

## Nos données personnelles, le Big Data

### LE BIG DATA, C'EST QUOI ?

Le Big Data c'est un peu la révolution du XXI<sup>e</sup> siècle version «études de marché». Ou comment exploiter la quantité d'informations immatérielles que les consommateurs laissent traîner un peu partout (données GPS, clics, likes, vidéos visionnées, commentaires, etc.).



On ne peut désormais ignorer que le contrôle de la donnée personnelle est un enjeu stratégique tant pour l'individu dont elle constitue un élément de l'identité numérique, que pour l'entreprise qui ne peut plus raisonnablement envisager de se développer sans la capter pour l'exploiter et la valoriser.

La question du contrôle des données est donc cruciale pour l'individu car la quantité des données qui sont en permanence

collectées, traitées et stockées sur chaque personne permet d'aller très loin dans sa connaissance et contribue à un profilage très précis avec la possibilité - en fonction des catégories de données traitées - de produire des modèles probabilistes pour en apprendre davantage sur ses croyances religieuses, ses opinions politiques, son mode de vie, son orientation sexuelle, ses goûts et habitudes et bien d'autres aspects de sa vie personnelle et intime.

La question de la confiance devient encore plus fondamentale pour l'entreprise lorsqu'il est question de bâtir une relation durable avec ses clients, afin de répondre à des doutes ou des craintes qui pourraient freiner le développement de ce qui semble être appelé à devenir un levier majeur de création de valeur, tant pour les citoyens que les entreprises.



## LE BIG DATA ET LA CULTURE

Tout connaître de son public pour coller au plus près à ses désirs : le fantasme de l'industrie culturelle se réaliserait-il enfin grâce à la maîtrise du Big Data ?

En fait, dans le secteur de la culture, le big bang du Big Data a déjà commencé, et les premières applications donnent le vertige ! A New York, des chercheurs de l'université Stony Brook ont conçu un algorithme capable de prévoir avec 84 % d'exactitude le succès d'un livre.

Dans le même temps, l'application musicale Shazam étudie différentes techniques pour recouper les résultats de recherche de ses 300 millions d'utilisateurs, et prédire ainsi qui seront les stars de demain.

Les données personnelles du public, garantes du succès commercial d'une œuvre ?

De quoi faire rêver les acteurs d'un secteur de la culture dont l'économie souffre encore du passage à la civilisation numérique. La nature même du Big Data étant de donner des probabilités, les maisons d'édition peuvent par exemple savoir - via les ebooks - comment leurs lecteurs se comportent, quelles pages les ont intéressés ou pas, ce qui pourrait amener l'auteur à écrire dans telle ou telle direction !...



## La dictature de la demande !

En 2013, la maison d'édition britannique Penguin a ainsi lancé un nouveau format, les «essais courts», après avoir constaté que ses lecteurs d'ebooks ne finissaient que très rarement ce type d'ouvrages.

Autrement dit, l'auteur se place désormais directement au service de son lecteur... De quoi entrevoir un futur dans lequel le public serait le seul prescripteur digne de ce nom, la création artistique n'étant plus alors que l'objet désincarné d'une dictature de la demande, et à forcer ainsi l'industrie à ne sortir que des produits présentant zéro risque.



## Et la création dans tout ça ?

Ce qui est inquiétant, c'est que les gens qui ont entre les mains ces nouvelles techniques d'exploitation des data ne les maîtrisent pas vraiment.

La tentation du recours massif au Big Data ne risque-t-elle pas en effet de « tuer dans l'œuf » toute la richesse culturelle, dans la tentative constante d'adaptation de la culture aux goûts du public, et de créer une industrie capable de fournir quelques chefs d'œuvre au milieu d'une offre ultra-formatée ?

Ils ne se rendent pas compte qu'ils ne font que créer de l'existant et vont un peu à l'encontre de la créativité.

Car la création doit s'inscrire dans quelque chose de beaucoup plus vaste, dans une sorte de « rupture qu'on ne peut pas toucher du doigt » avec de simples statistiques.

## Les artistes en résistance contre le Big Data

L'emprise des données numériques sur la création devrait aussi donner naissance à un nouvel art de résistance, avec des artistes plus confidentiels qui luttent contre l'emprise du Big Data en détournant ses codes, quitte à passer pour réactionnaires, technophobes ou élitistes auprès du public. La liberté de créer est peut-être à ce prix, même si le Big Data ne remettra jamais en question notre besoin de création, qui demeure un processus essentiellement humain.

Le Big Data appliqué à l'art ouvre en tout cas des perspectives vertigineuses en ce qui concerne la relation entre l'artiste et le public. Aux États-Unis, un designer travaille ainsi à la mise au point d'une application capable de répertorier les rêves de ses utilisateurs, le genre d'outil susceptible d'offrir une belle matière aux futurs artistes du Big Data... Et d'inspirer, dans le même temps, un art de résistance.

Concernant le Big Data dans le domaine de la culture la guerre a déjà commencé, et elle promet d'être homérique !

## LE BIG DATA ET LES OPERATEURS TELECOMS

Si l'industrialisation de l'analyse des données est un phénomène relativement récent, elle prend de plus en plus d'ampleur dans les télécoms françaises, même si elles ont un temps de retard sur certains de leurs concurrents étrangers.

Le marché global du Big Data devrait frôler les 190 milliards de dollars en 2019 (source : International Data Corporation). La part détenue par les opérateurs télécoms reste néanmoins marginale.

Pourtant les opérateurs télécoms sont assis sur un véritable tas d'or : les données personnelles de leurs clients. Smartphone en poche, le consommateur transmet de précieuses informations en permanence : adresse, déplacements, coordonnées bancaires, contacts téléphoniques, temps passé sur Internet... Les opérateurs savent tout, ou presque, de la vie de leurs clients connectés via leur smartphone ; et même s'ils font partie des acteurs les mieux placés dans le domaine pour collecter des informations ils restent très prudents, notamment pour respecter les contraintes juridiques.



Car comme pour tout gisement, il faut savoir l'exploiter !

Les données ont besoin d'être affinées, retraitées, pour pouvoir être utilisées en interne ou revendues, tout en respectant un environnement juridique national et européen contraignant, visant à protéger la vie privée des consommateurs : les données utilisées doivent être rendues anonymes et agrégées, elles ne doivent pas pouvoir être associées à une personne en particulier.

### L'utilisation des données

Les opérateurs télécoms utilisent d'abord ces données en interne, à des fins de marketing. «Le but est de proposer à nos clients l'offre la plus adaptée au mode de consommation, et de le faire au bon moment», explique ainsi le directeur Big Data et Innovation chez Bouygues Telecom.



Le Big Data est aussi un vecteur d'amélioration de la qualité des réseaux, pour détecter les zones dans lesquelles surviennent des coupures de communication, en définir les causes et intervenir pour améliorer la couverture ou la capacité de leur réseau.

La maintenance préventive se développe aussi, visant à anticiper les défaillances ou les pannes. En envoyant des messages à ses clients pour qu'ils éteignent leur Livebox en cas de risque d'orage, Orange a ainsi fait sensiblement baisser le nombre d'incidents : en un an 30.000 Livebox de moins endommagées.



### S'adapter à un marché en croissance

A terme, les opérateurs ajusteront en temps réel les capacités de leurs réseaux aux déplacements des utilisateurs : aux États-Unis, les opérateurs commencent à adapter la capacité de leurs antennes mobile lors de certains événements pour coller au mieux aux besoins de leurs clients.

Ceci peut se faire en commercialisant les données collectées : SFR a ainsi développé un outil, «Géostatistique», destiné notamment aux collectivités locales, aux transporteurs et à la grande distribution pour analyser et comprendre les flux de déplacements des personnes dans une ville ou entre des villes.

Ces informations sont aussi utilisées par des chaînes de magasins franchisés pour expliquer les différences de fréquentation des magasins et mieux cerner le profil des consommateurs.

La prochaine étape pourrait être des propositions de contenus ciblés en fonction des consommateurs : SFR réfléchit déjà

à mettre en avant des offres presse spécifiques sur son kiosque, en attendant des contenus vidéo adaptés à chacun, avec les publicités associées.

### ET AU FINAL ?

Rien n'est donc plus mal nommé que la donnée : elle est non seulement produite, mais aussi récupérée, exploitée. Si le travail involontaire des internautes fait l'objet de lumineuses analyses universitaires, le monde politique ou syndical n'a pas encore intégré cette dimension à son analyse - et encore moins à ses revendications.

Pourtant, les formes matérielles et immatérielles de l'exploitation s'imbriquent étroitement : le travail numérique n'est qu'un maillon d'une chaîne passée aux pieds des mineurs contraints d'extraire le coltan requis pour la fabrication des smartphones, aux poignets des ouvrières qui les assemblent, aux roues des chauffeurs sans statut d'Uber, au cou des manutentionnaires d'Amazon pilotés par des algorithmes.



Au final, la question n'est peut-être plus de savoir si il y a débat autour du contrôle des données numériques mais si des forces progressistes prendront part à cet affrontement. Des revendications comme la réappropriation démocratique des moyens de communication en ligne, l'émancipation du travail numérique, la propriété et l'usage socialisés des données prolongent un combat vieux de deux siècles... Et déjouent le fatalisme qui situe inéluctablement l'avenir au croisement de l'État-surveillant et du marché prédateur !