

Big data :

Entre innovation et

protection

La donnée connaît un tournant majeur avec le World Wide Web, accéléré lui-même par sa seconde génération, le Web 2.0. La majorité du contenu des données est fournie aujourd'hui par les utilisateurs des services plutôt que par leurs fournisseurs. Une grande part du patrimoine immatériel de l'organisation est constituée de données. La culture de la sécurisation des données s'impose également. Si l'Open Data est une sorte d'innovation dans la gouvernance publique, favorisant la responsabilité, l'efficacité, la transparence, l'innovation et la création de valeur au niveau économique et social, les objets connectés ouvrent la voie également à une surveillance personnalisée à grande échelle.

Taoufik Benkaraache

Professeur HDR, Hassan II Mohammedia Casablanca

Chercheur associé à *Economia*-HEM Research Center

Directeur du laboratoire de recherche en intelligence stratégique (LIS), Université Hassan II de Casablanca

Ghizlaine Salam

Professeur d'économie, chercheuse au laboratoire LIS, Université Hassan II de Casablanca

Le XXI^e siècle, appelé aussi siècle de l'information ou siècle du numérique, est en train de vivre plusieurs révolutions qui changent progressivement les modes de vie, les comportements, le management... L'introduction des technologies numériques dans le quotidien des personnes et des organisations a généré une accélération et une diversification des échanges informationnels via des réseaux de plus en plus nombreux. Les entités connectées subissent ou procréent une nouvelle forme d'intelligence collective (*Sagesse des foules*, James Surowiecki, 2004), émanant des agrégats de données et d'informations produits par ces nouvelles communautés connectées. Tous les domaines et secteurs sont concernés : politique, économie, industrie, médecine, management, éducation... Les systèmes décisionnels cherchent à exploiter cet « océan de données » (appelé communément données massives ou encore Big Data) créé par les uns et les autres, involontairement ou délibérément, pour extraire l'information pertinente et anticiper le changement. La donnée se propose alors comme matière première indispensable pour comprendre son environnement, créer de la valeur et innover.

Face à ce constat, innovation et Big Data sont devenus deux concepts à la mode, souvent (parfois même trop) utilisés. S'il est certain que les Big Data sont une conséquence directe de l'innovation et l'avancée technologique, on peut se demander

quels sont leurs apports dans les systèmes d'innovation des organisations. Comment peut-on les intégrer comme facteur essentiel à l'écosystème d'innovation, en plus des facteurs classiques tels que la culture entrepreneuriale, l'esprit de l'initiative ou encore la collaboration et le partage ? Quelle(s) innovation(s) se cache(nt) réellement derrière les Big Data ?

Au début, il y a les données, toutes les données...

Pour interroger le caractère innovant des données massives, il est essentiel de s'arrêter sur la courte histoire du phénomène Big Data. Plusieurs experts de la donnée rappellent que les assises de ce phénomène sont aussi anciennes que l'histoire de la donnée elle-même : captation, collecte et stockage des données chez les civilisations antiques ; avancées durant les XVII^e, XIX^e et XX^e siècles dans la recherche en statistique et analyse des données ; avènement au XX^e siècle des premiers centres de données (Data Center). Le tournant dans l'histoire de la donnée est incontestablement le World Wide Web, réseau de données mondialement interconnecté, accéléré lui-même par sa seconde génération : le Web 2.0, où la majorité du contenu est fourni par les utilisateurs des services plutôt que les fournisseurs de services eux-mêmes. La multiplication des objets connectés a accéléré l'histoire pour passer à l'ère numérique de la donnée, les Big Data. Les innovations technologiques et numériques se sont multipliées dans cette dernière décennie pour offrir des solutions et services de management de données plus volumineuses, plus variées (données structurées ou non structurées), de manière plus rapide, plus fiable, avec plus de valeur et de visibilité (on parle ici des 6V qui caractérisent les Big Data : d'abord les 3V comme volume, vitesse, variété, initialement introduits par Gartner, ensuite augmentés de trois autres « V » comme, véracité, valeur et visualisation). Le dernier épisode du feuilleton est l'émergence de l'intelligence artificielle (IA) dans les processus décisionnels : il ne suffit pas de stocker la data, encore faut-il être capable de l'analyser et extraire de la valeur pour la prise de décision. Les algorithmes d'analyse des données (ou de Datamining) ont accompagné la puissance de traitement des processeurs et permettent actuellement un apprentissage très avancé à partir des données massives. Plus il y a de données, plus la machine apprend ; la donnée devint alors la ressource la plus convoitée. Son utilisation n'est pas limitée et ne diminue pas avec le temps. Son usage n'implique aucune rivalité : le fait qu'elle soit consommée par une personne n'empêche pas sa consommation simultanée par d'autres personnes, et plus elle est utilisée, plus sa valeur augmente ! Nous sommes donc face à une ressource de nature singulière, disponible en grandes quantités et dont l'utilisation pourrait être source de nouvelles formes d'intelligence et d'innovation pour les organisations. Les corrélations entre ces données ouvrent des perspectives pour les entreprises comme pour les États.

L'innovation par les données

Il est donc incontestable que cette révolution des Big Data entraîne une nouvelle culture managériale axée sur la maîtrise de l'information. De nouvelles formes de management et de gouvernance sont alors nées : on peut parler ici d'une innovation organisationnelle par la donnée. La culture de la donnée est devenue centrale dans les approches managériales des organisations apprenantes. Les cellules de veille se multiplient pour systématiser la collecte et la surveillance de l'environnement. Le partage des données est de plus en plus facilité à travers les réseaux internes. Une grande part du patrimoine immatériel de l'organisation est constituée de données. La culture de la sécurisation des données s'impose également. De nouveaux métiers et spécialistes apparaissent : les « Data scientists ». L'avenir appartient aux organisations et pays qui ont anticipé les formations dans les spécialités de la donnée. Les États-Unis et la Chine forment des milliers de Data scientists à l'aide de programmes de formation innovants.

En termes de valeur, il est difficile de mesurer ou quantifier la valeur des données, car elle réside dans la manière dont on les utilise. Cette valeur résulte des idées, représentations, solutions et services générés grâce à ces données, qui sont

sources d'innovations et de nouveaux modèles économiques et nouvelles activités pour nous rendre la vie meilleure.

L'Open Data, source d'innovation utile

L'ouverture des données (ou Open Data) est une autre forme de révolution des Big Data à l'échelle des organisations publiques. Il s'agit là d'une sorte d'innovation dans la gouvernance publique incitée, d'une part, par le besoin de transparence et d'éthique et, d'autre part, par les impératifs de l'intelligence collective et collaborative dans la recherche de la performance dans l'accès aux services publics. La donnée provoque ici une innovation systémique dans la gouvernance de la chose publique. Elle favorise la responsabilité, l'efficacité, la transparence, l'innovation et la création de valeur au niveau économique et social.

L'exploitation des Big Data et Open Data est porteuse de valeurs et de changements dans plusieurs domaines économiques et sociaux : l'utilisation des données ouvertes peut aider, par exemple, dans les solutions innovantes en matière d'optimisation des systèmes de transport et de consommation de l'énergie (voir l'expérience du Royaume-Uni, le Rapport sur la révolution des données).

Dans le même registre, l'exploitation des données massives dans le domaine de la médecine (l'ensemble des données socio-démographiques et de santé, disponibles auprès de différentes sources qui les collectent pour diverses raisons) présente de nombreux intérêts : identification de facteurs de risque de maladie, aide au diagnostic et au suivi de l'efficacité des traitements, etc. Les Big Data se présentent ici encore comme porteuses d'innovations utiles en termes de prévention et d'amélioration des services de santé pour les individus et les populations.

Big Data : l'équilibre entre protection et innovation

L'exploitation des Big Data est un sujet sensible. Elle peut même devenir source de risques majeurs. En effet, les conséquences négatives que peut engendrer le pouvoir des détenteurs de ces gisements de données peuvent être source d'atteinte à la vie privée des personnes et des organisations, voire des pays.

D'abord, la question de la collecte et la captation des données personnelles, au même titre que la réemployabilité et la vente des données personnelles, est au cœur de la controverse. La société de surveillance se met en place. Les objets connectés permettant la collecte des données dans les espaces publics et privés. Les cartes à puces, cartes bancaires, géolocalisation... ouvrent la voie à une surveillance personnalisée à grande échelle. Nous entrons dans une société de la traçabilité. La sphère privée se réduit et devient de plus en plus transparente. Malgré le développement d'outils juridiques pour réglementer l'accès et la collecte des données personnelles (notamment le Règlement général sur la protection des données personnelles (RGPD) en Europe, 2018), le risque d'induction d'innovations organisationnelles et sociales nuisibles reste omniprésent. L'exemple de la Chine qui, grâce aux Big Data et aux systèmes de surveillance à grande échelle, a commencé à noter ses citoyens en bons et mauvais citoyens, pousse à réfléchir sur l'impact de cette innovation sur le comportement des citoyens.

La deuxième controverse qui alimente les débats actuels sur la frontière entre la protection et l'innovation dans l'utilisation des Big Data est celle de l'émergence d'un pouvoir artificiel, causé par l'utilisation et le recours quasi systématique aux algorithmes dans le Big Data mining (l'extraction des connaissances à partir des données massives brutes). L'intelligence artificielle développée par ces méthodes statistiques et mathématiques est en train de prendre le dessus sur l'Homme. Elle anticipe et oriente nos choix en fonction de simplifications de visualisation et modélisations

mathématiques opérées sur les données. Elle influence de plus en plus les comportements et les décisions.