

Introduction

Quel monde voulons-nous ? Jusqu'où pouvons-nous et devons-nous aller ?

Par le Dr Laure TABOUY, PhD

Neuroscientifique et éthicienne, équipe Éthique et épistémologie, CESP-INSERM U1018, Espace éthique APHP, Université de Paris-Saclay

Confiance, fiabilité, responsabilité, éthique, transparence, rigueur, confidentialité, disponibilité, intégrité : voici quelques mots clefs importants pour parler de la protection des données et de ses enjeux.

Concevoir l'avenir de la protection des données dans une économie globalisée et dans un contexte international incertain de guerre et de crise sanitaire nous demande d'être plus responsables, plus vigilants, plus réfléchis au regard de nos données. L'utilisation que nous en faisons et la protection de ces dernières figurent parmi les préoccupations majeures en 2022.

Mais de quoi parlons-nous ? Qu'est-ce qu'une donnée, ou, plus précisément, une donnée sensible ou une donnée chiffrée ? Est-ce qu'une donnée identitaire soulève les mêmes enjeux qu'une donnée de recherche ou de santé, une donnée cérébrale, démographique, d'interaction ou de comportement, ou encore transactionnelle, financière ou industrielle ? De plus, définir les données de la recherche n'est pas un exercice aisé, dans la mesure où elles sont de natures extrêmement diverses.

L'usage et la réutilisation des données en vue d'une confirmation de résultats de la recherche ou de l'apport d'une aide à la décision, ou d'accroître la confiance dans l'innovation, mais aussi en vue de favoriser le partage des données et la définition de modèles économiques de valorisation de celles-ci sont des enjeux cruciaux en 2022 pour tous les acteurs du monde de la recherche et des entreprises. L'attractivité de la France en dépend.

Mais il ne faut pas oublier que le *business* des données va bon train : le *business* de la vente mais aussi du vol des données des utilisateurs d'Internet ou des patients a des incidences sur les actions de la vie quotidienne des citoyens, sans qu'ils ne s'en rendent compte. Il peut ruiner les fondations d'une démocratie en favorisant l'addiction et la manipulation⁽¹⁾. Car les réseaux sociaux sont une drogue, en tant qu'ils déclenchent le relargage de la dopamine et favorise le renforcement du circuit de la récompense. Les outils technologiques, l'IA et les algorithmes sont aujourd'hui au cœur des innovations ; ils ont été étudiés et conçus pour permettre de contrôler l'attention des utilisateurs, tout en collectant leurs données.

Le respect de la vie privée et la protection des données sont des enjeux majeurs de notre société, où, dans la période actuelle, le piratage des données s'accroît, où des milliards de dollars sont investis dans but de monétiser ces données, où les risques d'usurpation d'identité et de manipulation des populations à grande échelle sont à prendre très au sérieux.

Les développements récents issus de la convergence de technologies émergentes se traduisent par une accélération de l'innovation se concrétisant par des applications dans les entreprises et le monde de la recherche que ce soit dans les domaines de la santé, de la finance, de l'aéronautique, de la cybersécurité, de la Défense, ou en matière de recrutement, d'accompagnement des projets, de consommation ou de déplacement, mais aussi dans le domaine de la recherche, toutes disciplines confondues en sciences humaines et sociales et sciences exactes, notamment en neurosciences et en génétique.

Au regard de la place que prend aujourd'hui la problématique de la protection des données dans notre société, le comité éditorial de la série *Réalités industrielles* des *Annales des Mines* a donc souhaité contribuer à la mise en œuvre des recommandations émanant des différentes autorités et des dispositions de la loi Informatique et libertés en dédiant le présent numéro à l'avenir de la protection des données dans les entreprises et dans la recherche dans un contexte international plus responsable.

⁽¹⁾ Voir le documentaire « Derrière nos écrans de fumée », <https://www.thesocialdilemma.com/>

Ce numéro des *Annales des Mines* s'articule autour de trois axes qui le structurent et ont pour intitulés :

1. L'utilisation des données : entreprises privées et recherche académique, ensemble pour un avenir plus responsable de la gestion des données.
2. L'économie des données : quelles utilisations des données au cœur des collaborations, des partenariats ou des plateformes collaboratives. Quelles données partager ? À qui en ouvrir l'accès ?
3. Les sciences humaines et les données : les enjeux de la mutation des sociétés à travers les données. Faire confiance au temps du numérique.

La problématique de la protection des données étant subordonnée à la singularité des spécificités de chaque entreprise et de la recherche, nous avons fait le choix, dans la première partie de ce numéro, de nous focaliser sur l'utilisation des données et sur les enjeux et les questions auxquels nous sommes confrontés en la matière.

L'obligation d'obtenir un consentement libre et éclairé des personnes avant une utilisation de leurs données ; la nécessité que les entreprises, les industriels, mais également les organismes de recherche soient rigoureux, fiables, responsables, transparents, intègres et honnêtes dans l'exercice de leurs activités ; le besoin urgent de réfléchir à la pertinence de la collecte des données et de leurs lieux de stockage, d'ouvrir le débat au sein de la société mais aussi des entreprises et des communautés de chercheurs, de légiférer, de mettre des bornes, des limites et de mettre en balance les bénéfices et risques en matière de partage des données : ce sont là des problématiques qui doivent être considérées de manière très sérieuse aujourd'hui. Ce sont des questions que prennent à bras le corps les délégués à la protection des données (DPO) présents dans les entreprises mais aussi certains cabinets d'avocats, qui sont les interlocuteurs privilégiés des entreprises et des organismes de recherche, et qui sont les garants du respect de la loi.

Que signifie une donnée nous appartenant en droit ? Que recouvrent les termes « souveraineté » et « confidentialité » des données ? Qui a le droit de réutiliser des données, notamment celles de clients, pour une autre utilisation que celle prévue à l'origine ?

Le respect de la confidentialité des données collectées dans le cadre des activités des entreprises ou de recherche, qu'elles soient civiles ou militaires, n'est plus une option. Des protections appropriées des données numériques comme non numériques présentes sur des espaces privés, de santé et relatives à notre identité individuelle, ainsi que des contenus sur Internet doivent être mises en œuvre pour nous aider dans notre compréhension des droits de l'homme qui s'y attachent.

Cette juste protection des données va permettre aux entreprises et aux organismes de recherche de continuer à être, au plan national et international, des acteurs influents majeurs de la valorisation et du transfert vers la société des innovations, dans toutes leurs dimensions. Il en va de la confiance que nos concitoyens accordent et vont continuer d'accorder aux entreprises, à l'État et aux chercheurs qui travaillent au sein des organismes de recherche académiques et privés.

Les enjeux de la protection des données qui sont au cœur de l'activité de nombre de filières industrielles, notamment de la *healthtech* et de la *medtech*, et de celle des laboratoires et autres organismes de recherche sont colossaux.

Quels sont les enjeux à relever par la France sur le plan de la protection des données pour qu'elle reste attractive aux yeux des entreprises et de la recherche ? Qu'est-ce qu'une donnée sensible en recherche, selon les disciplines ? Comment les organismes de recherche s'organisent-ils ? Comment ces différentes disciplines, notamment celles travaillant sur le microbiote et celles des neurosciences, s'organisent-elles à l'échelle européenne ?

Pour apporter des éléments de réponse à ces nombreuses questions, nous inviterons certains acteurs de la science des données à s'exprimer en tout début de la deuxième partie de ce numéro. Souvent associée à la collecte des données massives et à l'analyse de celles-ci, cette discipline interdisciplinaire utilise entre autres « des méthodes et des processus d'apprentissage automatique, des algorithmes pour extraire des connaissances et des idées de nombreuses données structurelles et non structurées. »

L'internationalisation des données est un véritable enjeu. Qu'est-ce que représentent ces données ? Qu'est-ce que l'anonymisation et le chiffrement des données ? Comment sont-elles classifiées et catégorisées ? Comment le *machine learning*, les algorithmes et la *blockchain* interviennent et permettent de représenter ces données ? Comment sont organisées ces données au sein d'une bibliothèque de données de recherche ?

Puis nous nous intéresserons plus précisément à l'utilisation de l'anonymisation et du chiffrement qui permettent d'apporter une réelle protection aux données stockées dans des *cloud*, sur des serveurs ou dans des entrepôts et dont le partage s'opère au sein de consortiums et de collaborations nationales et internationales, comme le Health Data Hub⁽²⁾ ou le Paris-Santé Campus⁽³⁾.

⁽²⁾ <https://www.health-data-hub.fr/>

⁽³⁾ <https://parisantecampus.fr/>

Assurer la protection des données issues de la recherche à l'ère du numérique est l'engagement pris par l'État pour permettre l'ouverture de la science⁽⁴⁾ : « La France s'engage pour que les résultats de la recherche scientifique soient ouverts à tous, aux chercheurs, aux entreprises et aux citoyens, sans entrave, sans délai, sans paiement ». L'enjeu pour le monde de la recherche est inédit et gigantesque ; il demande d'instaurer un véritable dialogue entre tous les acteurs concernés : « La science ouverte est la diffusion sans entrave des résultats, des méthodes et des produits de la recherche scientifique. »

Les différents projets de recherche internationaux, comme le Human Brain Project⁽⁵⁾ et l'International Brain Initiative⁽⁶⁾, mettent l'accent sur le développement des neurotechnologies qui visent à mieux comprendre les fonctions cérébrales et à intervenir sur elles. Le Health Data Cloud⁽⁷⁾ vise à fournir des services en lien avec les données cérébrales sensibles. Ce consortium comprend un ensemble d'infrastructures existantes partenaires du Human Brain Project/EBRAINS⁽⁸⁾ et de fournisseurs de services de données de santé qui ont récemment rejoint ce projet.

Propriété intellectuelle, domaine et données publics, protection des données personnelles... Les données sont à la croisée d'une multitude de régimes juridiques, parmi lesquels il peut être difficile de se repérer.

Enfin, parler de protection des données nécessite de s'attarder sur les enjeux de celle-ci à l'heure des réseaux sociaux et d'Internet.

Quels sont les enjeux et les limites de cette protection ? Quelles sont les questions et les craintes qu'elle soulève pour les futures générations ? Qu'en est-il du rôle et des responsabilités sociales que doivent assumer les citoyens, les industriels, les entreprises, les politiques, les chercheurs et les organismes de recherche publics et privés, dans la génération de ces données et le devenir de celles-ci ?

Dans la dernière partie de ce numéro, le lecteur pourra naviguer au cœur de la problématique des données et de l'innovation responsable, au travers des contributions de multiples acteurs, dont votre serviteur, spécialistes du droit français et européen, comme l'OCDE et la CNIL, de la cybersécurité sous la plume de représentants de la gendarmerie nationale et de l'armée de terre et, enfin, de l'éthique et de l'épistémologie à travers la position exprimée par la HTF (Human Technology Foundation).

Les questions soulevées sont nombreuses et demandent de faire preuve d'une exigence, d'une pertinence et d'une rigueur sans commune mesure pour traiter ces sujets si brûlants et si politiques. Les enjeux liés aux données, à leur protection, à leur collecte, à leur réutilisation peuvent être de différents ordres.

À la seule évocation des mésusages potentiels des données, apporter des limites éthiques, juridiques, économiques et sociétales à leur utilisation s'avère nécessaire pour que s'instaure une confiance informée des patients, des consommateurs et, plus largement, des citoyens, pour que les investissements et les innovations s'orientent vers des biens et services à forte utilité économique, sociale et environnementale. L'évocation de tels risques permet aussi d'éveiller les consciences à la nécessité de protéger ce que les hommes sont en eux-mêmes, car derrière les données, ne l'oublions pas, il y a des êtres humains, avec leurs histoires singulières.

Les rôles joués par la CNIL et l'AMF, les recommandations de l'OCDE, les dispositions du RGPD ou du Data Act sont centraux à l'échelle française et européenne, en tant qu'ils régissent les aspects juridiques de la protection des données.

L'accélération de l'essor des technologies et des innovations rend indispensable l'engagement de réflexions sur les enjeux qu'elles soulèvent au regard des données, et de travailler à la conception interdisciplinaire de certifications, de normes, de dispositifs d'autorisation, de systèmes d'évaluation et de surveillance et de cadres de gouvernance adaptés aux valeurs sociologiques, économiques, éthiques et juridiques de la France et de l'Europe. Compte tenu du large spectre que recouvrent les questions soulevées par l'usage des données, la Commission européenne, au travers du Règlement général de protection des données (RGPD)⁽⁹⁾, entré en application le 25 mai 2018, mais aussi l'OCDE^{(10) et (11)}, au travers de ses recommandations n°463, adoptée le 6 octobre 2021, et n°433, adoptée le 13 décembre 2016, et la France, par l'intermédiaire de la CNIL⁽¹²⁾, et l'Autorité des marchés financiers⁽¹³⁾ ont pris leurs responsabilités en adoptant les lois et recommandations qu'exige la protection des données.

⁽⁴⁾ <https://www.ouvrirlascience.fr/>

⁽⁵⁾ <https://www.humanbrainproject.eu/en/>

⁽⁶⁾ <https://www.internationalbraininitiative.org/>

⁽⁷⁾ <https://www.healthdatacloud.eu/>

⁽⁸⁾ <https://ebrains.eu/>

⁽⁹⁾ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

⁽¹⁰⁾ OECD, Recommendation of the Council on Health Data Governance, OECD/LEGAL/0433.

⁽¹¹⁾ OECD, Recommendation of the Council on Enhancing Access to and Sharing of Data, OECD/LEGAL/0463.

⁽¹²⁾ <https://www.cnil.fr/professionnel>

⁽¹³⁾ <https://www.amf-france.org/>

Comme le souligne Marie-Laure Denis dans l'éditorial qui introduit ce numéro : « Ce cadre, qu'est le RGPD, a toutefois d'abord été pensé en référence à des systèmes informatiques bien identifiés, situés dans des centres de données fixes et créés spécifiquement pour la mise en place du traitement de données projeté. »

Les réflexions éthiques et épistémologiques liées aux données, au *Big Data*, à leur protection et à leurs usages s'expriment en général par le biais de questionnements et d'intuitions. Le rapport « *DATA altruisme, une initiative européenne* »⁽¹⁴⁾ aborde les enjeux des données sous l'angle de leur mise au service du bien commun.

Réfléchir aux enjeux éthiques nécessite de s'interroger, de s'ouvrir à la démarche de questionnement sur les valeurs et les finalités, les principes et les normes, les contextes et les conséquences de nos actions à des moments où ces dernières sont ambiguës, en tension, voire en conflit avec notre environnement. Une telle réflexion, essentielle à l'innovation responsable, permet alors de poser factuellement le débat.

C'est sur ces trois aspects que ce numéro se clôturera, en nous questionnant sur ce que ces données disent de l'homme, et sur ce que cela change dans notre compréhension de l'humain.

De même, la question de la protection des données sera abordée à la lumière de l'éthique de la recherche, de l'intégrité scientifique, de l'épistémologie et de la neuroéthique, pour en venir à nous questionner sur les enjeux que la science ouverte sous-tend.

Car l'idée que les sciences « s'ouvrent », renvoie à des pratiques de mise en culture des dites sciences, d'une présence de celles-ci dans un grand nombre d'espaces sociaux et sociétaux. Le monde de la recherche aussi bien privée qu'académique et celui des entreprises sont-ils capables de rester autonomes et indépendants, tout en travaillant conjointement pour le bien commun et en interdépendance avec la société et sont-ils prêts à ouvrir largement leurs données ?

⁽¹⁴⁾ <https://www.human-technology-foundation.org/fr-news/rapport-data-altruisme>