atteinte d’un cancer avant même qu’elle ne le révèle à ses proches. Aucune loi ne protège réellement les consommateurs contre de tels agissements. Jusqu’à ce que des lois pour la protection de la confidentialité soient votées, ces pratiques continueront.

Quelle est la place du RGPD pour le big data ?

Droit à l'oubli, portabilité des informations: les nouveautés introduites par le RGPD Autre nouveauté, le RGPD a consacré de nouveaux droits au profit de la personne concernée qui peuvent constituer des enjeux pour le responsable de traitement en cas de traitement de type big data, à savoir:

• la consécration au niveau européen du droit à l'oubli; • l'introduction de la portabilité des informations.

En effet, lorsque le responsable du traitement a partagé les données d'une personne dans le cadre d'un traitement big data et si cette personne exerce son droit à l'oubli, le responsable de traitement doit prendre toutes les mesures raisonnables en vue d'informer les tiers qui traitent les données publiées sous sa responsabilité afin qu'ils procèdent à l'effacement de tous liens vers ces données et de toute copie ou reproduction de ces données. Une liste précise des personnes avec qui les données sont partagées devra donc être tenue par le responsable de traitement.

Quant au droit à la portabilité des données, il implique pour le responsable du traitement de prévoir la

possibilité de restituer l'ensemble des données à caractère personnel d'une personne concernée dans un format électronique structuré et couramment utilisé.

Pour les mettre en oeuvre, il convient d'anticiper et de respecter les obligations nouvelles du RGPD de façon à traiter les volumes de données traitées en toute légalité.

Source :
https://www.daf-mag.fr/Thematique/reglementation-1243/breves/tribune-quelle-place-big-data-ere-rgpd-328415.htm

On remarque donc qu'il y a de nombreux risques concernant l'utilisation du big data, ces risques sont particulièrement importants pour ses utilisateurs notamment sur le respect de la vie privée et des données personnelles

Glossaire : - EPIC : Etablissement public à caractère industriel et commercial

From:
https://wiki.sio.bts/ - **WIKI SIO : DEPUIS 2017**

Permanent link:
https://wiki.sio.bts/doku.php?id=3a

Last update: **2020/07/26 16:27**