LA VIE PRIVÉE À L'ÈRE DES BIG DATA

DANGERS & OPPORTUNITÉS DE LA RÉVOLUTION NUMÉRIQUE

Les études du Centre Jean Gol *www.cjg.be*

Cette étude a été portée par **Corentin de Salle**, Directeur du Centre Jean Gol.

En complément de ce dernier, elle a bénéficié de la rédaction d'un grand nombre d'auteurs, à savoir **Stéphane Tellier**, juriste et collaborateur au Centre Jean Gol, **Jérôme De Cooman**, Assistant à la Faculté de Droit de l'Université de Liège, **Nicolas Petit**, Professeur ordinaire à l'Université de Liège, Research Professor, University of South Australia (UniSA) et Visting Fellow, Stanford University Hoover Institution. Ainsi qu'**Emmanuelle Duquenne**, juriste, consultante dans le secteur des nouvelles technologies, **Arnaud Lombardo**, ingénieur commercial et conseiller Numérique du ministre Pierre-Yves Jeholet, **Laurent Hublet**, ingénieur commercial, philosophe et Managing Director de BeCentral et **Pascal Leduc**, avocat bruxellois spécialisé en droit intellectuel.

Je les en remercie ainsi que le professeur **Yves-Alexandre de Montjoye** (Imperial College, London), le professeur **Antoinette Rouvroy** (Facultés Universitaires Notre Dame de la Paix à Namur), le professeur **Alain Strowel** (UCLouvain, Université de Saint-Louis, KULeuven et Munich IP Law Center), **Charlotte Dereppe**, conseillère Vie Privée au cabinet du ministre Philippe De Backer et **Nathalie Ragheno**, premier conseiller à la FEB qui ont participé à cette étude..

Je vous souhaite une excellente lecture de ce numéro des Études du Centre Jean Gol.

RICHARD MILLER
Administrateur délégué

Les Études du Centre Jean Gol sont le fruit de réflexions entre collaborateurs du CJG, des membres de son comité scientifique, des spécialistes, des mandataires et des représentants de la société civile.
Accessibles à tous, elles sont publiées sous version électronique et sous version papier.

RESPONSABLES SCIENTIFIQUES

Olivier Chastel, Président du CJG
Richard Miller, Administrateur délégué du CJG
Corentin de Salle, Directeur du CJG

résumé

La révolution des big data constitue à la fois une menace pour la protection de la vie privée et un potentiel économique exceptionnel. Dans cette étude, nous entendons traiter ces deux questions de concert.

Nos données sont convoitées par des Etats trop zélés dans la lutte antiterroriste mais aussi par des acteurs commerciaux observant nos comportements pour dresser un profil très personnalisé qui, comparé à des profils similaires, leur permet de prédire nos achats futurs. Précurseur, le Règlement Général de Protection des Données (RGPD) apporte une réponse probablement satisfaisante à ce problème en Europe mais les moyens de notre nouvelle Autorité de Protection des Données devraient être renforcés. Quoi qu'il en soit, le danger réside moins dans la violation de notre vie privée que dans celui de voir les algorithmes prendre, avec notre consentement, progressivement le contrôle de nos vies. Une plus grande transparence des algorithmes devrait être exigée et une charte des algorithmes adoptée.

Les big data sont un Eldorado mais l'Europe et la Belgique sont encore à la traîne. Nous devons former massivement les travailleurs de ces filières du futur, investir dans les infrastructures permettant d'extraire et de traiter ces données. Afin de favoriser la création de valeur par les PME, les pouvoirs publics doivent rendre publiques leurs données encore trop peu exploitées et favoriser les partenariats public-privés équilibrés.

Une étude réalisée par

CORENTIN DE SALLE, STÉPHANE TELLIER, JÉRÔME DE COOMAN, NICOLAS PETIT, EMMANUELLE DUQUENNE, ARNAUD LOMBARDO, LAURENT HUBLET & PASCAL LEDUC

INTRODUCTION

Corentin de Salle

Début 2014, Yannick Boloré, patron de Havas, était invité à San Francisco par les dirigeants de Google. A l'atterrissage, un SMS l'avertit qu'il bénéficie d'une réduction de 15% sur les sushis au saumon à un restaurant situé à proximité de son hôtel. Il se fait que c'est son plat favori. Etonné, il s'enquiert le lendemain auprès des responsables de Google pour savoir s'il s'agit d'une coïncidence. Non, répondent-ils : l'algorithme savait que c'était votre plat favori (car vous en consommez souvent), il vous a géolocalisé, il a compris, par l'examen de vos mails, que vous descendiez à tel hôtel et a généré une publicité ciblée. Boloré rétorque :

> *« - Mais que faites-vous de la vie privée ? - Ah oui, la privacy, c'est vrai qu'en Europe, vous en parlez beaucoup ».* [1]

Cette réflexion, prononcée dans le pays qui a pourtant inventé la notion juridique de « privacy » à la fin du XIX^{ème} siècle, est interpellante. A l'heure des big data, force est de constater que la notion de vie privée, qui fait pourtant partie des libertés fondamentales, est aujourd'hui menacée. Pire : on relativise son importance et on questionne même sa pertinence. C'est ce qui fit Eric Schmidt, le CEO de Google en 2009 dans une interview de la chaîne CBS :

> *« Si vous avez fait une chose que vous ne voulez pas que tout le monde sache, il serait peut-être bon de commencer par ne pas faire cette chose ».*[2]

Le corolaire de cette assertion est l'idée suivante : « si vous n'avez rien à vous reprochez, que vous importe qu'on sache ce que vous faites ». Cette idée est une idée fausse et dangereuse car on confond ici « rien à cacher » et « rien à se reprocher ». Nous avons tous besoin de lieux pour nous ressourcer et partager avec des proches des choses qui n'ont pas à être divulguées. Par ailleurs, il est tout à fait possible de lutter efficacement contre le terrorisme sans empiéter sur la vie privée des citoyens. Nous devons préserver avec vigilance cet acquis que nous devons aux révolutions libérales du XIX^{ème} siècle.

Nous entendons montrer que **la préservation de la vie privée et le développement d'une économie de la donnée sont néanmoins tout à fait conciliables**.

Grâce à la grande diversité de profils des auteurs de la présente étude, nous avons pu mener cette recherche collective avec un quadruple éclairage : juridique, technique, économique et philosophique.

Dans un **premier chapitre**, j'ai tenté de retracer, de l'Antiquité à nos jours, l'histoire de la notion de « vie privée » et j'ai expliqué en quoi elle était une valeur aussi fragile que précieuse.

Dans un **second chapitre**, j'ai montré en quoi elle était menacée à l'ère numérique dans nos sociétés occidentales : par l'Etat qui, dans sa lutte antiterroriste, a pu adopter des programmes de surveillance de masse. Mais aussi par les géants du web à l'affut d'informations permettant d'anticiper nos comportements d'achat. Il existe heureusement des manières de se défendre.

Dans un **troisième chapitre**, **Stéphane Tellier**, juriste et collaborateur au Centre Jean Gol, dresse un panorama éclairant des législations protectrices de la vie privée en Europe et en Belgique. Il explicite aussi le rôle des autorités de surveillance. Il expose en détail le RGPD entré en œuvre il y a 6 mois. Avec **Pascal Leduc**, avocat spécialisé en droit intellectuel, il propose une analyse critique de ce règlement. A travers une présentation du régime du « Safe Harbour » et puis du « Privacy Shield », il examine aussi la difficile conciliation des approches américaines et européenne de la protection des données à caractère personnel.

Dans le **quatrième chapitre**, je définis les notions de big data et d'algorithmes. Contrairement à ce qu'on croit parfois, big data et algorithmes ne sont pas neutres. Un data n'est pas une donnée brute. Elle est « produite ». Les algorithmes sont conçus par des humains qui ont des finalités. Ces algorithmes sont construits par nous mais nous construisent en retour. Une idéologie technicienne veut nous faire croire qu'ils nous permettraient d'accéder directement au réel, que la réalité parlerait à travers eux. Les algorithmes ne visent pas à comprendre le réel mais à établir des corrélations qui invitent à l'action.

Dans un **cinquième chapitre** consacré à la « gouvernementalité algorithmique », je m'interroge sur le danger que les algorithmes puissent nous « gouverner » un jour, prendre contrôle de nos vies. Je le fais à travers l'analyse de la pensée très stimulante du professeur Antoinette Rouvroy. Je conclus le chapitre par une analyse critique de ses idées principales.

---

1 Anecdote rapportée dans M. Dugain & Ch. Labbé, ***L'homme nu. La dictature invisible du numérique****, Plon, 2018 (2016), pp.21-22*

2 Gl. Greenwald, ***Nulle part où se cacher, L'affaire Snowden par celui qui l'a dévoilé au monde****, JC Lattès, 2014, p.239*

Dans un **sixième chapitre** fort original, **Laurent Hublet**, managing director de BeCentral, ingénieur commercial et philosophe, s'interroge sur la donnée et entend dissiper un certain nombre de malentendus**.** En résumé, la donnée n'est pas donnée – elle est co-produite. Ce n'est pas donné de co-produire des données**.** Une même donnée peut avoir de multiples donnants**.** Différents mesurants peuvent co-produire une même donnée**.** La donnée fait généralement sens quand elle est multiple**.** La valeur de la donnée est extrinsèque. Ces points de départs ouvrent la porte de nombreuses questions nécessitant un vrai débat public. Quelle sont les données à caractère « non équivoquement » personnel ? Quelles doivent être les forteresses imprenables de l'intimité numérique ? Quel participant à la production de données peut décider de quoi dans l'utilisation faite de celles ? Qui détermine la valeur de la donnée ? Selon quel mécanisme ? Comment peut-elle être équitablement répartie entre ses co-producteurs ?

Avec le même esprit iconoclaste, les juristes **Nicolas Petit** et **Jérôme de Cooman**, respectivement professeur ordinaire et assistant en droit à l'Université de Liège, démontent trois idées fausses dans le **septième chapitre**: les données sont le nouvel or noir (1) ; les données personnelles ayant une valeur marchande, les individus devraient pouvoir les vendre (2) ; les réglementations spécifiques sur la protection des données (comme le « RGPD ») favorisent les monopoles au détriment des petites et moyennes entreprises (3).

La juriste **Emmanuelle Duquenne**, ancienne collaboratrice de cabinets ministériels et travaillant actuellement dans le secteur des nouvelles technologies défend l'idée, dans un **huitième chapitre**, que la révolution des big data est un nouvel Eldorado. Elle explique en quoi un grand nombre de secteurs sont profondément transformés par cette révolution et tous les bienfaits économiques que nous pouvons en attendre.

Dans un **neuvième chapitre**, **Arnaud Lombardo**, ingénieur commercial et Conseiller au Numérique auprès du Ministre wallon de l'Économie, de l'Industrie, de l'Innovation, de l'Emploi et du Numérique, traite de la thématique de l'Open Data. Il constate que les données publiques sont encore beaucoup trop faiblement partagées alors même qu'il s'agit d'une ressource essentielle pour le développement économique. Il encourage à la création de partenariats public-privé et estime que ces échanges devraient se faire en veillant à restaurer une symétrie entre les partenaires.

Dans un **dernier chapitre**, nous formulons un certain nombre de recommandations afin de sensibiliser davantage les citoyens, les former et les mettre en garde à propos de certains dangers inhérents à cette révolution. Nous formulons des pistes pour améliorer la protection de la vie privée. **Emmanuelle Duquenne** dresse une liste de propositions pour accompagner et stimuler l'économie de la donnée. Arnaud Lombardo préconise diverses mesures concrètes pour « ouvrir » les données.

I. HISTORIQUE ET SIGNIFICATION DE LA VIE PRIVÉE

Corentin de Salle

La vie privée sous la forme où nous la connaissons est un produit de la tradition libérale. C'est une conséquence logique de la consécration de la liberté individuelle durant la période moderne. Mais, dès l'Antiquité, cette notion existait déjà sous une autre forme.

Les Grecs distinguaient en effet entre :

- ***l'oikos*** *(littéralement la « maison ») désignant les activités domestiques, lesquelles pouvaient se déployer sur un domaine gigantesque dont étaient partie prenante les membres de la famille et les esclaves, ce qui explique d'ailleurs que ce terme a donné naissance au mot « économie ». C'est la sphère privée ;*

- *la* ***polis*** *(littéralement la « cité ») désignant l'activité qui se déploie dans la sphère publique mais à laquelle seuls les « hommes libres » avaient accès. C'est la sphère publique.*

Que signifie « *être libre* » à cette époque ? La philosophe Hannah Arendt[3] a étudié de près cette question. Originellement, la liberté se confond avec la liberté de mouvement : le pouvoir d'aller et venir où bon nous semble. Peu de gens bénéficiaient de ce luxe authentique. Se mouvoir dans l'espace public était limité à des privilégiés. Cela implique donc de n'être soumis à la contrainte d'aucun autre homme mais également de pouvoir s'éloigner de la sphère de contrainte toute entière du foyer et de la « famille ».

Cette faculté de pouvoir « se détacher » du foyer permet de comprendre cette idée un peu bizarre que pour un Grec, seul peut être libre celui qui est prêt à risquer sa vie. L'homme libre peut se détacher de son foyer, de la famille dont il est responsable. Il est prêt à assumer ce risque.

Au contraire, la *philopsychia* ou « *amour de la vie* » (notion péjorative) était communément attribué aux serviteurs et aux esclaves, comme un trait qui les séparait des hommes libres.

Telle est la distinction qu'opéraient les Grecs entre domaine public et domaine privé.

- Le public signifie tout ce qui paraît en public, tout ce qui peut être vu et entendu de tous, tout ce qui jouit de la plus grande publicité possible.

- Par contre, le domaine privé, c'est le domaine où l'on se cache. C'est l'endroit des fonctions biologiques. C'est là où l'on subvient à ses besoins naturels. L'endroit où l'on fait les choses les plus triviales.

En gros, la vie privée est l'endroit où l'on ne fait rien qui soit digne de paraître en public. Il faut, nous dit Hannah Arendt, comprendre le mot « privé » au sens originel de « privatif ». Vivre une vie essentiellement privée, c'est vivre une vie privée de toutes les choses essentielles à une vie humaine. C'est le cas des femmes et des esclaves qui ne pouvaient pas quitter le foyer. C'est seulement la présence des égaux qui peut garantir la publicité. Ceci permet de comprendre l'origine de l'espace public.

Pour le dire autrement, la vie privée était autrefois la règle et la vie publique l'exception. Les Grecs ne ressentaient pas le besoin de « protéger » cette sphère car elle n'était pas menacée. Le pouvoir politique ne songeait même pas à s'immiscer dans cette sphère. Quelques siècles plus tard, le droit romain consacrait d'ailleurs cet état des choses par un principe : le droit s'arrête aux frontières de la « domus ». Dans la domus, le « pater familias » est seul maître. Il a droit de vie ou de mort sur tout ce qui vit. Par ailleurs les Grecs n'avaient pas la même notion que nous de l'intimité parce que quantité de personnes y prenaient part dans la sphère domestique (famille, proches, clients, esclaves, etc.).

Selon l'historien Philippe Ariès, espace privé et espace public se confondent au Moyen Age. L'individu est pris dans un système de solidarités collectives, féodales et communautaires. Il est prisonnier, étouffé par la « grégarité domestique ».[4] Les récits des chevaliers errants et solitaires mus par leurs seuls désirs sont une échappatoire à la société féodale dont les exigences de survie comprimaient les aspirations à la liberté des personnes. En un sens, la littérature courtoise est annonciatrice de l'individualisme.

3 H. Arendt, ***Qu'est-ce que la politique ?****, Seuil, 1958, p.83 et s.*

4 Ph. Ariès & G. Duby, ***Histoire de la vie privée, T.2 De l'Europe féodale à la Renaissance****, Seuil, 1985, p.514*

C'est entre le XVI et le XVIII[ème] siècle que s'opère un processus de privatisation.[5] Il y a plusieurs raisons à cela nous dit Philippe Ariès. D'abord, parce que l'Etat assure progressivement, avec le développement de l'appareil judiciaire, le « contrôle du paraître »[6] et permet à l'homme de se consacrer à lui, à rester chez lui et y entretenir un commerce agréable avec une petite « société » d'amis bien choisis. Jusqu'alors, les individus doivent constamment défendre leur « honneur » par de fréquents duels et faire étalage de richesses (étalage qui peuvent les ruiner) car l'individu n'était pas comme il était réellement mais comme il paraissait ou plutôt comme il parvenait à paraître. Il devait constamment « garder la face ». Richelieu interdisit les duels sous peine de mort. Il adopta des lois somptuaires pour interdire le luxe de l'habit et les dépenses excessives. A partir de ce moment, l'individu se replie chez lui et organise des petits salons dans lesquels vont se déployer des conversations privées. On voit d'ailleurs se développer toute une série de lieux où l'individu peut s'isoler : par exemple, la « ruelle » ou « l'alcôve »,[7] c'est-à-dire l'espace entre le lit et le mur de la chambre. Les chambres à coucher au moyen âge sont encore pleines de gens mais cet espace, d'abord réservé aux aventures galantes, devient un lieu privé. Il y aura aussi, progressivement, les jardins clos et puis les « études » et les « bibliothèques » où il peut « se retirer ». Montaigne en parle plusieurs fois dans ses Essais.

Ensuite, parce que l'alphabétisation fait des progrès et, avec elle, la pratique de la lecture silencieuse et solitaire (par rapport à la lecture à voix haute et collective qui a lieu durant la messe). Au Moyen Age, hormis quelques moines, la lecture à voix haute était la seule manière de lire (dans les châteaux, les veillées de campagne, etc.). Dès lors, l'individu est soustrait au contrôle de la communauté et peut intérioriser les réflexions que lui inspirent ces lectures. Il peut progressivement se forger un moi intime.

Montaigne se retire dans sa bibliothèque

« Chez moy, je me destourne un peu plus souvent à ma librairie, d'où tout d'une main je commande à mon mesnage. Je suis sur l'entrée et vois soubs moy mon jardin, ma basse court, ma court, et dans la pluspart des membres de ma maison. Là, je feuillette à cette heure un livre, à cette heure un autre, sans ordre et sans dessein, à pieces descousues ; tantost je resve, tantost j'enregistre et dicte, en me promenant, mes songes que voicy. Elle est au troisiesme estage d'une tour. Le premier, c'est ma chapelle, le second une chambre et sa suite, où je me couche souvent, pour estre seul. Au dessus, elle a une grande garderobe. C'estoit au temps passé le lieu plus inutile de ma maison. Je passe là et la plus part des jours de ma vie, et la plus part des heures du jour (...) qui me plaist d'estre un peu penible et à l'esquart, tant pour le fruit de l'exercice que pour reculer de moy la presse. » (Les Essais, Livre III, Chapitre III, Des trois commerces, 1595).

Enfin, c'est à cette période que des formes nouvelles de religion se mettent en place au XVI et XVII[ème] siècle. Elles permettent de développer la piété intérieure, de procéder à des « examens de conscience ». La prière se transforme en une méditation solitaire dans un oratoire privé ou sur un « prie-Dieu », un meuble adapté pour prier à genoux dans un coin de chambre.

C'est à partir des diverses révolutions libérales du XVIII[ème] siècle que se forge progressivement cet espace protégé qui arbitrera la vie privée. Et cela dans le sillage de toutes les libertés individuelles consacrées par les révolutions anglaise, américaine et française, notamment l'habeas corpus et la règle de l'inviolabilité du domicile garantie en France dès 1791 et dont la violation est sanctionnée sévèrement part l'article 184 du Code pénal. Par ailleurs, l'évolution économique sépare et éloigne progressivement le lieu de travail et le domicile. C'est aussi à partir de cette période que commence à se développer et à prospérer la classe moyenne.

Mais c'est en Amérique que sera proclamé pour la première fois « un droit à la vie privée ». Dans un article de référence publié en 1890 dans la prestigieuse Harvard Law Review et intitulé « **The right to privacy »**[8] (Le droit à la vie privée), **Samuel Warren et Louis Brandeis** ont consacré le « **right to be let alone** » (le droit d'être laissé tranquille). Cet article de 27 pages est l'un des textes les plus influents de l'histoire du droit américain. Ce n'est rien moins que l'acte de naissance du droit à la vie privé dans le monde anglo-saxon. Pour la petite histoire, cet article a pour origine la fureur de Samuel Warren consécutive au fait que la presse avait fait intrusion dans la cérémonie de mariage de sa fille. La presse américaine dans son ensemble a, par la suite, chèrement payé cette erreur.

Les deux auteurs partent du principe fondamental selon lequel « les individus doivent bénéficier d'une pleine protection de leur personne et de leur propriété » et estiment que ce droit a eté reconfiguré avec succès en raison des progrès politiques, sociaux et économique. Ainsi, le « droit à la vie » permet de protéger le citoyen contre toute agression mais il a été étendu pour protéger le citoyen contre la crainte d'une agression. Le « droit de propriété » protégeait initialement les biens tangibles mais il a été étendu aux biens intangibles tels que les produits de l'esprit protégé désormais par les droits d'auteur.

Or, le problème disent-ils, c'est que la presse est devenue de plus en plus invasive. Les commérages sont devenus un commerce. Sans vergogne, la presse amène tout sur la place publique, en ce compris le détail des relations sexuelles.

---

5 Ph. Ariès & G. Duby, ***Histoire de la vie privée, T.3 De la Renaissance aux Lumières****, Seuil, 1985, p.161*

6 Ibidem, pp.10-11

7 Ibidem, pp. 221 et s.

8 S.D.Warren & L. D. Brandeis, The right to Privacy, ***Harvard Law Review****, Vol. 4, N°5, Dec.15, 1890*

Dès lors, il importe, disent-ils, de regarder s'il n'existe pas un principe qui puisse être invoquer pour protéger le « droit à la vie privée ». On pense tout de suite, disent-ils, aux lois qui répriment la calomnie et la diffamation. Mais, selon eux, elles ne sont pas suffisantes car elles condamnent seulement les comportements qui portent atteinte à la réputation de quelqu'un. Elles nécessitent que les personnes dont on a violé la vie privée souffrent directement dans leur rapport avec une ou plusieurs autres personnes. Or, notre réputation peut très bien demeure intacte alors même que nous souffrons du fait que des éléments privés de notre vie aient été dévoilés publiquement.

Est-ce que le droit de propriété intellectuelle ne pourrait pas fournir une telle base se demandent-ils ? **Ils pensent que ce droit qui protège les pensées, les sentiments et les émotions** pour peu qu'ils aient été exprimés dans des écrits ou par le medium artistique **est simplement une application particulière « du droit », plus général, « de l'individu d'être laissé tranquille »**.

Le droit de propriété ne fournit pas une base suffisante lui non plus. Certes, ce droit de propriété permet d'empêcher une publication mais c'est uniquement pour protéger le droit du créateur au profit dérivé d'une publication. C'est cela la source de ce droit. En soi, le droit de propriété seul ne permet pas d'interdire une publication.

Dès lors, la seule base légitime pour enraciner le droit à la vie privée réside bien, selon ces auteurs, dans le droit de l'individu à être laissé tranquille.

Ils énumèrent un certain nombre de limitations. Ainsi, ce droit à la vie privée ne s'applique pas aux publications qui présentent un intérêt public. Si quelqu'un se présente à un mandat public, on réprimera bien toute atteinte à la vie privée, au dévoilement d'habitudes, d'actes et de relations qui n'ont aucune connexion légitime avec son aptitude à exercer un mandat. On ne réprime pas la publication de faits qui ont le consentement de la personne. Etc.

Ces principes inspirent toujours le droit à la vie privée aujourd'hui. Aux Etats-Unis et, comme on le verra, dans notre pays également.

POURQUOI LA PROTECTION DE LA VIE PRIVÉE EST-ELLE SI IMPORTANTE ? POUR DIVERSES RAISONS.

D'abord, pour préserver la **dignité** des gens. Par pudeur, pourrait-on dire.

Ensuite, parce que dévoiler ces choses qui doivent rester secrètes, c'est **rendre les gens vulnérables**. Cela peut conduire à miner leur autorité s'ils exercent des responsabilités. Cela revient à rendre plus difficile à endosser le rôle social qu'ils doivent endosser dans leur vie professionnelle. Cela peut aussi conduire à révéler leurs faiblesses et à permettre à des personnes peu scrupuleuses d'exploiter ces dernières pour les manipuler, les escroquer, voler leur identité ou leur faire du tort.

Enfin, protéger la vie privée est important parce que tout le monde a besoin d'un **refuge**, d'un endroit où il peut se ressourcer sans se soucier de ce qu'il dit, de ce qu'il fait et de ce qu'il pense. Selon Jérémie Zimmerman, hacker et cofondateur de « La Quadrature du Net », l'intimité :

> *« Ce sont ces endroits où vous êtes vraiment vous-même. Soit seul, soit avec ceux que vous choisissez pour partager cette intimité. L'intimité, c'est être jugé au sens figuré ou au sens propre. Où vous pouvez expérimenter des nouvelles idées, des nouvelles pratiques sans être jugé par vos pairs (...). C'est parce que vous avez l'opportunité, grâce à l'intimité, d'expérimenter des choses nouvelles sans être jugé que vous pouvez faire des erreurs sans conséquences. C'est ainsi que vous pouvez développer votre personnalité. C'est cela qu'on se fait voler avec notre intimité ».*[9]

Jean-Claude Ameisen, président du Comité consultatif national d'éthique en France, donne une définition qui va dans le même sens :

> *« La vie privée, ce n'est pas ce que l'on dissimule, c'est l'espace non public, quelque chose dont nous avons besoin pour ensuite jouer notre rôle sur l'agora. Elle est aussi vitale, socialement, que le sommeil l'est biologiquement ».*

Selon ces deux auteurs, la vie privée aurait donc à la fois un rôle de ressourcement et un rôle propédeutique. Elle permettrait de nous « entraîner » à développer telle ou telle idée ou aptitude dans un cadre protecteur.

Prolongeons la réflexion. La vie privée ne serait-t-elle pas dans un rapport spécifique vis-à-vis des autres libertés ? C'est ce que pense le professeur et juriste Yves Poullet, spécialisé dans le droit informatique. Selon lui, la vie privée est non pas une liberté fondamentale à côté des autres libertés mais une condition des autres libertés.[10] Notamment la liberté d'expression et la liberté de déplacement. Si je sais, dit-il, que je suis constamment espionné, je n'oserai plus m'exprimer comme je le souhaite, même dans des cadres plus intimes et plus privés. Si je me sens contrôlé à tout moment, comment, dès lors, me déplacer comme je le désire ?

---

*9 Documentaire « **Nothing to hide** », 2017, 16'15", www.youtube.com/watch?v=djbwzEIv7gE&t=2911s*

*10 Y. Poullet, Le point de vue de la société sur le sentiment de tracabilité in **La vie privée des citoyens et la protection des données face aux nouvelles technologies : les enjeux**, Sénat de Belgique, 17 octobre 2016, p.65*

II. LA VIE PRIVÉE À L'ÈRE NUMÉRIQUE

Corentin de Salle

Il a fallu attendre les révolutions libérales du XVIII[ème] pour assister à la naissance de la notion de vie privée et le XIX[ème] pour qu'elle soit consacrée juridiquement. **Depuis lors, la vie privée a toujours été menacée.** Elle a été foulée du pied dans les régimes totalitaires. Les démocraties occidentales ont connu plusieurs scandales d'écoutes téléphoniques : de l'affaire Watergate aux écoutes secrètes de l'Elysée en passant par les écoutes durant la guerre d'Algérie, ces atteintes ont toujours existé.

Depuis les attentats des tours jumelles au début de notre siècle, l'administration Bush a autorisé la NSA à écouter les communications électroniques des Américains sans obtenir les commissions rogatoires exigées par la législation. Le Patriot Act, signé le 21 octobre 2001, a autorisé la NSA et la CIA à solliciter, au nom de la sécurité nationale, les grands géants du Net pour obtenir d'eux des informations sur les internautes. En juin 2013, le célèbre lanceur d'alerte Edward Snowden, ancien employé de la CIA et de la NSA, dévoila que la NSA collectait les relevés téléphoniques de millions de clients américains de Verizon, l'un des principaux opérateurs téléphoniques d'Amérique, en application d'un arrêt ultraconfidentiel rendu par une cour de justice. Des millions de personnes étaient ainsi espionnées en masse alors même qu'elles n'étaient suspectées d'aucun méfait. Le même mois, Edward Snowden dévoilait le programme PRISM : ce dernier permettait à la NSA de puiser dans les données utilisateurs d'Apple, Google, Gmail, Skype, Facebook, Yahoo et d'autres groupes.

En octobre 2013, un article du quotidien allemand Der Spiegel nous apprenait que la NSA avait espionné le GSM d'Angela Merkel. Et Wikileaks révélait que les GSM non sécurisés de Jacques Chirac, Nicolas Sarkozy et François Hollande avaient, eux aussi, été espionnés par cette agence.[11]

La stratégie de Keith B. Alexander à la tête de la NSA était claire : « tout collecter ». Il l'avait mise en application à partir de 2005 en Irak pour l'étendre à tous les pays au nom de la lutte contre le terrorisme. Le système qu'avait bâti le gouvernement américain d'Obama à l'époque était proprement démentiel. Son objectif n'était rien moins que l'élimination complète, à l'échelle planétaire, de toute vie privée électronique.[12]

Aujourd'hui, ces choses sont connues. Ce programme PRISM a été dénoncé par la presse internationale, condamné par les instances européennes. Le 2 juin 2015, Obama signa l'USA Freedom Act qui remplaça le Patriot Act et qui réduisit la capacité du gouvernement à collecter des données. Ce document stipule que le gouvernement ne peut collecter des données qu'après avoir soumis une requête publique à la Cour FISA (Foreign Intelligence Surveillance Court, aussi appelée FISC, cour fédérale américaine créée par la loi Foreign Intelligence Surveillance Act (FISA) de 1978 pour superviser les demandes mandats autorisant la surveillance, par les agences fédérales judiciaires américaines de présumés agents de renseignements étrangers sur le sol américain. C'est largement grâce au courage d'Edward Snowden et des journalises du Guardian et du Washington Post que cela a été rendu possible.

Néanmoins, les atteintes à la vie privée ne sont pas uniquement le fait des Etats surveillant leurs citoyens. Aujourd'hui, ce sont des compagnies privées qui détiennent la majeure partie des informations sur la vue privée des internautes. Et cela a été rendu possible grâce à des services auxquels les internautes ont volontairement confié quantité d'informations relatives à leur vie privée. On assiste à une forme d'abdication collective à ce sujet.

A cet égard, Andreas Weigend, l'ancien directeur scientifique d'Amazon affirmait :

> *« Ce que le KGB ne pouvait obtenir des gens sous la torture, les gens le publient désormais volontairement et en connaissance de cause sur Facebook ».*[13]

Pourquoi ? Parce que, affirme Weigend, en confiant nos données, nous obtenons en échange de bien meilleurs produits et services. Selon lui, il est temps, aujourd'hui « de redéfinir ce qu'on entend par vie privée ».

Mark **Zuckerberg**, fondateur et CEO de Facebook, partage la même analyse sur cette évolution. Selon lui, **a vie privée n'est désormais plus une norme sociale**, ou, plus exactement, il affirme :

> *« Les gens aiment réellement désormais non seulement partager plus d'informations et de différentes sortes mais aussi plus ouvertement et avec plus de gens (...). Cette norme sociale est quelque chose qui a évolué avec le temps ».*[14]

11 M. Dugain & Ch. Labbé, ***L'homme nu. La dictature invisible du numérique****, Plon, 2018 (2016), p.70*

12 Gl. Greenwald, ***Nulle part où se cacher, L'affaire Snowden par celui qui l'a dévoilé au monde****, JC Lattès, 2014, p.137*

13 BBC News, ***Is privacy dead in an online world ?****, 6 October 2017, www.bbc.com/news/technology-41483723*

14 The Guardian, ***Privacy no longer a social norm, says Facebook founder****, 11 Jan 2010, www.theguardian.com/technology/2010/jan/11/facebook-privacy*

Pourquoi les individus fournissent-ils leurs données privées avec une relative indifférence ? Comment expliquer ce phénomène ? La juriste Antoinette Rouvroy fournit plusieurs explications.[15]

1. Parce qu'ils adorent leurs **objets connectés.** Ce sont quasiment des « objets transitionnels » (comme les peluches des enfants) qui les éloignent de l'angoisse de la séparation ;

2. Parce que, de cette manière, ils peuvent être en contact, observer et **surveiller** leurs proches.

3. En raison de l'idéologie de la **transparence** : on se sentirait tenu d'affirmer qu'on n'a rien à cacher (et donc rien à se reprocher)

4. Parce que nous vivons de plus en plus dans une « **société de notation** » : les gens s'évaluent, se likent, se mesurent dans les plus infime détails de leur vie quotidienne (sport, alimentation, santé, soirées, voyages, etc.).

Nous pensons pouvoir ajouter à cette liste une cinquième raison : les sociétés commerciales nous offrent des **produits et services gratuitement** à condition que nous leur fournissions des données et même que nous les autorisions à les communiquer à d'autres. Nous le faisons et nous signons sans les lires des contrats parce que nous sommes impatients d'accéder à un service, une application où même une simple vidéo un peu alléchante. Tout le monde ou presque, juristes y compris, signe les yeux fermés.

On parle parfois du « **pêché originel d'internet** ». De quoi s'agit-il ? Lorsqu'Internet a commencé à se développer, on a cru un moment que les particuliers allaient acheter beaucoup de services en ligne. En réalité, il est vite apparu que les gens préféraient visiter les sites et communiquer entre eux. Un certain nombre d'entreprises ont alors réalisé qu'exploiter les données de leurs utilisateurs pourrait être un moyen rentable pour financer, grâce à la publicité, les activités qu'ils proposaient gratuitement.

On connaît la sentence : « **Si quelque chose sur internet est gratuit, c'est que c'est vous le produit** ».

Cela dit, les compagnies commerciales ne s'intéressent pas tant à qui vous êtes, ce que vous faites et ce que vous pensez qu'à la question de savoir si on peut anticiper et orienter votre comportement de consommateur. Comme l'écrit Luciano Floridi, personne ne se soucie de savoir qui vous êtes sur le Web, du moment qu'on puisse vous classer dans telle ou telle catégorie d'acheteur.

Nous sommes, dit-il, un peu comme les « âmes mortes » de Gogol mais « avec un portefeuille ». Il existe, dès lors, un marché qui vend ces âmes mortes. Ce sont les compagnies qui récoltent les profils, qui les ordonnancent, les traitent et les revendent. Mais, contrairement à un mythe largement répandu (et auquel le chapitre VI tente de tordre le cou), la valeur marchande d'une personne (ou, plutôt, ses caractéristiques, son « âme digitale) ne vaut pas grand-chose. Luciano Floridi s'est amusé à faire le calcul : à titre individuel, une personne vaut… 0,3723 dollars.[16] Comme on le verra dans le chapitre consacré aux big data et aux algorithmes, les données ont peu de valeur et même d'intérêt à titre individuel. C'est quand elles sont massives et qu'on parvient à établir des corrélations entre elles qu'on parvient à dégager de la valeur.

Dès lors, le danger que représentent les grandes compagnies, ne résiderait pas dans le fait qu'elles pénètrent dans votre vie privée pour la divulguer. Il réside plutôt, comme nous le verrons, dans le fait qu'ils risquent de **gouverner vos comportements grâce à leurs algorithmes**.

Mais, sans encore trop anticiper sur notre réflexion du chapitre IV, nous pouvons constater qu'à l'ère du big data, ce qui constitue notre vie privée, n'intéresse pas les algorithmes. Certes, ils traquent, enregistrent et analysent nos moindres comportements (le nombre de kilomètres que nous marchons, que nous courons, les achats que nous faisons, les lieux et l'heure où nous les faisons, les endroits où nous déjeunons, les films et les livres que nous achetons, etc.) mais ils n'ont plus besoin de connaître qui nous sommes. Ces données sont d'ailleurs anonymisées. Comme l'explique le professeur Dominique Cardon, le marketing n'a plus besoin de connaître notre état civil, notre âge, notre domicile, notre profession, nos emprunts, etc. pour vendre ses produits. Il se basera juste sur le profil numérique qui traduit nos préférences (et qui permet d'anticiper nos besoins et envies futurs) sans se préoccuper de qui nous sommes.[17]

Comment expliquer cela ? Ce qui intéresse les marketeurs, c'est de s'adresser à une cible (les jeunes, les mères de famille de plus de 40 ans, les indépendants, les retraités, les citadins fortunés, etc.). Dès lors, auparavant, il était intéressant de savoir dans quelle catégorie on pouvait ranger telle ou telle personne et on s'employait alors à le deviner grâce à certains indices. On pénétrait ainsi dans la vie privée des gens. Aujourd'hui, on ne s'intéresse plus à la personne en tant qu'elle serait membre d'un groupe. L'algorithme, sur base des requêtes formulées, des sites visités, des achats passés, etc. par un internaute, parvient à le classer dans un profil qui permet d'anticiper ce qu'il désirera et fera. Pourquoi ? Parce que ce profil est partagé par un certain nombre de personnes qui lui ressemblent : « **le futur de l'internaute est prédit par le passé de ceux qui lui ressemblent** ».[18]

---

*15 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, **Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet**, Larcier, 2018, pp. 418-420*

*16 L. Floridi, **The fourth revolution. How the infosphere is reshaping human reality**, Oxford University Press, 2014, pp. 98-100*

*17 D. Cardon., **A quoi rêvent les algorithmes. Nos vies** à **l'heure des big data**, Seuil, 2015, p.37*

*18 D. Cardon., **A quoi rêvent les algorithmes. Nos vies** à **l'heure des big data**, Seuil, 2015, p.34*

L'internaute laisse constamment des traces. Nous émettons constamment des données que nous le voulions ou non. Ce sont des indications très précieuses pour ces compagnies. On n'enverra pas des publicités qui vous intéresseront probablement en raison de la catégorie sociale, assez grossière, à laquelle vous êtes censé appartenir. Plus de calcul de probabilité ici. L'algorithme collera au plus près de votre comportement réel et ne vous proposera que des choses qui risquent réellement de satisfaire à un besoin que vous éprouvez ou que vous n'éprouvez pas encore mais qui vous correspond parce que, en traquant tous vos comportements depuis des années, **ces algorithmes finissent par vous connaître mieux que vous-mêmes**. **Mais, pour cet algorithme, vous êtres juste une adresse de protocole informatique. Il se désintéresse de votre nom, de votre âge**, du quartier où vous vivez, du pays où vous résidez, de votre profession et même de vos revenus. Vous n'êtes qu'un comportement régulier et prévisible à ses yeux. **En un sens, votre vie privée est mieux protégée car lui, ses concepteurs et les vendeurs s'en désintéressent totalement**.

Cela dit, ce nouveau modèle est réellement en rupture avec l'ancien système de récolte des données. Selon la professeur Antoinette Rouvroy,[19] les grands principes européens de la protection des données entrent en opposition frontale avec la logique des big data. En effet, ces principes traditionnels sont les suivants :

- **la minimisation dans la collecte des données** : on ne prélève que ce qui est nécessaire au projet ;
- **la finalité** : on ne collecte ces dernières qu'en raison d'un objectif déterminé, déclaré et légitime ;
- **la limitation dans le temps** : on supprime les données une fois le but atteint.

Les big data, au contraire, sont prélevées en fonction des principes suivants :

- **la maximalisation** : on collecte tout sans discrimination ;
- **l'absence de finalité** : la collecte est automatique et ne correspond à aucun objectif ; déterminé. Ce n'est qu'en cours de route que l'utilité se manifestera à partir du « data mining » (l'analyse des donnée massives par des procédés automatiques ou semi-automatiques) ;
- **la conservation illimitée** : on n'efface aucune donnée car la logique est d'accroître la quantité vu que la valeur de la base de données augmente avec le temps à mesure qu'elle s'accroît et que les algorithmes mettent en relation des points de données éloignés dans le temps et dans l'espace.

Il faut effectivement se montrer nuancé. Ces principes traditionnels datent évidemment d'une période antérieure à la révolution des big data. La vie privée est mieux protégée grâce à des techniques d'anonymisation et de pseudonymisation. Nous verrons dans le chapitre suivant que le récent Règlement Général de Protection des Données (RGPD) permettra probablement, même s'il est encore trop tôt pour le dire, d'assurer une protection efficace de la vie privée.

Mais il est vrai aussi qu'aux mains d'un Etat totalitaire, les données récoltées aujourd'hui pourraient se révéler un instrument d'oppression d'une puissance qui n'a jamais existé dans l'histoire. Comme le disait Isabelle Falque Pierrotin, présidente de la Commission Nationale Informatique et Liberté, « c'est probablement la première fois dans l'histoire que la technologie rend possible une surveillance de masse réelle mise en place par des acteurs publics et privés ».[20] En effet, le risque d'espionnage de nos vies est une réalité. Les données que les géants du net récoltent sur nous sont transférées dans le « cloud computing », c'est-à-dire dans le cœur de gros « data centers » principalement situés aux Etats-Unis.

Mais ce ne sont pas uniquement des règlements qui défendent la vie privée. Les citoyens sont les premiers à devoir lutter contre cette menace. Malheureusement, il arrive de plus en plus fréquemment que des personnes affirment que la vie privée, finalement, est un concept obsolète et qu'on pourrait très bien s'en passer.

Trois arguments sont généralement utilisés pour relativiser voire contester la pertinence de la protection de la vie privée :

- *« La vie privée est un frein à l'innovation et au développement économique » ;*
- *« De toute façon, je n'ai rien à cacher » ;*
- *« La vie privée est une entrave à la sécurité à l'heure de la menace terroriste ».*

Examinons ces trois arguments.

1. PREMIER ARGUMENT : « LA VIE PRIVÉE EST UN FREIN À L'INNOVATION ET AU DÉVELOPPEMENT ÉCONOMIQUE ».

On entend parfois, par exemple, que le fait de récolter des données massivement sans état d'âme (et en l'absence de législations trop contraignantes sur la vie privée) aurait permis aux Etats-Unis aujourd'hui (et ce serait également le cas de l'Asie) de faire des avancées considérables dans le domaine de l'intelligence artificielle car les données massives sont le principal aliment pour faire tourner ces programmes. On dit aussi que si nous sommes tellement en retard au niveau de l'économie numérique et qui si les plus grands géants numériques (GAFAM) sont américains, c'est parce qu'ils peuvent se procurer ces données beaucoup plus facilement que nous et développer ce secteur en conséquence.

*19 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, **Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet**, Larcier, 2018, pp. 427*
*20 Documentaire « **Nothing to hide** », 2017, 1h11'49', www.youtube.com/watch?v=djbwzEIv7gE&t=2911s*

C'est une idée fausse. L'avance incontestable des Etats-Unis dans le domaine numérique s'explique par d'autres causes. Le Centre Jean Gol y a consacré une analyse.[21] Elles tiennent principalement dans la culture entrepreneuriale longue de plusieurs décennies qui s'est développée dans la Silicon Valley et d'excellents choix industriels qui ont été faits par les investisseurs.

Quoi qu'il en soit, pour les raisons explicitées quelques paragraphes plus haut, le développement de l'intelligence artificielle en particulier et de l'économie numérique en général, peut très bien s'opérer avec des données anonymisées ou pseudonymisées. Il n'y a strictement aucun lien de causalité entre la violation de la vie privée et l'innovation économique.

2. SECOND ARGUMENT : « DE TOUTE FAÇON, JE N'AI RIEN À CACHER »

On entend parfois : « on copie mes données. Où est le mal ? Où est le dommage ? ». « De toute façon, je n'ai rien à cacher ». A cela, le lanceur d'alerte américain Thomas Drake réplique :

> *« Si vous dites, 'je n'ai rien à craindre car je n'ai rien à cacher. Donc, tout va bien'. Alors je réponds : 'Donnez-moi les clés de votre voiture, de votre domicile, vos mots de passe, vos comptes bancaires, votre carnet de santé, toutes vos données personnelles. Je les garderai dans une boîte sécurisée'. Pas une personne, parmi les milliers que j'ai interrogées, au cours de nombreuses conférences n'a accepté ma proposition. Pas une ! Personne n'était prêt à me confier ses données. Et, pourtant, je leur demandais leur consentement… ».* [22]

Les gens, même s'ils ne font rien de mal, seraient-ils d'accord qu'on capte leurs conversations privées et qu'on les diffuse publiquement sur le net ? Seraient-ils d'accord qu'on place une caméra dans leur salle de bain ?

En réalité, les gens confondent « je n'ai rien à cacher » et « je n'ai rien à me reprocher ». Ils ne se rendent pas compte que beaucoup de choses de notre vie privée peuvent être utilisées à notre détriment. Non, ce qui doit être protégé ce n'est pas seulement la somme de nos turpitudes. Ce qui doit être protégé, c'est notre intimité.

Par ailleurs, est-on si sûr qu'on n'a strictement « rien » à se reprocher ? Tout adulte a probablement déjà commis une infraction pénale dans sa vie. Même mineure, comme par exemple le vol d'un matériel de bureau ou une infraction au code de la route. Enfin, comme tout un chacun s'exprime généralement de manière relâchée dans sa vie privée, il est toujours possible d'exploiter des éléments de la vie privée de quelqu'un pour lui prêter des intentions qu'il n'a pas ou pour le désigner à la vindicte de la foule. Comme le disait le cardinal de Richelieu qui avait coutume d'espionner ses semblables (c'est, du moins, une phrase qu'on lui attribue) :

> *« Qu'on me donne six lignes écrites de la main du plus honnête homme, j'y trouverai de quoi le faire pendre ».*

Ce réflexe « je n'ai rien à cacher » est une phrase d'enfant gâté vivant dans une démocratie où les droits sont protégés. Non seulement, nos droits et libertés ne sont protégés que si cette vigilance se transmet de génération en génération. Mais cette phrase est profondément égoïste car elle méconnaît le sort de quantité de gens qui sont persécutées dans le monde pour leurs opinions et convictions. Quantité de gens qui combattent la dictature, la théocratie, l'intolérance, etc. doivent agir et s'exprimer clandestinement pour ne pas finir en prison ou dans la pièce d'un bourreau.

Comme le dit Edward Snowden :

> *« Dire que la protection de son intimité ne nous intéresse pas est sans doute l'acte le plus asocial qui soit. Ce n'est pas parce qu'on est pas intéressé à user d'un droit qu'il faut affirmer qu'il n'a aucun intérêt. C'est comme dire que « la liberté d'expression n'a aucun intérêt car je n'ai rien à dire. Quand vous occupez le somment de la pyramide de l'échelle sociale - imaginons que vous êtes blanc, mâle, âgé - vous n'avez pas besoin de ces droits car la société est dessinée de façon à répondre à vos besoins. Ce sont toujours les minorités qui encourent le plus de risques ».*[23]

Enfin, ceux qui profèrent cette phrase ne mesurent pas les multiples dangers qui peuvent exister dans une société où la protection de la vie privée serait abolie. Par exemple celui d'une pratique qui, comme c'est aujourd'hui le cas en Chine, consisterait à « scorer » les habitants, à leur donner une note de fiabilité rendue publique. Note qui serait attribuée à chacun grâce à toutes les données concernant son comportement récoltées grâce à la surveillance de masse.

En laissant faire les choses en estimant que l'on n'a « rien à cacher », on permettrait que certains (des banques, des compagnies d'assurance, des agences immobilières, etc.), en recoupant toute une série d'informations sur notre compte, finissent, par exemple, par attribuer à chacun une « note de solvabilité ».

---

21 Ch. Cockshaw, ***Le modèle de la Silicon Valley est-il pertinent pour l'Europe ?****, Analyse du Centre Jean Gol, 2018*

22 Documentaire « ***Nothing to hide*** *», 2017, 13'18", www.youtube.com/watch?v=djbwzEIv7gE&t=2911s*

23 Documentaire « ***Nothing to hide*** *», 2017, 18'45", www.youtube.com/watch?v=djbwzEIv7gE&t=2911s*

3. TROISIÈME ARGUMENT : « LA VIE PRIVÉE EST UNE ENTRAVE À LA SÉCURITÉ À L'HEURE DE LA MENACE TERRORISTE ».

Cet argument est évidemment le principal argument qui a été utilisé par la NSA et tous les faucons américains le lendemain des attentats du 11 septembre pour faire passer le Patriot Act, le programme PRISM et toutes les mesures de surveillance de masse dont nous avons parlé.

Par rapport à cet argument, on peut rétorquer que les libertés fondamentales ne sont pas négociables. C'est la fameuse phrase de Benjamin Franklin :

> *« Un peuple prêt à sacrifier un peu de liberté pour un peu de sécurité ne mérite ni l'une ni l'autre et finit par perdre les deux » (Lettre de l'Assemblée de Pennsylvanie au Gouverneur de Pennsylvanie, 1755).*

Une société qui sacrifie ses libertés est d'ailleurs l'un des objectifs généralement poursuivis par les groupes terroristes. Cela conduit à la mise en place d'une Etat de surveillance et un tel Etat est profondément nocif. C'est contre un tel Etat que combattent à juste titre des gens tels que Edward Snowden et Julian Assange.

Par ailleurs, il apparaît que la surveillance de masse n'est pas efficace pour prévenir des attentats terroristes. Ce qui est autrement plus efficace, c'est la surveillance ciblée. Une étude de la *New American Foundation*, organe politiquement situé au centre estimait que le programme PRISM n'avait eu « aucun impact perceptible sur la prévention d'actes de terrorisme ».[24]

Le Washington Post relevait pour sa part que « les méthodes d'application de la loi et d'investigation traditionnelles suffisaient à fournir le renseignement ou la preuve nécessaires à l'ouverture d'une enquête ».[25]

La « collecte totale » n'a pas permis de détecter l'attentat à la bombe du marathon de Boston en 2012 ou les tentatives d'attentat à la bombe du Times Square, contre le métro new-yorkais ou contre l'avion de ligne au-dessus de Detroit. Tous ces attentats ont pu être déjoué par les forces de police et des passants vigilants. Cette collecte totale n'a pas permis d'éviter non plus les nombreuses fusillades mortelles sur le territoire américain.

Côté français, le constat est le même. En 2012, naissait la DCRI (Direction Centrale du Renseignement Intérieur), rebaptisé depuis en DGSI, qui fusionnait le contre-espionnage et les renseignements généraux. Le résultat a été un affaiblissement considérable du renseignement de proximité qui était pourtant excellent en raison d'un impressionnant maillage territorial. Le DGSI prenait la NSA comme le modèle à suivre car elle pratiquait la surveillance absolue. Ce choix stratégique n'a pas empêché la France d'échapper par la suite à quatre attentats djihadistes et notamment le massacre du 13 novembre 2015, le plus meurtrier jamais commis en France avec 130 morts.[26]

En matière de contre-terrorisme, il apparaît que la surveillance de masse est contreproductive : elle complique la détection et l'endiguement des attentats. Elle ne fait que masquer les complots bien réels à propos desquels échangent entre eux de vrais terroristes.[27]

RÉSISTER À L'ESPIONNAGE DES GAFAM

Il est faux de penser que nous ne pouvons rien faire et que nous aurions déjà perdu la bataille. Il existe en effet des outils alternatifs à ceux proposés par les GAFAM. Selon Alison Macrina, responsable de « Tor Project » (à savoir un moteur de recherche qui ne récolte pas des données de ses utilisateurs) :

> *« pour toutes les technologies actuelles, pour tout produit Microsoft, Apple, Google, Facebook, etc., il y a probablement une alternative et, dans la plupart des cas, l'alternative est très bonne. Il y a un certain nombre de réflexes à avoir comme, notamment, utiliser des moteurs de recherche qui n'enregistre pas vos données comme Disconnect ou DuckDuck go. Car Google est un puissant appareille de surveillance qui nous offre de très bons services, mais qui garde toutes nos données. Je recommande l'utilisation de Signal comme messagerie instantanée. Tor Browser est le navigateur internet qui permet de protéger certaines informations identifiables vis-à-vis de l'hébergeur, d'un site, de tous ceux qui surveillent votre trafic internet ou encore de votre entreprise ou d'un Etat ».*[28]

A la base, TOR est né de l'US Navy au milieu des années 90. La marine américaine désirait un instrument qui rende ses communications intraçables. Par la suite, lorsque ce programme a été arrêté, il a été récupéré et transformé par « l'Electronic Frontier Foundation », une association américaine qui défend les libertés sur internet. Ce moteur de recherche anonymise les connexions et masque à la fois leur contenu, leur point d'origine et leur destination. Il serait aujourd'hui utilisé par deux millions de personnes.

---

24 Gl. Greenwald, ***Nulle part où se cacher, L'affaire Snowden par celui qui l'a dévoilé au monde****, JC Lattès, 2014, p.286*
25 Ibidem
26 M. Dugain & Ch. Labbé, ***L'homme nu. La dictature invisible du numérique****, Plon, 2018 (2016), pp.16-17*
27 Gl. Greenwald, ***op.cit.****, p.288*
28 Documentaire « Nothing to hide », 2017, 1h14', www.youtube.com/watch?v=djbwzEIv7gE&t=2911s

TOR permet également d'accéder au Deep Web par une porte dérobée. Le Deep Web (ou Web profond ou « Web caché ») est souvent assimilé au Dark Net. Parfois à dessein pour dissuader les gens de sortir du réseau bien balisé où on peut suivre tous leurs mouvements. Certes, le Dark Net existe aussi et il est effectivement utilisé par des mafias, des malfaiteurs et autres acteurs se livrant à des activités illégales. Mais, le Deep Web, même s'il contient le Dark Net, est avant tout un espace clandestin, dans lequel naviguent un grand nombre de personnes qui désirent échapper à la surveillance omniprésente des grands géants du Net et des agences de renseignement. Des dissidents politiques, des militants des droits de l'homme, des lanceurs d'alerte, des journalistes, etc.[29]

Selon Fabrice Epelboin, entrepreneur et enseignant à Sciences-Po Paris, si la moitié de la population discutait en utilisant des outils de cryptage, elle protégerait l'autre en noyant les autorités sous une masse de conversations cryptées « non-intéressantes ». Selon Jérémie Zimmerman, hacker et cofondateur de « La Quadrature du Net », nous avons besoin de logiciels libres, de logiciels appartenant à l'humanité, pour lequel chacun a la même liberté que l'auteur pour comprendre ce que fait le programme et pour le modifier. Nous avons aussi besoin de services décentralisés pour savoir où vont nos données, pour savoir ce qui est enregistré et où cela est enregistré. Il est possible de mettre en place un système permettant les conversations où chacun génère une clé de chiffrement et l'échange avec son correspondant.[30]

*29 M. Dugain & Ch. Labbé, **L'homme nu. La dictature invisible du numérique**, Plon, 2018 (2016), p.188 et s.*
30 Documentaire « Nothing to hide », 2017, 1h16', www.youtube.com/watch?v=djbwzEIv7gE&t=2911s


NODE 04
NODE 05
NODE 02
BLOCK 01
BLOCK 01

III. NORMES ET AUTORITÉS ASSURANT LE RESPECT DE LA VIE PRIVÉE

Stéphane Tellier

1. QUELLES SONT LES NORMES EN DROIT BELGE QUI PROTÈGENT LA VIE PRIVÉE ?

INTRODUCTION ET APPROCHE HISTORIQUE

Le droit à la vie privée ne peut plus laisser personne indifférent tant il s'est immiscé dans toutes les sphères du droit et, au-delà de cela, dans le quotidien de chacun.

A l'ère du numérique, le développement des technologies de traitement des données, leur caractère transversal ainsi que leur dimension internationale et transfrontalière suscitent légitimement les craintes et les interrogations.

Le droit à la vie privée est notamment consacré par l'article 8 de la convention européenne des droits de l'homme. Cette disposition garantit le développement de la personnalité de chaque individu dans ses relations avec les autres sans ingérence extérieure. Il s'agit d'un véritable droit au développement personnel et à l'autonomie. Son champ d'application s'est considérablement élargi au fil du temps.

Selon F. Rigaux, le droit au respect de la vie privée n'était consacré par aucune des constitutions libérales du 19ème siècle, du moins en termes exprès[31]. C'est davantage au travers de l'évolution jurisprudentielle et de l'œuvre doctrinale qu'un ensemble de droits de la personnalité se sont cristallisés et se sont vus protégés, principalement au départ de l'inviolabilité du domicile et du secret des lettres. Ce n'est qu'au 20ème siècle que ces nouvelles prérogatives ont gagné une reconnaissance dans les constitutions nationales et les conventions internationales relatives aux droits de l'homme. Cette nouvelle notion contemporaine de droit à la vie privée s'est également forgée sur la base de la théorie des droits de la personnalité.

Cette évolution a élevé le droit à la vie privée au statut de véritable condition d'existence des autres libertés démocratiques[32].

LES SOURCES DU DROIT À LA VIE PRIVÉE

Le droit à la vie privée est consacré tant au niveau national qu'international.

Au niveau national, le siège du droit au respect de la vie privée se situe à l'article 22 de la Constitution belge depuis sa révision de 1994 : « Chacun a droit au respect de sa vie privée et familiale, sauf dans les cas et conditions fixés par la loi. La loi, le décret ou la règle visée à l'article 134 garantissent la protection de ce droit »[33]. Il faut ajouter à cet article, ses dispositions constitutionnelles mères, soit l'inviolabilité du domicile (article 15) et le secret des lettres (article 29).

Il y a un véritable parallélisme à faire avec les dispositions conventionnelles au niveau international et européen, conséquence de l'obligation dans le chef du législateur national de prendre des mesures positives pour veiller à protéger le droit à la vie privée.

La notion de respect de la vie privée s'exprime dans d'autres dimensions juridiques, comme le secret professionnel, la surveillance caméra ainsi que, bien sûr, la collecte de données électroniques et la protection des données personnelles que nous examinerons au paragraphe 5.

Comme nous le verrons *infra*, la législation belge est riche de dispositions relatives au droit à la vie privée et plus spécifiquement dans le cadre du traitement des données à caractère personnel.

Pour faire l'inventaire des sources de droit international et européen en matière de droit à la vie privée, il faut essentiellement citer :

- les articles 17 et 23 du Pacte international relatif aux droits civils et politiques du 16 décembre 1966, d'effet direct en droit belge[34] ;
- l'article 8 de la convention européenne des droits de l'homme et des libertés fondamentales (C.E.D.H.) du 4 novembre 1950, ayant valeur d'ordre public en droit belge[35] ;

31 F. Rigaux, ***La protection de la vie privée et des autres biens de la personnalité****, Bruxelles, Bruylant, L.G.D.J., 1990.*

32 Y. Poullet, « La protection des données : entre libertés, droits subjectifs et intérêts légitimes », in ***Liber Amicorum P. Martens****, Bruxelles, Larcier, 2007, p. 139.*

33 C'est au sens formel et non uniquement matériel qu'il faut entendre le terme « loi ». Seule une loi peut prévoir une limitation ou une exception au principe du droit à la vie privée. C. Arb., arrêt 202/2004 du 21 décembre 2004, point B.5.4.

34 J. Velu et R. Ergec, « La convention européenne des droits de l'homme », ***R.P.D.B.****, complément VII, Bruxelles, Bruylant, 1990, n°10 et n°100.*

35 J. Velu et R. Eergec, ***op. cit.****, n°99*

- l'article 7 de la Charte des droits fondamentaux de l'Union européenne, qui se calque sur l'article 8 C.E.D.H. ; l'article 8 de la Charte étant même spécifiquement dédié à la protection des données à caractère personnel : « 1. Toute personne a droit à la protection des données à caractère personnel la concernant. 2. Ces données doivent être traitées loyalement, à des fins déterminées et sur la base du consentement de la personne concernée ou en vertu d'un autre fondement légitime prévu par la loi. Toute personne a le droit d'accéder aux données collectées la concernant et d'en obtenir la rectification. 3. Le respect de ces règles est soumis au contrôle d'une autorité indépendante » ;

- la Convention internationale relative aux droits de l'enfant de 1989 qui édicte en son article 16 que : « *1. Nul enfant ne fera l'objet d'immixtions arbitraires ou illégales dans sa vie privée, sa famille, son domicile ou sa correspondance, ni d'atteintes illégales à son honneur et à sa réputation ; 2. L'enfant a droit à la protection de la loi contre de telles immixtions ou de telles atteintes* ».

LES COMPOSANTES DU DROIT À LA VIE PRIVÉE

La vie privée recouvre classiquement quatre composantes : la vie privée individuelle, le secret de la correspondance, la vie familiale et le domicile. Seules les deux premières seront évoquées, en lien direct avec le sujet de cette étude.

a) La vie privée individuelle

C'est la dimension première du droit au respect de la vie privée, son enjeu démocratique principal : le droit à l'intimité de la vie privée, classiquement défini comme le droit de « vivre à l'abri des regards étrangers ». Le droit protège cette intimité tant à l'égard des atteintes des autorités publiques que de celles provoquées par des tiers. C'est le cœur du droit à la vie privée[36].

Selon B. Docquir, la jurisprudence de la Cour européenne des droits de l'homme a défini la vie privée individuelle comme un véritable droit à l'autonomie personnelle et au libre développement de sa personnalité[37].

Si la notion de vie privée s'est considérablement élargie au fil du temps, c'est aussi grâce à la jurisprudence de la Cour européenne des droits de l'homme, laquelle a reconnu, dès un arrêt *Niemetz c. Allemagne* du 16 décembre 1992, qu'elle ne se limitait pas à « un cercle intime où chacun peut mener sa vie personnelle à sa guise » mais qu'elle comprend aussi « le droit pour l'individu de nouer et de développer des relations avec ses semblables ». Par conséquent, le droit à la vie privée s'applique également aux activités professionnelles ou commerciales, ce qui revêt une importance particulière lorsque l'on analysera cette notion à l'aune du traitement des données à caractère personnel.

b) Le secret de la correspondance

Aujourd'hui, comme le souligne B. Docquir, il serait plus exact de parler de secret des communications[38], tant la notion fondamentale de correspondance a évolué, en particulier grâce à la jurisprudence de la Cour européenne des droits de l'homme qui inclut tous les moyens de communication, orale ou écrite, quel que soit le procédé utilisé.

Le secret des lettres est protégé par une série de dispositions légales, au premier rang desquelles l'article 29 de la Constitution, mais aussi les articles 130 et suivants de la loi du 21 mars 1991 portant réforme de certaines entreprises publiques économiques. Le Code pénal et le Code d'instruction criminelle encadrent également les pouvoirs du Procureur du Roi et du Juge d'instruction pour intercepter et ouvrir les lettres confiées à la Poste. Au niveau privé, les lettres sont protégées par le biais des droits de la personnalité, du droit à la vie privée et même du droit de propriété de la lettre qui existe dans le chef du destinataire.

Dans la sphère des communications électroniques, le législateur est également intervenu pour encadrer les interceptions de communications.

LES LIMITES DU DROIT À LA VIE PRIVÉE

Le droit à la vie privée n'est pas un droit absolu. A côté des prérogatives que le droit à la vie privée accorde à l'individu, il faut examiner les conditions à respecter pour le limiter, au bénéfice des droits d'autrui ou dans l'intérêt général.

Si l'article 8 §2 de la convention européenne des droits de l'homme autorise qu'il puisse y avoir des limitations à la protection de la vie privée et familiale, du domicile et de la correspondance, le principe est que toute ingérence dans la vie privée doit être autorisée par la loi, servir un but légitime et être nécessaire dans une société démocratique[39]. Le fameux « RGPD » fait respecter aujourd'hui ces exigences de légalité, de légitimité et de proportionnalité.

Notons que cette règle s'applique non seulement entre l'Etat et les citoyens mais également dans les relations entre les personnes de droit privé, citoyens ou entreprises, ce que la jurisprudence belge reconnaît depuis longtemps et que l'article 2 de la loi « vie privée » du 8 décembre 1992 consacrait déjà.

---

36 P. Waschmann, « Le droit au secret de la vie privée », in F. Sudre, ***Le droit au respect de la vie privée au sens de la convention européenne des droits de l'homme****, coll. Droit et Justice, n°63, Bruxelles, Bruylant, 2006, pp. 119-155.*

37 B. Docquir, ***Actualités du droit de la vie privée****, Bruxelles, Bruylant, 2008, pp. 7-9.*

38 B. Docquir, ***op. cit.****, p. 10-12.*

39 Pour atteindre l'un des autres buts énumérés par la C.E.D.H., soit le bien-être économique, la sécurité nationale, la prévention des infractions, la protection de la santé ou de la morale ou encore la protection des droits et libertés d'autrui.

a) Une condition de légalité

Toute intrusion dans le droit à la vie privée doit être prévue par la loi. Cette « loi », qu'elle puisse être une règle non écrite ou jurisprudentielle dans l'acception du terme par la C.E.D.H. ou une loi au sens formel du terme dans l'acception de l'article 22 de la Constitution, doit rendre cette intrusion prévisible pour le citoyen.

La loi doit également être accessible et précise. Pour la Cour européenne des droits de l'homme, cela s'apprécie *in concreto* : « le niveau de précision de la législation interne – qui ne peut en aucun cas prévoir toutes les hypothèses – dépend dans une large mesure du contenu de l'instrument en question, du domaine qu'il est censé couvrir et du nombre et du statut de ceux à qui il est adressé »[40].

La condition de légalité ne se suffit pas à elle-même. L'existence d'une loi ne justifie pas à elle seule toute intrusion dans la vie privée. Encore faut-il que cette loi soit légitime et proportionnelle au but recherché.

b) La condition de légitimité

La limitation du droit à la vie privée ne sera autorisée, au regard de l'article 8 §2 de la convention, que si elle sert l'un des buts recherchés par elle : le bien-être économique, la sécurité nationale, la prévention des infractions, la protection de la santé ou de la morale ou encore la protection des droits et libertés d'autrui. Cette liste est limitative mais peut être déclinée sur la base d'une interprétation souple par la jurisprudence. Si la loi poursuit effectivement l'un de ces buts, elle sera légitime.

L'objectif est, dès lors, de trouver un juste équilibre entre les droits et les devoirs en opposition. Cela nous amène à la condition de proportionnalité.

c) La condition de proportionnalité

Toute limitation au droit à la vie privée doit être absolument « nécessaire dans une société démocratique ». Il ne suffit donc pas que son but soit légitime mais il faut également que la loi en question soit raisonnable et pertinente pour poursuivre ce but, en restant dans la mesure de ce qui est strictement nécessaire. Le sens de cette condition est évidemment d'empêcher toute application arbitraire et de préserver l'Etat de droit. Là aussi, cette condition va devoir s'apprécier *in concreto*.

LA PROTECTION DE LA VIE PRIVÉE À TRAVERS LA PROTECTION DES DONNÉES À CARACTÈRE PERSONNEL

Ce sont bien entendu toutes les possibilités de profilage, d'identification et d'analyse des comportements, des habitudes, des choix et des préférences des citoyens qui appellent un pouvoir de contrôle sur cette masse d'informations très personnelles accumulées grâce à l'évolution des nouvelles technologies.

a) La législation applicable

Nous avons vu *supra* les sources fondamentales du droit à la vie privée. Nous nous attarderons ici plus spécifiquement à la protection des données à caractère personnel.

Au niveau national :

- La loi du 11 mars 2003 sur certains aspects juridiques des services de la société de l'information, prévoit des règles en matière de courrier électronique, y compris le courrier électronique non sollicité.
- La loi du 13 juin 2005 relative aux communications électroniques, fixe les principes applicables au secret des communications, au traitement des données et à la protection de la vie privée des utilisateurs.
- La loi du 3 décembre 2017 portant création de l'Autorité de protection des données.
- La loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel.

Au niveau européen, plusieurs directives européennes encadrent la protection des données à caractère personnel, comme :

- la directives 95/46/CE du Parlement européen et du Conseil du 24 octobre 1995 traçant les lignes directrices fondamentales du traitement des données en Europe, avant son abrogation au 25 mai 2018 par le RGPD ;
- la directive 2002/58/CE du Parlement européen et du Conseil du 12 juillet 2002 relative au traitement des données dans le cadre des communications électroniques ;
- la directive 2006/24/CE du Parlement européen et du Conseil du 15 mars 2006 sur la conservation des données générées ou traitées dans le cadre de la fourniture de services de communications électroniques accessibles au public ou de réseaux publics de communications ;
- la directive 2009/136/CE du Parlement européen et du Conseil du 25 novembre 2009 modifiant la directive 2002/22/CE concernant le service universel et les droits des utilisateurs au regard des réseaux et services de communications électroniques ;
- la directive 2016/680 du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales et à la libre circulation de ces données.

40 C.E.D.H., 26 octobre 2000, ***Hassan et Tchaouch c. Bulgarie.***

Le droit de l'Union européenne recèle également des dispositions réglementaires :

- le règlement 45/2001/CE du Parlement européen et du Conseil du 18 décembre 2000 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les institutions et organes communautaires et à la libre circulation de ces données ;

- et, last but not least, le règlement 2016/679 du 27 avril 2016, dit "GDPR" pour *General Data Protection* Regulation (Règlement général de protection des données ou « RGPD » en français) qui sera analysé de plus près, notamment quant à son efficacité ou son impact financier sur la santé économique des entreprises, dans un autre chapitre de cette étude.

Cela démontre que le droit communautaire s'est très tôt inquiété du traitement des données à caractère personnel dans le secteur des nouvelles technologies et qu'il a accompagné les évolutions de la technologie depuis les premiers pas de l'internet jusqu'à nos jours.

Au niveau international, citons les normes internationales pour la protection de la vie privée et des données à caractère personnel adoptées à Madrid en 2009 lors de la conférence internationale des commissaires à la protection des données, les normes ISO (normes pour la protection de la vie privée et des données à caractère personnel de l'ISO)[41], les lignes directrices de l'OCDE régissant la protection de la vie privée et les flux transfrontières de données de caractère personnel[42] ainsi que les textes de référence relatifs à la protection de la vie privée et des données à caractères personnel adoptés par l'ONU[43].

b) Le régime juridique du traitement des données à caractère personnel, désormais encadré par le Règlement général sur la protection des données personnelles (RGPD)

Selon nous, il convient de lire toutes les dispositions relatives aux données à caractère personnel et à leur traitement à l'aune du principe de proportionnalité qui émane de l'article 8 de la convention européenne des droits de l'homme. Elles doivent également être encadrées par des principes de loyauté et de transparence.

Pour vérifier la licéité du traitement de données à caractère personnel, le Règlement général sur la protection des données personnelles (RGPD) indique que le responsable du traitement doit avoir une bonne raison de traiter les données personnelles. Autrement dit, il doit pouvoir justifier du motif pour lequel il traite ces données.

Le RGPD prévoit six bases juridiques possibles à un traitement de données personnelles :

1. le responsable du traitement a obtenu des personnes concernées, un consentement *exprès* à ce traitement de leurs données personnelles ;

2. le responsable du traitement traite ces données personnelles parce qu'il y est tenu, en vertu d'une obligation légale ;

3. le traitement est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée ;

4. le traitement est nécessaire à l'exécution d'une mission d'intérêt public (ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement) ;

5. le traitement est nécessaire aux fins des intérêts légitimes poursuivis par le responsable du traitement ;

6. le traitement est nécessaire à l'exécution d'un contrat auquel la personne concernée est partie (ou à l'exécution de mesures précontractuelles prises à la demande de celle-ci).

Le motif 5) ne s'applique pas aux traitements effectués dans le secteur public : le RGPD l'interdit. Quant au motif 6), il relève plutôt du secteur privé et s'appliquera rarement aux traitements effectués au sein d'une structure publique, où les relations entre personnes physiques et structures publiques sont régies par la loi, rarement par des conventions ou contrats.

Par ailleurs, six principes généraux s'imposent également depuis l'entrée en vigueur du RGPD :

1. Licéité, loyauté et transparence du traitement

Les données personnelles doivent être « traitées de manière licite, loyale et transparente au regard de la personne concernée ». Cela signifie que les données ne doivent pas avoir été collectées et ne doivent pas être traitées, sans que la personne concernée en ait connaissance. Ce principe nécessite aussi de fournir aux personnes concernées plusieurs informations (sur le traitement de leurs données, mais aussi sur leurs droits).

2. La limitation des finalités

Les données doivent être « collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités ».

---

41 ISO/CEI 27001 Technologies de l'information - Techniques de sécurité - Systèmes de gestion de sécurité de l'information – Exigences + ISO/CEI 27002 Technologies de l'information - Techniques de sécurité - Code de bonnes pratiques pour la gestion de la sécurité de l'information.

42 Consultez : www.privacycommission.be/sites/privacycommission/files/documents/OCDE-lignes-directrices-protection-vie-priv%C3%A9e-flux-transfronti%C3%A8res.pdf.

43 Il y en a 3 : Principes directeurs pour la règlementation des fichiers informatisés contenant des données à caractère personnel (1990)

Toutefois, le RGPD ajoute que des données peuvent être traitées « à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques », même si elles n'avaient pas été initialement collectées à cette fin, sans qu'un tel traitement ultérieur soit considéré incompatible avec la finalité initiale de la collecte. C'est ce qui permet, par exemple, aux chercheurs de consulter des fonds contenant des données personnelles, sans enfreindre le droit des personnes concernées.

3. La minimisation des données

Seules doivent être collectées les données strictement nécessaires à la finalité du traitement.

4. L'exactitude des données

Les données à caractère personnel collectées doivent rester « exactes et, si nécessaire, tenues à jour ». Elles doivent sinon être rectifiées ou effacées.

5. La limitation de la conservation

Les données personnelles ne doivent pas être conservées au-delà de la durée nécessaire à la finalité du traitement.

6. L'intégrité et la confidentialité des données

Enfin, et c'est la grande nouveauté du RGPD, le responsable du traitement doit prendre « les mesures techniques ou organisationnelles appropriées » pour garantir la sécurité des données personnelles, y compris la protection contre le traitement non autorisé ou la perte des données.

Cette obligation de prendre les mesures techniques ou organisationnelles appropriées est la pierre angulaire du RGPD.

Le responsable du traitement doit, pour garantir l'intégrité et la confidentialité des données personnelles traitées, vérifier que l'organisation (humaine) et les moyens techniques (souvent, informatiques) soient suffisamment sûrs : cela suppose de sensibiliser et former les personnes chargées du traitement, de conclure si nécessaire des contrats avec plusieurs d'entre elles, et de mettre en œuvre des moyens techniques robustes (authentification avant accès, cryptage des données, etc.). Dans certains cas (données sensibles, traitement de données à grande échelle, etc.), le RGPD exige que soit menée une étude d'impact approfondie.

Comme auparavant, le responsable du traitement doit informer les personnes concernées de l'existence du traitement de leurs données et leur en indiquer la finalité. Il doit aussi informer les personnes concernées de leur droit d'accès, de rectification et d'effacement des données les concernant. Il s'agit d'un véritable « droit à l'oubli ».

Les données sensibles, ou « catégories particulières de données à caractère personnel » comme les nomme le Règlement européen, sont celles qui portent sur l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que les données génétiques, biométriques, concernant la santé, la vie sexuelle ou l'orientation sexuelle d'une personne physique.

Le traitement de ces catégories particulières de données est, en principe, interdit, sauf exceptions.

En effet, cette disposition sise à l'article 9.1. concernant les données sensibles ne s'applique pas si l'une des conditions suivantes est remplie (art. 9.2 RGPD) :

> *« a) la personne concernée a donné son consentement explicite au traitement de ces données à caractère personnel pour une ou plusieurs finalités spécifiques, sauf lorsque le droit de l'Union ou le droit de l'État membre prévoit que l'interdiction visée au paragraphe 1 ne peut pas être levée par la personne concernée;*
>
> *b) le traitement est nécessaire aux fins de l'exécution des obligations et de l'exercice des droits propres au responsable du traitement ou à la personne concernée en matière de droit du travail, de la sécurité sociale et de la protection sociale, dans la mesure où ce traitement est autorisé par le droit de l'Union, par le droit d'un État membre ou par une convention collective conclue en vertu du droit d'un État membre qui prévoit des garanties appropriées pour les droits fondamentaux et les intérêts de la personne concernée;*
>
> *c) le traitement est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée ou d'une autre personne physique, dans le cas où la personne concernée se trouve dans l'incapacité physique ou juridique de donner son consentement;*
>
> *d) le traitement est effectué, dans le cadre de leurs activités légitimes et moyennant les garanties appropriées, par une fondation, une association ou tout autre organisme à but non lucratif et poursuivant une finalité politique, philosophique, religieuse ou syndicale, à condition que ledit traitement se rapporte exclusivement aux membres ou aux anciens membres dudit organisme ou aux personnes entretenant avec celui-ci des contacts réguliers en liaison avec ses finalités et que les données à caractère personnel ne soient pas communiquées en dehors de cet organisme sans le consentement des personnes concernées;*
>
> *e) le traitement porte sur des données à caractère personnel qui sont manifestement rendues publiques par la personne concernée;*
>
> *f) le traitement est nécessaire à la constatation, à l'exercice ou à la défense d'un droit en justice ou chaque fois que des juridictions agissent dans le cadre de leur fonction juridictionnelle;*

g) le traitement est nécessaire pour des motifs d'intérêt public important, sur la base du droit de l'Union ou du droit d'un 'État membre qui doit être proportionné à l'objectif poursuivi, respecter l'essence du droit à la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée[44]*;*

h) le traitement est nécessaire aux fins de la médecine préventive ou de la médecine du travail, de l'appréciation de la capacité de travail du travailleur, de diagnostics médicaux, de la prise en charge sanitaire ou sociale, ou de la gestion des systèmes et des services de soins de santé ou de protection sociale sur la base du droit de l'Union, du droit d'un État membre ou en vertu d'un contrat conclu avec un professionnel de la santé et soumis aux conditions et garanties visées au paragraphe 3;

i) le traitement est nécessaire pour des motifs d'intérêt public dans le domaine de la santé publique, tels que la protection contre les menaces transfrontalières graves pesant sur la santé, ou aux fins de garantir des normes élevées de qualité et de sécurité des soins de santé et des médicaments ou des dispositifs médicaux, sur la base du droit de l'Union ou du droit de l'État membre qui prévoit des mesures appropriées et spécifiques pour la sauvegarde des droits et libertés de la personne concernée, notamment le secret professionnel;

j) le traitement est nécessaire à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques, conformément à l'article 89, paragraphe 1, sur la base du droit de l'Union ou du droit d'un État membre qui doit être proportionné à l'objectif poursuivi, respecter l'essence du droit à la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée. »

Le RGPD n'impose aucune exigence formelle concernant la communication des informations à la personne concernée. En revanche, un principe de responsabilité s'applique : le responsable du traitement doit pouvoir démontrer qu'il a satisfait à ses obligations et notamment à l'obligation d'information.

Le RGPD précise que les informations relatives au traitement doivent être communiquées d'une façon concise, transparente et compréhensible. Par conséquent, il convient d'utiliser des termes clairs et simples. Le RGPD impose également au responsable du traitement des données de faire particulièrement attention au langage utilisé lorsque celle-ci collecte des données concernant des enfants. Il devra examiner au cas par cas comment mettre en œuvre son obligation d'information afin de satisfaire au mieux au principe de transparence.

Les données à caractère personnel ne peuvent être utilisées, c'est-à-dire traitées, de manière libre puisqu'elles constituent un aspect de la vie privée des personnes physiques que cette réglementation tend à protéger.

Pour répondre au principe de licéité des traitements des données à caractère personnel, une des hypothèses de traitement licite est celle du consentement donné par la personne dont on traite les données à caractère personnel. Le règlement renforce les exigences liées à ce consentement puisqu'il le définit comme étant la manifestation de volonté, libre, spécifique, éclairée et univoque par laquelle la personne concernée accepte, par une déclaration ou par un acte positif clair, que des données à caractère personnel la concernant fassent l'objet d'un traitement. La personne concernée pourra également retirer son consentement.

Une autre nouveauté a fait son apparition : la tenue d'un registre des traitements. Le règlement impose désormais l'obligation pour le responsable du traitement de tenir un registre des activités de traitements. Cette obligation remplace l'obligation de notification préalable des traitements prévue par la loi du 8 décembre 1992.

La désignation d'un délégué à la protection des données est également une des mesures phares du Règlement. Les responsables du traitement qui sont des autorités publiques ou des organismes publics doivent désigner un délégué à la protection des données en fonction de certaines conditions. Il est, entre autres, chargé d'informer et de conseiller le responsable du traitement, de contrôler le respect de la réglementation et de conseiller le responsable du traitement quant à la réalisation d'une analyse d'impact.

Enfin, il y a désormais lieu de notifier toute violation des données à caractère personnel. Les conditions de traitement des données à caractère personnel imposent que les données soient traitées de façon à garantir une sécurité appropriée des données à caractère personnel. Malgré toutes les mesures qui peuvent être prises par le responsable du traitement, nul n'est à l'abri d'une faille de sécurité comme la perte, l'altération ou la divulgation de données. Désormais, avec le Règlement, en cas de violation de données à caractère personnel, le responsable du traitement en notifie la violation en question à l'autorité de contrôle compétente dans les meilleurs délais et, si possible, 72 heures au plus tard après en avoir pris connaissance, à moins que la violation en question ne soit pas susceptible d'engendrer un risque pour les droits et libertés des personnes physiques.

Par ailleurs, le règlement prévoit aussi la notification à la personne concernée de la violation de ses données à caractère personnel lorsque la violation est susceptible d'engendrer un risque élevé pour les droits et libertés d'une personne physique.

En résumé, l'on peut dire que ce nouveau Règlement amène de nombreuses nouveautés, qui s'intègrent logiquement dans la continuité de la réglementation protégeant les données à caractère personnel tout en évoluant sur des points précis.

44 La Belgique a établi une liste exhaustive des traitements entrant dans cette exception prévue à l'article 9.2.g du RGPD, dont certains se retrouvaient déjà dans la loi de 1992.

c) La loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel

La loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel contient essentiellement des dispositions complémentaires qui revêtent un intérêt pour certains services publics ou groupes professionnels. Ainsi, des règles spécifiques sont notamment prévues pour les forces armées, les services de renseignement et de sécurité, les institutions chargées de l'aide aux délinquants sexuels, Child Focus, les journalistes, etc.

Elle abroge la loi du 8 décembre 1992 relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel.

Même si un règlement européen est directement applicable dans le droit de chaque Etat membre, le RGPD laisse la place à des dispositions complémentaires nationales pour préciser certains éléments, comme des conditions d'application, comme nous le verrons *infra*.

Quant à son champ d'application territorial – élément non harmonisé par le RGPD lui-même –, la loi s'applique aux traitements effectués dans le cadre des activités d'un établissement d'un responsable du traitement ou d'un sous-traitant établi sur le territoire belge, que le traitement ait lieu ou non sur le territoire belge. La loi porte également sur les traitements des données à caractère personnel relatives à des personnes concernées qui se trouvent sur le territoire belge, par un responsable du traitement ou un sous-traitant établi hors UE[45].

Comptant plus de 280 articles, elle transpose le RGPD mais également la directive 2016/680 relative aux traitements de données liées à la commission d'infractions pénales et à leurs sanctions.

En exécution de l'article 8.1 du RGPD, le traitement des données à caractère personnel d'un enfant est licite lorsque le consentement a été donné par des enfants âgés de 13 ans ou plus. Cela signifie par exemple qu'un enfant âgé de 13 ans ou plus peut donner son consentement librement pour la création d'un profil sur différents médias sociaux. En comparaison avec ses voisins, la Belgique a opté pour l'âge applicable le plus bas possible.

Le législateur belge a estimé nécessaire de prévoir des mesures de protection supplémentaires en cas de traitements de données génétiques, biométriques ou de santé[46] (article 9), comme, par exemple, le fait que le responsable du traitement doive déterminer les catégories de personnes ayant accès aux données, sur la base d'une description de leur mission en lien avec le traitement de ces données.

L'article 10 traite des cas de levée de l'interdiction de traitements de données relatives à des condamnations et des infractions pénales. Le personnes, physiques ou morales, de droit privé ou de droit public, ont le droit de traiter ce type de données si cela s'avère nécessaire dans le cadre de leur contentieux. Les avocats sont, par exemple, exemptés de cette interdiction dans le cadre de la défense de leurs clients. Le consentement de la personne concernée lèvera cette interdiction ainsi que le fait d'avoir, soi-même, manifestement rendu les données concernées publiques.

Les articles 19 à 23 portent sur les spécificités liées au secteur public. A titre d'exemple, lorsqu'un organisme privé traite des données à caractère personnel pour le compte d'une autorité publique fédérale, il doit désigner un délégué à la protection des données (DPD) en cas de risque élevé dans le cadre du traitement.

Les traitements journalistique, scientifique, artistique ou littéraire n'entrainent pas l'obligation d'information sise à l'article 13.

Dans la même logique, l'article 58, qui traite de l'Autorité de protection des données, réduit considérablement les pouvoirs de cette instance dans le but de préserver la liberté d'expression et le secret des sources.

Au niveau des voies de recours, le législateur a opté pour l'action en cessation comme réaction judiciaire aux violations du RGPD et ce, tant dans le chef de la personne concernée (ou d'une association qui représenterait ses intérêts) que dans celui de l'autorité de contrôle. Le demandeur pourra réclamer, dans le cadre d'une action distincte, la réparation de son dommage au regard de la responsabilité contractuelle ou extracontractuelle[47].

Le législateur belge a par ailleurs affiné les sanctions déjà fixées par la réglementation européenne. Il prévoit ainsi notamment qu'aucune amende administrative ne peut être infligée aux autorités. Ces dernières ne risquent par conséquent que des sanctions administratives non pécuniaires et/ou pénales. En revanche, les entreprises publiques qui offrent des services sur le marché belge risquent encore des amendes administratives.

Il est permis de se demander si les autorités publiques risquent une sanction réelle en cas de violation des dispositions légales en matière de protection des données à caractère personnel. Pour B. Salovic, O. Guerginov et T. Léonard, « les interdictions ou limitations de traitement butteront sur le principe de continuité du service public »[48].

De nouvelles sanctions pénales ont en outre été intégrées à la législation belge.

Par exemple, tant le responsable du traitement que le sous-traitant ou son préposé ou mandataire peuvent encourir une amende de 250 EUR à 15.000 EUR dans les cas spécifiques suivants :

---

*45 B. Salovic, O. Guerguinov et T. Léonard, « La Belgique transpose (enfin) le RGPD (GDPR) », **Droit et Technologies**, 5 sept. 2018, droit-technologie.org/actualites/Belgique-transpose-enfin-rgpd-gdpr.*

46 Ces garanties étaient déjà prévues dans l'arrêté royal du 13 février 2001 en exécution de la loi de 1992.

47 B. Salovic, O. Guerguinov et T. Léonard, « La Belgique transpose (enfin) le RGPD (GDPR) », op. cit.

48 Ibidem.

- Les données à caractère personnel sont traitées sans base juridique, y compris les conditions relatives au consentement et au traitement ultérieur ;

- Les données à caractère personnel sont traitées en violation des conditions générales de traitement (finalité justifiée, caractère adéquat, pertinent et non excessif, etc.) par négligence grave ou avec intention malveillante ;

- Le traitement ayant fait l'objet d'une objection est maintenu sans raison juridique impérieuse ;

- Le transfert de données à caractère personnel à un destinataire dans un pays tiers ou à une organisation internationale est effectué en violation des garanties, conditions ou exceptions prévues dans le RGPD ou la loi belge sur la protection de la vie privée, et ce, par négligence grave ou avec intention malveillante ;

- La mesure correctrice adoptée par l'autorité de contrôle visant la limitation temporaire ou définitive des flux n'est pas respectée ;

- La mesure correctrice adoptée par l'autorité de contrôle qui a été imposée pour mettre le traitement en conformité avec les dispositions du RGPD n'est pas respectée ;

- Il a été fait obstacle aux missions légales de vérification et de contrôle de l'autorité de contrôle compétente, de ses membres ou de ses experts ;

- De la rébellion (au sens de l'article 269 du Code pénal) a été commise à l'encontre des membres de l'autorité de contrôle ;

- La certification est revendiquée ou des sceaux de certification en matière de protection des données sont utilisés publiquement alors que ces certifications, labels ou marques n'ont pas été délivrés par une entité accréditée ou ceux-ci sont utilisés après que la validité de la certification, du sceau ou de la marque a expiré.

En cas de condamnation éventuelle à une infraction telle que décrite ci-dessus, le tribunal peut ordonner l'insertion du jugement, intégralement ou par extraits, dans un ou plusieurs journaux, dans les conditions qu'il détermine, aux frais du condamné.

2. QUELLES SONT LES AUTORITÉS QUI ASSURENT LE RESPECT DES NORMES PROTÉGEANT LA VIE PRIVÉE EN DROIT BELGE ET EN EUROPE ?

AU NIVEAU EUROPÉEN

Il existe une instance européenne dédiée à la protection des données ainsi qu'une série d'instances issues de la coopération entre les Etats européens.

a) Le Contrôleur européen de la protection des données

Le Contrôleur européen de la protection des données (CEPD) a pour principale mission de protéger les personnes dont les données font l'objet d'un traitement par les institutions ou organes communautaires, tels la Commission Européenne, le Conseil et le Parlement européen, mais aussi la Cour de Justice, le Comité des régions ou la Banque européenne d'investissements par exemple.

Le CEPD conseille également ces instances sur les questions de protection des données dans toute une série de domaines d'activités. Cette autorité est établie par le Règlement (CE) 45/2001 qui est en cours de révision afin d'être mis en conformité avec le RGPD.

L'objectif principal est de veiller à évaluer correctement les conséquences des évolutions technologiques. Pour ce faire, le CEPD :

- suit les évolutions technologiques susceptibles d'avoir une incidence sur la protection de la vie privée et des données telles que l'informatique en nuage, les systèmes de gestion des informations personnelles, les données massives ou les logiciels malveillants ;

- analyse et comprend les incidences potentielles de la politique axée sur la technologie et des mesures législatives proposées ;

- fait avancer le débat public en faisant rapport sur les nouvelles technologies liées à la protection de la vie privée et des données ;

- offre un soutien aux institutions de l'UE, aux autorités nationales chargées de la surveillance et de la protection des données ainsi qu'au grand public, en servant de point de référence pour répondre aux questions technologiques liées à la protection de la vie privée et des données ;

- dispose des connaissances et outils technologiques nécessaires pour effectuer des inspections efficaces des systèmes informatiques et autres solutions utilisées pour traiter des données à caractère personnel ;

- fournit des orientations, exerce un pouvoir d'influence, sensibilise et apporte des conseils en ce qui concerne les évolutions technologiques pertinentes du point de vue de la protection de la vie privée et des données et la bonne mise en œuvre des principes de prise en compte du respect de la vie privée dès la conception et de prise en compte du respect de la vie privée par défaut.

b) Le Comité européen de la protection des données (European Data Protection Board – « EDPB »)

Le Comité européen de la protection des données peut adopter des documents d'orientations générales afin de clarifier les dispositions des actes législatifs européens en matière de protection des données et des décisions contraignantes envers les autorités de contrôle nationales afin de garantir une application cohérente de ses dispositions.

Il se compose de représentants des autorités nationales chargées de la protection des données et du Contrôleur européen de la protection des données (CEPD).

c) L'Autorité de contrôle commune Système d'Information Douanier (SID)

Le système d'information douanier («SID») a été créé par un règlement du Conseil du 13 mars 1997. Il s'agit d'un système informatique centralisant les informations douanières permettant de poursuivre et de rechercher plus efficacement les infractions aux réglementations douanière et agricole.

Le SID a pour objectif de permettre aux administrations nationales des douanes de s'échanger et de diffuser des informations sur les activités de contrebande et les demandes d'intervention. Son but est d'aider à rechercher et poursuivre les infractions graves aux lois nationales en renforçant l'efficacité des procédures de coopération et de contrôle des administrations douanières des États membres.

Le SID comporte deux bases de données, l'une relevant du cadre des actions de la Communauté européenne, l'autre, des actions intergouvernementales. La base juridique de la base de données inter-gouvernementale, à savoir la Convention sur l'emploi de l'informatique dans le domaine des douanes (Convention SID) énonce les procédures à suivre pour utiliser les technologies de l'information à des fins douanières. Elle décrit les grands paramètres d'informations qui peuvent être stockées, la manière dont les informations peuvent être amendées, les systèmes de sécurité et contient des dispositions en matière de protection des données.

L'article 18 de la Convention SID crée une autorité de Contrôle Commune des Douanes. Elle est composée de deux représentants de l'autorité ou des autorités de protection des données de chaque Pays membre signataires de la Convention susmentionnée, parmi lesquelles l'Autorité de protection des données belge.

L'autorité de contrôle commune est compétente pour :

- contrôler le fonctionnement du SID ;
- examiner toutes les difficultés d'application ou d'interprétation susceptibles de surgir pendant le fonctionnement du système ;
- étudier les problèmes susceptibles de surgir pendant le fonctionnement du système ;
- étudier les problèmes susceptibles de se présenter lors de l'exercice d'un contrôle indépendant par les autorités de contrôle nationales des Etats membres ou lors de l'exercice des droits d'accès au système dont peuvent se prévaloir les particuliers ;
- définir des propositions visant à trouver des solutions communes à des problèmes qui se présenteraient.

d) Le contrôle des traitements Europol

Une Autorité de contrôle commune (ACC), composée de représentants des autorités de contrôle nationales (dont notre Autorité de protection des données belge), est chargée de surveiller l'activité d'Europol afin de s'assurer que le stockage, le traitement et l'utilisation des données dont disposent les services d'Europol ne portent pas atteinte aux droits des personnes.

Elle contrôle en outre la licéité de la transmission des données qui ont pour origine Europol. L'ACC s'acquitte notamment de cette tâche en effectuant des inspections au sein d'Europol.

L'ACC a pour responsabilité de déterminer si Europol observe les principes de la protection des données dans un certain nombre de domaines spécifiques. Les tâches spécifiques visées comprennent l'examen et le commentaire de l'ouverture de fichiers d'analyse spécifiques, le contrôle des autorisations de transmission de données provenant d'Europol, l'examen des questions concernant la mise en œuvre et l'interprétation de la convention Europol, l'analyse des règles régissant la transmission de données à caractère personnel par Europol à des organes tiers et des Etats non membres, et l'établissement de propositions harmonisées de solutions communes aux problèmes existants.

L'ACC a également pour responsabilité de faire respecter les droits des individus concernant leur information à caractère personnel. Cela comprend la prise en considération des recours des personnes qui ont sollicité l'accès à des informations les concernant mais qui ne sont pas satisfaites de la réaction d'Europol.

e) Le contrôle des traitements SIS

La Convention Schengen a également créé une Autorité de contrôle commune composée des représentants de chaque autorité nationale de contrôle.

Le Système d'information Schengen (SIS) est un système informatique mis en place dans le cadre de la réalisation de l'espace Schengen. Il constitue l'une des mesures dites compensatoires à la libre circulation prévue par l'acquis de Schengen. Ces mesures compensatoires ont été adoptées en vue d'assurer un niveau de protection au moins équivalent à celui existant avant l'abolition des frontières. Le SIS permet aux autorités responsables des États Schengen d'échanger des informations pour le contrôle des personnes et des objets aux frontières ainsi que pour la délivrance de visas et de permis de séjour.

L'Autorité de Contrôle Commune est chargée :

- du contrôle de la fonction de support technique du SIS ;
- de la bonne exécution des dispositions de la Convention ;
- de l'analyse des difficultés d'application ou d'interprétation pouvant survenir lors de l'exploitation du SIS ;
- de l'étude des problèmes pouvant se poser lors du contrôle indépendant effectué par les autorités de contrôle nationales des parties contractantes ou à l'occasion de l'exercice du droit d'accès au système ;
- de l'élaboration de propositions harmonisées en vue de trouver des solutions communes aux problèmes existants.

Au niveau national, chaque État Schengen a désigné une autorité de contrôle chargée d'exercer un contrôle indépendant du fichier de la partie nationale du SIS et de vérifier que le traitement et l'utilisation des données intégrées dans le SIS ne sont pas attentatoires aux droits de la personne concernée.

En Belgique, l'autorité de contrôle est l'Autorité de protection des données. Celle-ci doit donc s'assurer que l'enregistrement de données par les autorités belges dans le SIS est conforme aux dispositions de la Convention Schengen.

f) Le « Global Privacy Enforcement Network » (GPEN)

Le *« Global Privacy Enforcement Network » a été créé afin de promouvoir la collaboration transfrontière entre les autorités chargées de protéger la vie privée.*

En juin 2007, les gouvernements membres de l'OCDE ont adopté une Recommandation relative à la coopération transfrontière dans l'application des législations protégeant la vie privée. Cette Recommandation fait appel aux États membres pour qu'ils constituent un réseau informel des Autorités chargées de protéger la vie privée et mentionne un certain nombre de tâches auxquelles le réseau devrait s'atteler :

- débattre des aspects pratiques de la coopération pour l'application des lois protégeant la vie privée ;
- échanger des pratiques exemplaires face aux problèmes transfrontières ;
- œuvrer à la définition de priorités communes en matière d'application des lois ;
- soutenir des initiatives et campagnes conjointes en matière d'application des lois et de sensibilisation.

Au cours de l'été 2008, les Autorités chargées de protéger la vie privée ont commencé à échanger leurs expériences et à débattre des aspects pratiques de la coopération pour l'application des lois protégeant la vie privée par le biais d'une plate-forme Internet.

g) Eurodac

Eurodac est un système d'information établi pour la comparaison des empreintes digitales des demandeurs d'asile et des immigrants clandestins afin de faciliter l'application de la Convention de Dublin qui permet de déterminer l'État responsable de l'examen d'une demande d'asile.

Le Système Eurodac se compose d'une unité centrale gérée par l'autorité de protection des données européenne, d'une base de données centrale informatisée d'empreintes digitales et de moyens électroniques de transmission entre les États membres et la base de données centrale. Seules les autorités nationales compétentes en matière d'asile ont accès à la banque de données centrale.

Les trois catégories de personnes dont les données à caractère personnel sont collectées par Eurodac sont les suivantes : les demandeurs d'asile âgés de plus de 14 ans, les étrangers appréhendés à l'occasion du franchissement irrégulier d'une frontière extérieure et les étrangers se trouvant illégalement sur le territoire d'un Etat membre.

Les données suivantes sont enregistrées : l'État membre d'origine, les empreintes digitales, le sexe, et un numéro de référence utilisé par l'État membre d'origine. Dans le cas d'un signalement, un échange additionnel de données est effectué à travers le système DubliNet. DubliNet est un réseau sûr de communications électroniques entre les autorités nationales qui traitent les demandes d'asile. Les deux États membres concernés peuvent échanger à travers DubliNet des données à caractère personnel différentes de celles d'Eurodac telles que le nom, la date de naissance, la nationalité, la photo, des détails sur les membres de la famille et dans certains cas des adresses.

Le Contrôleur européen à la protection des données contrôle le traitement des données à caractère personnel dans la banque de données (unité centrale) et leur transmission aux États membres. Les autorités chargées de la protection des données dans les États membres contrôlent les traitements de données effectués par les autorités nationales, ainsi que la transmission de ces données à l'unité centrale.

Afin d'assurer une approche coordonnée, un groupe de coordination de contrôle composé du Contrôleur européen de la protection des données et des autorités nationales de protection des données se réunit à intervalles réguliers afin d'examiner les problèmes communs que pose le fonctionnement du système Eurodac et de recommander des solutions communes.

AU NIVEAU NATIONAL : LA NOUVELLE AUTORITÉ DE PROTECTION DES DONNÉES SUCCÉDANT À LA COMMISSION POUR LA PROTECTION DE LA VIE PRIVÉE

Instituée par la loi du 3 décembre 2017, une nouvelle Autorité de protection des données (APD) exerce, depuis le 25 mai 2018, sa mission d'organe de contrôle indépendant institué auprès de la Chambre des représentants. Comme l'indique la loi, elle « succède » à la défunte Commission de la protection de la vie privée et dispose de la personnalité juridique. C'est une vraie réorganisation de la Commission de la protection de la vie privée et de nouveaux pouvoirs pour cette institution.

Le RGPD a, en effet, prévu de revoir les missions et les pouvoirs de sanctions des autorités de contrôle des Etats membres. L'effectivité des nouvelles règles, contraignantes et protectrices, n'est possible que grâce à l'accroissement des pouvoirs de contrôle et de sanction de l'autorité de contrôle. Aussi, le règlement attribue aux autorités de contrôle, notamment, des pouvoirs d'enquête et le pouvoir de prendre des mesures correctrices.

Elle a pour mission de veiller au respect des principes fondamentaux de la protection des données à caractère personnel, en tant qu'entité fédérale disposant de la personnalité juridique à cet effet.

L'Autorité de protection des données contrôle la manière dont une personne ou une instance se sert des données à caractère personnel des citoyens et les informe de leurs droits et obligations à cet égard. Par ailleurs, elle joue un rôle d'intermédiaire dans le cadre de demandes liées à des traitements de données à caractère personnel, elle traite des réclamations, procède à des contrôles et peut également imposer des sanctions lorsque le RGPD n'est pas respecté.

L'APD est structurée en un « comité de direction », un « centre de connaissances » et une « chambre contentieuse », dont les membres doivent encore être nommés par la Chambre des représentants. Une loi du 24 mai 2018 modifiant la loi du 3 décembre 2017 portant création de l'Autorité de protection des données a prévu que les membres de l'ancienne Commission de la protection de la vie privée exerceraient les tâches et compétences de l'Autorité de protection des données pendant une période transitoire. Le mandat des membres de la Commission de la protection de la vie privée prendra fin le jour où les membres du comité de direction prêteront serment et signeront la déclaration selon laquelle il n'y a pas de conflit d'intérêts.

Le comité de direction de l'APD décidera d'un plan stratégique et d'un plan de gestion, y compris des priorités annuelles de l'institution. Dans ce cadre, le comité de direction demandera l'avis d'un nouveau ***« conseil de réflexion »*** et soumettra le plan stratégique à la consultation publique pendant au moins deux semaines.

Pour correspondre aux exigences du RGPD, l'APD a connu une grande extension de ses pouvoirs d'investigation, de contrôle et de sanction.

La Belgique s'est donc dotée de cette nouvelle institution nationale rendue obligatoire par le RGPD, même si elle n'est pas encore totalement en ordre de marche et n'en est qu'à ses premiers balbutiements.

En effet, comme souligné *infra*, les éléments qui précèdent sont, à ce stade, encore très théoriques et ne connaissent pas encore une grande effectivité. Non seulement, l'APD n'est pas encore opérationnelle pour des raisons de nature institutionnelle et organisationnelle, mais elle doit, en même temps, faire face à une pression due à l'existence de ses nouvelles missions.

Cette difficulté existait déjà à l'époque de la défunte Commission de la vie privée, puisqu'elle ne disposait déjà pas, dans les faits, des moyens pour contrôler efficacement le respect de l'ancienne loi de 1992 ni pour infliger des sanctions en cas de violation de cette législation.

Quant à l'effectivité donnée aux normes de protection de la vie privée en Belgique, il faut souligner également que les juridictions de l'Ordre judiciaire ne leur accordent, à n'en pas douter, assez de crédit. Les absences de poursuite, les classements sans suite ou les acquittements le démontre. Dans les autres cas, les dommages et intérêts accordés en cas de violation sont généralement très symboliques, ne tenant aucun compte des conséquences psychologiques et financières des infractions au droit de la vie privée et du traitement des données à caractère personnel. Gageons que l'esprit du RGPD soufflera sur les cours et tribunaux du pays et que cette situation changera. Mais, sans doute, la justice opposera-t-elle aussi un manque de moyens, à l'instar de l'APD.

Dans la situation actuelle, eu égard au prix de la défense de ces droits, faire appel à la justice pour dénoncer une infraction coûte cher et, comme souligné *supra*, rapporte peu.

En conclusion, même si sa lecture n'est pas aisée tant il est complexe – ce qui n'en facilite pas le respect le plus scrupuleux –, l'on peut néanmoins affirmer que le RGPD a tout pour être efficace en théorie. Mais la sanction de ses violations sera-t-elle effective dans la pratique ? Rien n'est moins sûr.

Que penser du RGPD ?

Pascal Leduc

Les autorités belges assurant le respect de la vie privé sont-elles efficientes et correctement instrumentées en Belgique ?

Le RGPD a considérablement renforcé les missions de l'Autorité de Protection des Données (ex-Commission Vie Privée) ainsi que ses pouvoirs d'investigation, de contrôle et de sanction.

Elle est l'instance de recours en cas de plainte infructueuse d'une personne concernée auprès d'un responsable de traitement, concernant le traitement de ses données.

Elle doit notamment traiter toutes les plaintes contre les GAFA et autres moteurs de recherches concernant ce nouveau droit qu'est *le droit à l'oubli.*

Le RGPD impose également au responsable du traitement de notifier systématiquement à l'APD, toute atteinte sérieuse à la sécurité des données ou toute violation des données (accès non autorisé ou destruction accidentelle ou volontaire des données traitées par le responsable du traitement).

Depuis l'entrée en vigueur du RGPD, l'APD doit faire face à une inflation de plaintes et de notifications, que de son propre aveu, elle est incapable de gérer avec ses équipes actuelles, qui n'ont été que très peu renforcées.

Récemment, le président de l'APD s'est plaint dans la presse que quatre mois après l'entrée en vigueur du RGPD, il avait reçu plus de plaintes sur cette période que pour toute l'année 2016.

Certaines plaintes attendent depuis plus de six mois d'être traitées au sein des services de l'APD, et sans renforcement drastique de ses équipes, les sanctions extrêmement lourdes prévues par le GDPR en cas de violation de ses dispositions essentielles (jusqu'à 4% du chiffre d'affaire du responsable du traitement), resteront longtemps lettre morte en Belgique, à défaut de moyens et de personnel pour les faire appliquer.

Le RGPD pourra-il protéger adéquatement la vie privée - comparaison avec le système anglo-saxon ?

Le RGPD est une règlementation générale qui traite tous les aspects liés à la protection des données à caractère personnel et s'applique, sous réserve de législations sectorielles plus contraignantes, notamment dans le secteur des communications électroniques (directive e-Privacy qui règlementent notamment l'envoi de publicités par courrier électronique).

L'objectif du cadre réglementaire tracé par le RGPD est de protéger les libertés fondamentales de l'individu, notamment son droit à la vie privée, qui en Europe est un droit garanti par et dans la Constitution et ce, à l'inverse de la solution qui prévaut aux Etats-Unis.

Le RGPD offre une souplesse suffisante pour concilier la protection des libertés de l'individu avec les intérêts légitimes des entreprises ou des autres responsables du traitement, qui doivent dans certaines circonstances, pouvoir traiter librement les données des personnes concernées.

La protection offerte par le RGPD bénéficie à tous les citoyens et non pas uniquement à certaines catégories de ceux-ci, comme c'est le cas aux Etats-Unis, qui ne réglementent le traitement des données que de certaines catégories de citoyens telles que les consommateurs, les enfants ou d'autres parties « faibles » dans une relation contractuelle.

Cette conception universelle et égalitaire de la protection de la vie privée, explique pourquoi l'Europe va sans doute réussir à imposer les standards contenus dans le RGPD aux Etats-Unis et au reste du monde civilisé, car dès lors que tous les citoyens européens sont protégés par cette législation, aucune activité de traitement concernant ces citoyens ne peut être effectuée sans respecter le RGPD, même si les activités de traitement sont mises en oeuvre en dehors du territoire de l'Union européenne.

Il n'est pas étonnant dans ces circonstances, que plusieurs grandes sociétés américaines comme Facebook ou d'autres GAFA, ont dans un premier temps isolé leurs activités européennes de leurs activités concernant le marché américain, en les logeant dans une société séparée (par exemple : Facebook Ireland).

Certains responsables de ces groupes ont toutefois pris conscience de l'absurdité de maintenir deux régimes différents et souhaitent calquer la protection des données de leurs clients américains sur celle accordée aux clients européens qui sont protégés par le GDPR.

Si le RGPD offre un cadre complet pour la protection des données à caractère personnel, il n'en reste pas moins que les obligations imposées par ce nouvel instrument aux entreprises se sont considérablement alourdies ; le centre de gravité du dispositif de protection qui auparavant était centralisé au niveau de l'APD, s'est en effet déplacé au niveau du responsable du traitement, que le législateur européen a voulu responsabiliser en lui imposant de restaurer et de maintenir en permanence le contrôle sur ses propres données.

Le renforcement considérable des droits offerts à la personne concernée par un traitement de données à caractère personnel, notamment afin que celle-ci puisse vérifier l'utilisation régulière et proportionnée de ses données à caractère personnel, oblige en quelque sorte le responsable du traitement à rendre en permanence compte de l'usage qui est fait des informations personnelles de ses clients, fournisseurs, et employés, etc., et le contraint, sous peine de sanctions, à prendre les mesures de gestion et de sécurité adéquates afin de satisfaire à ces obligations de *compliance*.

Certaines formalités imposées aux entreprises en vue de satisfaire aux prescriptions du GDPR, constituent un coût disproportionné par rapport aux bénéfices que les personnes concernées peuvent en retirer en termes de protection de leurs données, même si dans la plupart des cas, les coûts de mise en conformité seront non récurrents et vite amortis une fois les procédures internes mise en place et opérationnelles.


Technology
Innovation
SYSTEM
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Innovation
Branding
Solution
Marketing
Analysis
Ideas
Success
Management

Le « Safe Harbor » et le « Privacy Shield » : une difficile conciliation des approches américaine et européenne de la protection des données à caractère personnel

Stéphane Tellier

L'accord *Safe Harbor*, ou la « sphère de sécurité » était, lorsqu'il était encore en vigueur, un ensemble de principes qui permettaient à des entreprises américaines – celles dépendant de l'autorité du département du Commerce des États-Unis, c'est-à-dire les principales entreprises sauf les banques et les compagnies d'assurance – de certifier qu'elles respectaient la législation de l'Espace économique européen (EEE) afin d'obtenir l'autorisation de transférer des données personnelles de l'EEE vers les États-Unis.

La Directive européenne 95/46/CE – abrogée en 2018 par le Règlement général sur la protection des données personnelles –, interdisait déjà, à l'époque de la mise en œuvre du Safe Harbor, le transfert de données personnelles vers des États non membres de l'EEE qui protégeraient les données personnelles à un niveau inférieur à celui de l'EEE.

Les États-Unis et l'EEE partagent l'objectif d'améliorer la protection des données de leurs concitoyens, mais n'abordent pas ce thème de la même manière, loin s'en faut, les Etas européens entendant être beaucoup plus stricts quant à l'enregistrement et au traitement des données à caractère personnel.

Afin de faire le lien entre ces deux approches très différentes de respect de la vie privée et de permettre aux entreprises et aux organisations américaines de se conformer à cette Directive européenne, le département du Commerce des États-Unis, en concertation avec la Commission européenne, avait instauré un cadre juridique dénommé *Safe Harbor*.

PRINCIPES DU SAFE HARBOR

- **Notification** : les individus situés dans l'EEE devaient être informés du fait que leurs données allaient être collectées et de la façon dont ces données allaient être utilisées.
- **Choix** : les individus devaient avoir la possibilité de refuser que les données les concernant soient transférées à des tiers ou utilisées dans un but autre que celui auquel la personne avait consenti précédemment.
- **Transfert à des tiers** : le transfert de données à de tierces parties ne pouvait se faire que si elles garantissaient le même niveau de respect des principes de protection de données personnelles.
- **Sécurité** : l'entreprise devait prendre les mesures nécessaires pour protéger les informations collectées contre la suppression, le mauvais usage, la divulgation ou l'altération de ces données.
- **Intégrité des données** : l'entreprise s'engageait à n'utiliser les données collectées que dans le but pour lequel l'utilisateur avait donné son accord.
- **Accès** : les individus devaient pouvoir accéder aux informations les concernant et les corriger ou les supprimer s'ils le souhaitaient.
- **Application** : l'entreprise devait mettre tout en œuvre pour que ces règles soient effectivement appliquées et devait contrôler leur respect.

CERTIFICATION

Après avoir été certifiée sur la base de cet engagement, l'entreprise devait être à nouveau certifiée tous les douze mois. Elle pouvait contrôler elle-même qu'elle se conformait à ces principes, ou faire appel à un tiers pour effectuer cette évaluation. Ce processus de contrôle exigeait que les employés qui l'effectuaient soient dûment formés et qu'un dispositif permettant de gérer les éventuels litiges soit mis en place.

La *Federal Trade Commission* a été chargée de contrôler ce programme. Les contrôles ont permis de déboucher tant sur des actions en justice que sur des procédures transactionnelles.

L'accord *Safe Harbor* a regroupé jusqu'à près de 4.000 entreprises américaines, dont Microsoft, General Motors, Amazon, Google, Hewlett-Packard et Facebook.

INVALIDATION

Le 6 octobre 2015, la Cour de justice de l'Union européenne invalida l'accord *Safe Harbor*. La Cour a considéré que les États-Unis n'offraient pas un niveau de protection adéquat aux données personnelles transférées, et a rappelé que « la législation permettant aux pouvoirs publics d'accéder de manière généralisée au contenu des communications électroniques doit être considérée comme compromettante pour l'essence même du droit fondamental au respect de la vie privée » et qu'un État membre devait pouvoir vérifier si les transferts de données personnelles entre cet État et les États-Unis respectaient les exigences de la directive européenne sur la protection des données personnelles.

La Commission européenne et le gouvernement des États-Unis ont alors entamé des discussions et ont trouvé, le 2 février 2016, un accord politique, qui conduisit la Commission européenne à publier un projet de Décision, déclarant une équivalence de protection.

Le « G29 », Groupe de travail de l'article 29 sur la protection des données, dans un avis du 13 avril 2016, puis le Contrôleur européen de la protection des données (EDPS), dans un avis du 30 mai 2016, ont toutefois estimé que des points majeurs de préoccupation demeuraient et que des améliorations significatives étaient nécessaires.

Et, le 8 juillet 2016, la version finale du *Privacy Shield* (« Bouclier de protection des données UE – États-Unis ») fut approuvée par la plupart des États membres de l'UE (à l'exception de l'Autriche, la Croatie, la Slovénie et la Bulgarie) et la Commission adopta la décision d'équivalence le 12 juillet 2016.

LE PRIVACY SHIELD

Ce « bouclier » se compose d'une série d'engagements de la part du gouvernement fédéral des États-Unis et d'une décision de la Commission européenne . La Commission a accepté, le 12 juillet 2016, ces dispositions relatives à la protection de la vie privée car elles correspondaient au même niveau de protection des données appliqué en Union européenne.

Cet accord ne constitue pas un traité international, mais se compose d'une série de dispositions, qui réglemente la protection des données personnelles transférées depuis un État membre de l'Union européenne vers les États-Unis. Il était devenu indispensable suite à l'invalidation du *Safe Harbor*.

Le *Privacy Shield* n'est encore, en février 2016, qu'une version d'un « package » de textes – parfois contradictoires – publié conformément à la réglementation. L'accord, lui-même une « décision d'adéquation » de la Commission européenne et d'autres textes, devait être incorporé dans le processus législatif européen.

Le *Privacy Shield* comprend des principes de protection des données à suivre par les entreprises américaines, ainsi que des garanties écrites du gouvernement américain, c'est-à-dire des garanties et des restrictions sur l'accès aux données édictées par les autorités.

Les entreprises américaines concernées doivent s'engager à respecter ces obligations pour être inscrites dans la liste des entreprises certifiées, comme antérieurement avec la liste du *Safe Harbor*.

Les États-Unis ont assuré à la Commission européenne de procéder à des mesures de contrôle efficaces contre les entreprises sous peine de sanctions, allant jusqu'à leur suppression de la liste des entreprises bénéficiaires. La divulgation de données à des tiers est donc désormais liée par des exigences plus strictes.

En outre, la Commission européenne a déclaré que le gouvernement américain s'est engagé formellement vis-à-vis de l'Union européenne, par l'intermédiaire du Bureau du Directeur des services de renseignement, pour que l'accès aux données personnelles des citoyens de l'UE ne soient plus justifiées que par des raisons de sécurité nationale.

Les citoyens de l'UE peuvent saisir un médiateur (*Ombudsperson*), faisant partie du Département d'État, pour enquêter sur les violations et déterminer si une entreprise a agi illégalement.

Les revendications des citoyens de l'UE doivent également être admises, à l'instar de celles des entreprises américaines. Les plaignants doivent assigner l'entreprise dans les 45 jours. Un éventuel litige débouche sur une méthode de résolution des conflits.

Par ailleurs, les citoyens peuvent se tourner vers les autorités nationales de protection des données qui agissent avec la coopération de la commission des plaintes, la *Federal Trade Commission*. Les entreprises de traitement des données personnelles doivent être conformes aux recommandations des autorités nationales de protection des données des États membres de l'UE.

La Commission européenne doit préparer annuellement un rapport sur l'efficacité du *Privacy Shield* et le transmettre au Parlement européen ainsi qu'au Conseil européen. Cet audit, mandaté par la Commission, est réalisé en coopération avec le département du Commerce américain ainsi que des experts des services de renseignement américains et les autorités européennes de protection des données, avec le concours des organisations non-gouvernementales et d'« autres tiers intéressés », tous conviés lors d'un congrès sur la vie privée.

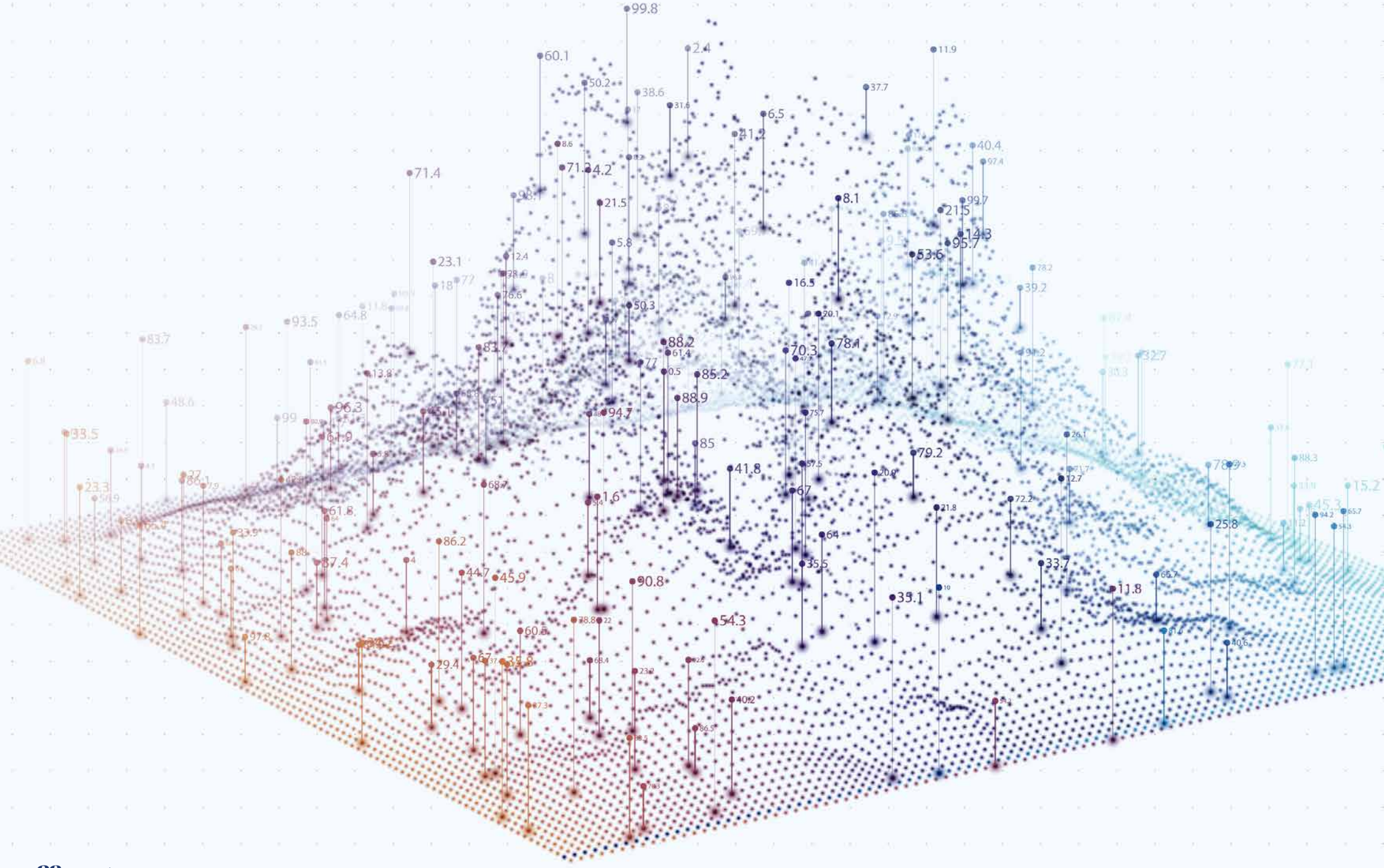
L'accord — la presse américaine a parlé d'un « deal » — a fait l'objet, dès le début, de critiques. Maximilian Schrems, l'avocat autrichien pourfendeur du Safe Harbor, a avancé ses arguments pour invalider le *Privacy Shield* à l'occasion du *Cloud Independance Day*, juste avant sa ratification. Il a souligné que le *Privacy Shield* pose des problèmes à la fois dans le domaine commercial et sur le plan des politiques publiques. Il dénonce l'utilisation à très grande échelle des données personnelles sur le plan commercial et souligne que les sociétés européennes continuent à ne pas disposer des mêmes droits que les sociétés américaines.

Maximilian Schrems souligne que le problème existe toujours au niveau de la surveillance de masse, le texte ne recelant aucun changement fondamental par rapport au *Safe Harbor* et rappelle les six exceptions : la détection et la lutte contre certaines activités de puissances étrangères, l'anti-terrorisme, la lutte contre la prolifération nucléaire, la cyber sécurité, la détection et la lutte contre les menaces visant les États-Unis et les forces armées alliées et, enfin, la lutte contre les menaces de crimes transnationaux, alors que cela serait, selon lui, contraire à la jurisprudence de la Cour de justice de l'Union européenne.

Il doute également de l'impartialité du médiateur désigné et encadré par le ministère des Affaires étrangères américain.

Le *Privacy Shield* est déjà remis en question, preuve que ces deux approches de la protection des données personnelles – l'approche européenne et l'approche américaine – ne sont pas facilement conciliables, à l'heure pourtant de l'utilisation de masse de Facebook, Google et d'autres opérateurs américains par les consommateurs européens, impliquant le transit, l'enregistrement et le traitement de leurs données personnelles sur des serveurs extra-européens.

Sans doute le grand enjeu sociétal et juridique du traitement des données à caractère personnel mériterait-il la conclusion d'un véritable traité international et non un si fragile accord entre deux visions et deux contextes trop différents.



IV. BIG DATA & ALGORITHMES

Corentin de Salle

QU'APPELLE-T-ON « BIG DATA » ?

On définit souvent les big data au moyen d'une plaisanterie : les big data seraient « tout ce qui est trop gros pour tenir dans un seul ordinateur ».[49]

Selon Rob Kitchin,[50] les data procèdent « en abstrayant le monde en catégories, mesures et autres formes représentatives - des nombres, des caractères, des symboles, des images, des sons, des ondes électromagnétiques, des octets - qui constituent les briques avec lesquelles on crée l'information et la connaissance ».

Selon le professeur Alain Strowel[51] qui explicite la définition de Kitchin, **les data ne sont pas neutres, objectives ou pré-analytiques.** Elles sont toujours configurées par des instruments, des pratiques, des contextes conçus pour les générer, les sélectionner, les représenter et les analyser (ainsi un « degré » est soit Celsius soit Fahrenheit). Dès lors, écrit Alain Strowel, se baser sur des données comme s'il s'agissait d'un matériau brut est une illusion qui conduit nécessairement à des biais cognitifs et à des discriminations car les manières par lesquelles les data ont été extraites, sélectionnées et représentées sont toujours le résultat de complexes systèmes de mesure et de valeurs élaborés par des hommes.

Si ce terme est devenu tellement utilisé aujourd'hui, ce n'est pas parce que la récolte des données serait un phénomène neuf. Bien au contraire, le moissonnage des données est aussi vieux que l'humanité. Les premiers documents écrits de l'histoire servaient en réalité à enregistrer des données utiles à l'administration (recensement, récoltes, têtes de bétail, etc.). Mais ce qui est neuf, c'est la quantité colossale de données qu'il est désormais possible d'enregistrer, de stocker et surtout la puissance de calcul avec laquelle on peut désormais les traiter et les exploiter.

Cela dit, le caractère colossal ou gigantesque des données n'est pas un caractère très indicatif des big data. En effet, depuis le XVIII^{ème} siècle, on ne cessait de s'exclamer, à chaque fois qu'une nouvelle technique de statistique apparaissait, sur le caractère quantitativement phénoménal des données qu'elle permettait de récolter et de traiter. Et ce fut chaque fois vrai. La grandeur est un caractère relatif. Elle ne nous apprend pas grand-chose. Ce qui, par contre, est très caractéristique de ces données, c'est qu'elles sont numérisées. Il ne faut pas confondre ici « numérisé » et « digitalisé ». La digitalisation est juste le processus par lequel une chose est scannée. La numérisation implique que ce scan soit passé par un logiciel de reconnaissance et que toutes les données soient dès lors exploitables.

Un exemple de big data, ce seront toutes des données nécessaires à une politique de l'urbanisme (pour mettre en œuvre, par exemple, un smart city). Cela peut être également des données permettant de détecter la préparation d'un attentat terroriste. Ou alors, des habitudes de consommation énergétique dans tel pays, telle région, telle ville, tel quartier, etc.

Nous générons chaque jour quantité de données. Une profusion. Elles prolifèrent. On laisse constamment des traces. Il y a quelque chose d'organique dans cette production ininterrompue. Cela ressemble à un processus naturel. Il y a quelques années, des chercheurs de Berkeley[52] estimaient que l'humanité avait accumulé environ 12 exabytes de données durant toute son histoire jusqu'à la généralisation des ordinateurs. Un exabyte correspond à 10^{18} bytes ou, plus prosaïquement à une vidéo format DVD dont le visionnage sans interruption prendrait 50.000 ans. Mais le chiffre aurait grimpé à 180 exabytes en 2006. Selon une étude plus récente,[53] le nombre serait passé à 1600 exabytes entre 2006 et 2011, franchissant ainsi la barrière du zetabyte (c'est-à-dire 1000 exabytes).

Et cela est destiné à s'accroître de façon phénoménale. Surtout lorsque nous aurons pleinement pénétré dans l'univers des objets connectés. Cela dit, **contrairement à ce qu'on croit parfois, tout n'est pas conservé sur internet.** Une quantité gigantesque de données est détruite à tout moment. Quantité de pages sont mises à jour et modifiées. « Sauver un document » signifie « remplacer l'ancienne version ». Même le premier site web de l'histoire a été détruit : en 1993, le Centre Européen de Recherche Nucléaire (CERN) annonçait que le World Wide Web qu'il avait créé serait dorénavant libre pour chacun. Vingt ans plus tard, le CERN organisa un évènement pour fêter l'anniversaire. Une équipe du CERN fut mobilisée pour recréer le toute première page web de l'histoire (avec son URL originelle, etc.) car elle n'existait plus depuis longtemps…

49 On dit souvent qu'on peut parler de "big data" une fois qu'on est dans le « tera », c'est-à-dire 10^{12}, soit 1000 milliards

50 R. Kitchin, ***The Data Revolution. Big Data, Open Data, Data Infrastructure & Their Consequences****, Sage, 2014*

51 Strowel Alain, ***Big Data and Data Appropriation in the EU****, in T. Aplin (ed.), Research Handbook on Intellectual Property and Digital Technologies, Edward Elgar, 2018, p.3*

52 P. Lyman & H.R. Varian, ***How much information****, 2003*

53 J. Gantz & D. Reinsel, ***Extracting value from chaos****, 2011*

La vitesse avec laquelle nos capacités de stockage augmentent est encore plus rapide que la prétendue loi de Moore.[54] Pourtant, cette vitesse est encore inférieure à la vitesse avec laquelle nous générons de nouvelles données. Il faut constamment effacer d'anciennes données pour en créer de nouvelles. Le nouveau chasse l'ancien. Heureusement, quantité de nos données ne sont pas très intéressantes. Souvent, nous dit Luciano Floridi, nous prenons 10 photos dans l'espoir qu'une d'entre elles sera bonne.[55]

D'ailleurs, une donnée n'a généralement quasiment aucun intérêt et aucune valeur économique. De manière individuelle, elle constate ou enregistre une situation (par exemple, un piéton passe dans telle rue à telle heure) et est très pauvre en informations. C'est lorsqu'elles sont prises en masse et mises en relation qu'elles acquièrent intérêt et valeur. C'est dans cette hypothèse qu'on peut établir entre elles des corrélations prédictives qui, elles, ont une valeur.

C'est **la** « **loi** » **de Metcalfe** qui, comme la prétendue loi de Moore, est plutôt une généralisation empirique qu'une loi. Elle énonce que la valeur d'un réseau est proportionnelle au carré du nombre des nœuds connectés du système. Ainsi, un réseau constitué de 2 ordinateurs a seulement une valeur de $2^2=4$ mais doubler le nombre d'ordinateur signifie quadrupler la valeur du réseau car $4^2=16$. Plus il y des nœuds, plus il devient utile d'être connecté et plus il devient onéreux de ne pas être connecté. Après 20 itérations, on peut parler d'une croissance qui ressemble pratiquement à une ligne verticale.

On parle souvent des « 3 V » du Big Data :

- grand **« volume »** de data ;
- grande **« variété »** des data ;
- grande **« vélocité »** de la génération et de la transmission des data.

On pourrait aussi parler de la grande « valeur » des data mais ce n'est pas nécessairement vrai pour toutes les catégories de data. Pour qu'elles acquièrent de la valeur, il faut, précise le professeur Alain Strowel, que ceux qui désirent les exploiter parviennent à identifier, dans le déluge des data, un certain nombre de modèles, de régularités et de corrélations. Le but étant ici de pouvoir en tirer des résultats et des prédictions. La valeur des data ne réside donc pas, selon Alain Strowel, en elles-mêmes mais dans l'utilisation qu'on en fait et dans le parfait timing de leur utilisation. On verra par ailleurs, dans le chapitre de Jérôme de Cooman et de Nicolas Petit que la valeur de ces données doit être relativisée.

Par quoi ces big data sont-elles traitées et analysées ? Par des algorithmes. Il est temps d'examiner cette notion fondamentale.

QU'EST-CE QU'UN ALGORITHME ?

Il s'agit, au sens large du terme, d'une **série limitée d'instructions et d'opérations permettant de résoudre un problème ou d'obtenir un résultat**.

Une recette de cuisine est un algorithme. Mais, ce terme est évidemment utilisé ici pour désigner des outils informatiques accomplissant en très peu de temps des opérations complexes portant sur une masse très importante de données. Ce sont des opérations qui, autrefois, étaient effectuées (et le sont encore) par des cerveaux humains. L'algorithme est conçu de manière très minutieuse : le programmeur décrit dans ses moindres détails comment procéder pour faire quelque chose.

Ce qui est important à comprendre ici, c'est que, de la même façon que les données ne sont pas neutres et objectives, les algorithmes sont conçus par des humains et qu'ils obéissent à des intentions. Ils donnent plus d'importance à certaines données que d'autres. **Ce qu'ils parviennent à découvrir et à mettre en valeur, ce n'est pas une description de la réalité. Mais la mise en valeur d'un aspect de celle-ci. Et, en particulier, la mise en valeur d'une opportunité** pour prendre une décision. Par exemple, pour réaliser un profit. Ou résoudre un problème sociétal.

L'algorithme met en avant un aspect du réel. Il n'épuise pas le réel. Nous savons depuis Aristote que ce dernier est toujours plus riche que ce qu'on peut dire sur lui.

Idéalement, avant de se prononcer sur le résultat qui apparaît suite à une opération algorithmique, il serait intéressant de se renseigner sur les intentions que poursuit cet algorithme. Mais la chose est rarement possible car un algorithme est généralement opaque. Et leur propriétaires (généralement des compagnies commerciales) préfèrent garder secrets les algorithmes qu'ils utilisent.

*54 On parle plutôt des **lois de Moore** car il y en deux. En 1965, Gordon Moore, l'un des fondateurs de la société Intel, a observé que la complexité des semi-conducteurs doublait tous les deux ans à coût constant depuis 1959, date de leur invention. En 1975, il réévalua sa prédiction en affirmant que le nombre de transistors des microprocesseurs sur une puce de silicium doublait tous les deux ans. Toujours à coût constant. L'idée de base ici est bien que les machines électroniques augmentent leur puissance tout en devenant de moins en moins coûteuses. Par la suite, cette loi a été étendue par d'autres auteurs à quantité d'autres choses : la puissance, la capacité, la vitesse, les capacités de stockage qui doubleraient tous les 18 mois. Depuis lors, cette « loi » (on parle plutôt ici encore d'observation empirique à ce stade) se serait vérifiée. Il s'agirait d'un phénomène de développement exponentiel. Un tel phénomène se caractérise par le fait que, au début, « on ne voit rien venir » jusqu'à ce que se produise une explosion passé un certain stade. Pour une analyse critique de la loi de Moore, confer C. de Salle, **Vers un transhumanisme libéral** ?, Centre Jean Gol, 2018*

*55 L. Floridi, **The fourth revolution. How the infosphere is reshaping human reality**, Oxford University Press, 2014, p.21*

Ces algorithmes vont, dès lors, exercer une influence importante sur la réalité. C'est ce que résume joliment Dominique Cardon quand il écrit : « **Nous fabriquons ces calculateurs, mais en retour ils nous construisent ».**[56]

En effet, ces calculateurs nous influencent car ils « fabriquent » la réalité que nous habitons. Comment ? En nous donnant une grille de lecture du réel. Et, comme toute bonne grille de lecture, cette dernière passe inaperçue aux yeux de la majorité des gens qui estiment simplement que les mesures et résultats qui sont générés par ces algorithmes constituent le réel lui-même. En réalité, les algorithmes organisent et orientent la réalité. Ils sélectionnent, pondèrent et valorisent des éléments du réel qu'ils hiérarchisent entre eux.

Dominique Cardon illustre les partis pris inhérents à toute méthode de calcul en examinant quatre façons de s'y prendre pour mesurer le succès d'acteurs sur internet. Ces quatre modes - ou familles - de calcul sont quatre méthodes qui ont été utilisées successivement dans l'histoire d'internet. Ils témoignent de l'évolution des algorithmes.

QUATRE MANIÈRES DE CLASSER L'INFORMATION NUMÉRIQUE SELON DOMINIQUE CARDON[57]

En résumé, il y a eu quatre manières de procéder pour mesurer le succès des acteurs sur internet :

On a d'abord mesuré la **popularité** des sites et, pour ce faire, on comptabilisait le nombre de « vues », en l'occurrence de « visiteurs uniques ». C'était utile quand il n'y avait encore qu'un petit nombre de sites. Quand les producteurs sont devenus plus nombreux, il a fallu trouver un autre système de mesures car la « quantité » de vues ne garantissait en rien la qualité de ces sites qui pouvaient être conformistes, consensuels et populaires.

	A CÔTÉ	AU-DESSUS	DANS	AU-DESSOUS
Exemples	Médiamétrie, Google Analytics, affichage publicitaire	PageRank de Google, Digg, Wikipédia	Nombre d'amis Facebook, Retweet de Twitter, notes et avis	Recommandation Amazon, publicité comportementale
Donnés	Vues	Liens	Likes	Traces
Population	Échantillon représentatif	Vote censitaire, communautés	Réseau social, affinitaire, déclaratif	Comportements individuels implicites
Forme de calcul	Vote	Classements méritocratiques	Benchmark	Machine Learning
Principe	Popularité	Autorité	Réputation	Prédiction

*Source: D. Cardon, **op.cit.**, p.18*

- On a alors mesuré **l'autorité** des sites, leur « force sociale » et cela par une mesure typique au monde des revues scientifiques et universitaires : un article fait d'autant plus autorité qu'il est cité par quantité d'autres articles. En l'occurrence, la manière de mesurer l'autorité d'un site, c'est en comptant le nombre de liens hypertextes qui renvoient vers ce site. Google a ainsi créé l'outil qui a fait sa fortune : le PageRank, lequel mesure l'autorité d'un site en fonction du nombre de sites qui s'y réfèrent. Le problème, c'est que cela a amené tout une série de spécialistes du référencement à fabriquer des sites qui parlent de leurs clients, des « fermes de faux sites » avec de faux contenus éditoriaux pour tromper l'algorithme.

- Dès lors, on a mesuré la **réputation** qui, elle, s'obtient de manière exclusivement démocratique, c'est-à-dire sous forme de « like ». Facebook est le paradigme de ce modèle d'évaluation. On a vu fleurir toute une série de systèmes de notations et avis. Le problème, c'est que ces systèmes sont accusés d'enfermer leurs utilisateurs dans des bulles, des micro-communautés de gens qui se ressemblent et qui ne rendent plus compte de la totalité de l'offre sur internet. Par ailleurs, ces métriques sont hétérogènes et peuvent difficilement être agrégées. De plus, on reste prisonnier de ce que prétendent les usagers du web : ils disent « aimer » ceci mais, en réalité, ils ne le pensent peut-être pas et préfèrent d'autres choses : par exemple, quantité de gens disent « aimer » Arte mais passent en réalité tout leur temps devant TF1. Il fallait détecter ce que les gens aiment vraiment.

- C'est la raison pour laquelle, on a mis en place cette métrique aujourd'hui dominante : la **prédiction** : l'algorithme observe l'internaute, le flique, le suit à la trace. Il observe ce qu'il lit, ce qu'il demande à Google, le temps qu'il passe sur tel ou tel site, les choses dont il parle sur les réseaux sociaux, ce qu'il commande, ce qu'il recommande, les publicités qui l'attirent, etc. Il dresse alors un profil de plus en plus spécifique et il peut alors comparer, parmi les millions d'autres profils qu'il construit et gère simultanément, ce profil à d'autres profils similaires. De cette façon, il peut alors prédire ce que l'internaute va faire. Comme nous l'avions déjà dit dans le chapitre sur la vie privée à l'ère numérique, Dominique Cardon explique que « le futur de l'internaute est prédit par le passé de ceux qui lui ressemblent ».[58]

*56 D. Cardon., **A quoi rêvent les algorithmes. Nos vies** à **l'heure des big data**, Seuil, 2015, p.7*
*57 D. Cardon., **A quoi rêvent les algorithmes. Nos vies** à **l'heure des big data**, Seuil, 2015, p.18*
*58 D. Cardon., **A quoi rêvent les algorithmes. Nos vies** à **l'heure des big data**, Seuil, 2015, p.34*

Glossaire

Algorithme : série limitée d'instructions et d'opérations permettant de résoudre un problème ou d'obtenir un résultat.

Anonymisation des données : procédé par lequel on détruit toute possibilité de pouvoir identifier à quel individu appartiennent telles ou telles données personnelles. Ce processus consiste à modifier le contenu ou la structure des données en question afin de rendre (quasi) impossible la « réidentification » des personnes, même après traitement. Et ce contrairement aux données pseudonymisées où cette réidentification est possible. Cela dit, beaucoup de spécialistes considèrent qu'une anonymisation absolue est impossible. On anonymise des données afin de protéger les individus et leurs données personnelles, généralement quand on diffuse ou partage des données jugées d'intérêt public comme les données ouvertes (open data).

Base de données : stock de données à propos d'un thème ou d'une activité et permettant de délivrer un ensemble sinon exhaustif du moins important d'informations sur ce thème ou cette activité.

Big Data (données massives ou mégadonnées) : ensemble volumineux et diversifié de données numérisées, nées des progrès considérables des capacités informatiques de stockage et de vitesse de traitement et qui, adéquatement traitées, permettent de révéler des corrélations inattendues, ont une valeur prédictive et offrent des perspectives extraordinaires dans quantité de domaines tels que la prospective, la gestion du risque, la lutte contre la criminalité, la gestion des réseaux électrique, la météorologie, etc.

Data mining : ensemble de procédés automatiques ou semi-automatiques mobilisés pour traiter, fouiller, explorer des données massives afin d'en extraire des informations ou des connaissances utiles. De manière générale, ce terme désigne toute recherche de corrélations sans hypothèses de départ.

Data scientists (spécialistes des données) : profession en pleine expansion de personnes qui rassemblent, agrègent, nettoyent, structurent des données massives et les analysent pour en produire de la valeur et répondre à une problématique spécifique.

Métadonnées : données, généralement produites automatiquement par la machine et qui sont des données à propos des données. Ce sont des données qui ne concernent pas directement le contenu ou la substance. C'est-à-dire des données qui en décrivent d'autres. Par exemple, la durée d'une conversation téléphonique, la liste des adresse IP des ordinateurs qu'un ordinateur à contacté, etc.

Open data (données ouvertes) : « toutes les données dont l'accès, la réutilisation ou la distribution est libre, sujettes tout au plus à une demande d'attribution ».[59]

Pseudonymisation des données : processus au terme duquel les données personnelles « ne peuvent plus être attribuées à un sujet spécifique sans l'utilisation d'informations additionnelles, pourvu que ces informations additionnelles soient conservées séparément et soit soumises à des mesures techniques et organisationnelles qui assurent que ces données personnelles ne soient pas attribuées à une personne identifiée ou identifiable » (article 4(5) du RGDP). Elle diffère de l'anonymisation car les données restent liées à la même personne dans tous les dossiers et systèmes informatiques sans que l'identité ne soit révélée.

Nous en sommes à ce stade-là. Les algorithmes détectent automatiquement des corrélations. Et ces dernières reposent sur un nombre tellement considérable de données qu'elles sont jugées objectives et incontestables. Mais, si l'on commence à croire cela, c'est que nous confondons l'instrument de mesure et le phénomène qu'il mesure. Or, il y a des réalités qui ne se manifestent pas, qui ne s'extériorisent pas et qui, dès lors, restent en dehors des radars. Ce glissement se fait insensiblement. Ainsi, les plaintes des femmes battues dans un pays deviennent le nombre de femmes battues dans ce pays. Mais, ce qui intéresse les propriétaires des algorithmes, ce n'est pas de comprendre le monde. Ce qui les intéresse, c'est d'exploiter les opportunités qu'il offre. Nous ne sommes pas dans le registre de la vérité mais dans celui de l'efficacité.

L'ABANDON DE LA CAUSALITÉ AU PROFIT DE LA CORRÉLATION

Une idée centrale, et provocante, à propos de big data, est que les données massives nous permettraient de prendre des décisions optimales en se passant de théorie. C'est la thèse de Chris Anderson développée dans un article célèbre publié en 2008 : « The End of Theory ».[60] Selon cet auteur, nous avons toujours dû recourir à des modèles et ces derniers s'avèrent toujours faux en dernière analyse. Aujourd'hui, heureusement, nous pouvons, si nous désirons agir, nous passer de tous ces modèles et nous contenter des corrélations :

> *« Oubliez la taxinomie, l'ontologie et la psychologie. Qui sait pourquoi les gens font ce qu'ils font ? Le fait est qu'ils le font et on peut l'enregistrer avec une fidélité sans précédent. Avec assez de données, les chiffres parlent d'eux-mêmes ».*[61]

59 Strowel Alain, ***Big Data and Data Appropriation in the EU****, in T. Aplin (ed.), Research Handbook on Intellectual Property and Digital Technologies, Edward Elgar, 2018, p.11*
60 Ch. Anderson, The End of Theory : the data deluge makes the scientific method obsolete, ***Wired Magazine****, 2008 www.wired.com/2008/06/pb-theory/*
61 Ch. Anderson, op.cit., 2008

Les « data scientists » essayent de découvrir des corrélations en faisant le moins d'hypothèses possibles. Ainsi, les algorithmes ont découvert une corrélation entre le fait d'acheter des feutres à placer en dessous des chaises pour éviter de rayer le parquet et le fait de conduire prudemment. Pour les assureurs automobiles, c'est une information importante pour savoir à qui, de préférence, accorder des réductions de primes.[62]

Cette pratique consacre la prééminence de la catégorie de la « corrélation » sur celle, plus traditionnelle (et plus scientifique), de la « causalité ». En résumé, selon l'idéologie à la base de ces pratiques, peu importe la cause d'un comportement, l'important est qu'on puisse établir une corrélation entre ce comportement et un ou plusieurs autres et qu'on puisse, dès lors, le prédire. Nous n'avons plus besoin, disent les idéologues des big data, d'expliquer les choses.

Ce qui intéresse évidemment considérablement les entreprises (mais on peut dire la même chose des gouvernants), c'est la capacité de prédire les désirs et les comportements des gens. A la rigueur, diront les entreprises, nous connaissons vos désirs mieux que vous. Ou plutôt vos pulsions.

Evidemment, procéder de cette manière offre d'indubitables avantages : on ne doit plus commander, payer et attendre la réalisation de longues études pour avoir une base sur laquelle prendre des décisions. Les algorithmes permettent de faire des découvertes totalement inattendues car ils travaillent sur un champ de données considérablement étendu dans le temps et dans l'espace.

Mais le prix à payer, c'est qu'on renonce à comprendre. Et donc, on risque de manquer de prudence en agissant. On risque de commettre des erreurs qui auront un impact important sur des gens. On risque aussi, même avec des algorithmes très précis et très fins, de faire preuve d'injustice.

Ainsi, imaginons un algorithme qui montre que, en moyenne, les gens qui habitent dans tel quartier et font leurs courses dans tel ou tel magasin sont des mauvais payeurs, la banque risque de refuser un prêt à une personne qui correspond à ce profil mais qui a pourtant toujours scrupuleusement honoré ses dettes en temps et en heure. Alors qu'une procédure plus classique consistant à examiner l'historique de ses payements lui aurait permis de réaliser cet emprunt.

BIAIS DE CONFIRMATION & ÉGALITÉ D'ACCÈS À L'INFORMATION

Un danger, accentué par ces algorithmes, c'est le biais de confirmation, c'est-à-dire la propension, délibérée ou non, à ne se focaliser que sur les sources d'informations qui confirment et renforcent notre propre vision du monde. Les individus qui se ressemblent s'enfermeraient dans une micro-communauté qui deviendrait une bulle. Une étude menée par l'Université d'Etat de l'Ohio conclut heureusement que cet effet est plus réduit qu'on ne le croit, du moins en ce qui concerne les Etats Unis.[63] Mais la professeur Antoinette Rouvroy y voit un danger pesant sur le débat démocratique.

Ces algorithmes menacent aussi, selon nous, l'égalité d'accès à l'information. N'y a-t-il pas un véritable problème ici en raison de la spécification des recherches en fonction du profil de l'utilisateur individuel ? Aujourd'hui deux personnes différentes n'accèdent plus aux mêmes informations en formulant des requêtes identiques sur un même moteur de recherche. En effet, les résultats vont dépendre de notre identité numérique, c'est-à-dire de tout notre passé sur le Net et des préférences que ce dernier révèle.

Il importe, pensons-nous, de veiller à ce que ces algorithmes ne détruisent pas notre « monde commun ».

Il existe aussi un autre danger mais qui provient cette fois non des GAFAM mais des Etats : la balkanisation d'internet.[64] Eric Schmidt et Jared Cohen mettent en garde contre le danger de censure d'internet, déjà bien présent en Chine, par certains Etats. Ce procédé est souvent appelé pudiquement « filtrage ». Même les pays démocratiques le pratiquent, parfois pour de bonnes raisons (ils bloquent, en permanence, certains sites comme ceux présentant de la pornographie infantile). Dans certains pays, il n'y a qu'un seul point d'entrée, un seul fournisseur d'accès à internet (FAI) et ce blocage est plus facile. On pourrait voir le Web se fissurer, se fragmenter en une multitude d'internet : un internet arabe, un internet russe, un internet américain, etc. L'internet chinois est le plus actif « filtreur » du monde. Facebook, Twitter, etc. sont bloqués par ce gouvernement sur son territoire. Aucune recherche n'est possible sur les manifestations de la place Tian'enmen, sur les droits de l'homme, sur le dalaï lama, sur le mouvement des droits tibétains, etc. Le pire, c'est que, pour l'usager chinois ordinaire, cette censure est invisible. Par ailleurs, le gouvernement chinois, selon une étude de Rebecca MacKinnon parue en 2010, rétribuerait près de 300.000 commentateurs anonymes pour publier des statuts qui s'extasient sur la sagesse et l'intelligence du gouvernement chinois.

---

*62 D. Cardon., **A quoi rêvent les algorithmes. Nos vies à l'heure des big data**, Seuil, 2015, p.52*

*63 E. Schmidt & J. Cohen, **A nous d'écrire l'avenir. Comment les nouvelles technologies bouleversent le monde**, Denoël, 2013, p.55*

*64 E. Schmidt & J. Cohen, **A nous d'écrire l'avenir. Comment les nouvelles technologies bouleversent le monde**, Denoël, 2013, p.125*

V. RECUL CRITIQUE SUR LA GOUVERNEMENTALITÉ ALGORITHMIQUE

Corentin de Salle

SERONS-NOUS UN JOUR, SOMMES-NOS DÉJÀ, GOUVERNÉS PAR LES ALGORITHMES ?

Pour aborder cette question, faisons un détour par les travaux intellectuellement très stimulants du professeur **Antoinette Rouvroy**. Nous reproduirons d'abord ici quelques-unes de ses idées principales telles qu'elles sont développées dans plusieurs de ses articles et conférences. Par la suite, nous essayerons d'adopter une perspective critique.

QU'EST-CE QUE LA GOUVERNEMENTALITÉ ?

Ce terme ne désigne pas spécifiquement les institutions et les hommes investis symboliquement d'un pouvoir. C'est beaucoup plus large. **Selon Michel Foucault, tout qui parvient à structurer par avance le champ d'action d'autrui gouverne**. Ainsi, la publicité gouverne, l'architecture d'une pièce gouverne, etc.

LES ALGORITHMES GOUVERNENT-ILS ?

PREMIÈRE IDÉE : « LES BIG DATA NE NOUS OFFRENT PAS UN ACCÈS DIRECT AU MONDE »

Les big data produites en un temps t représentent une masse de données tellement colossale que l'esprit humain est devenu incapable de se les représenter.[65] Comme seuls les ordinateurs sont à même de les appréhender et de leur donner un sens, on prétend parfois que, de toute façon, on n'a plus besoin de passer par une représentation. Nous pourrions, comme le prétend Chris Anderson, nous passer de tout modèle et nous référer directement à ce que les algorithmes nous en disent. Plus besoin, disent certains, de passer encore par des institutions pour comprendre, par exemple, le monde social. Plus besoin de passer par le droit, par des conventions, par des témoignages, par des aveux, etc. pour comprendre et interpréter le réel. Non. Les faits « parlent d'eux-mêmes ». Plus besoin de passer par le langage lui-même. Le monde parle de lui-même via des chiffres (des séries de 0 et de 1).

Or, cela est faux dit Antoinette Rouvroy : nous n'avons accès au monde qu'à travers la représentation que nous nous en faisons, c'est-à-dire uniquement à travers nos biais. Il existe des biais, des préjugés, des grilles de lectures. Ces biais sont identifiables et ne sont pas nécessairement négatifs. De toute façon, on ne peut pas s'en passer. Cette idée qu'on pourrait avoir accès au monde brut, au monde tel qu'il est, est une idée qui est complètement fausse.

On prétend enregistrer « le monde tel qu'il est ». En réalité, si le monde est tel qu'il est décrit, il ne l'est pas naturellement. Le monde est, nous dit-elle, structuré par des rapports de pouvoir, de domination, de force. Les données transcrivent cet état de fait sans expliciter les conditions qui l'ont engendré. Si ces données semblent brutes, c'est parce qu'elles sont décontextualisées, « désindexées » nous dit Antoinette Rouvroy. On découpe. On segmente. On rend les données « amnésiques de leur source, amnésiques de leur condition de production ».[66] Pourquoi ? Parce que les entreprises qui récoltent ces données sont obligés de les anonymiser pour ne pas tomber sous le coup des législations de protection des données à caractère personnel. On donne alors l'illusion qu'il s'agirait de faits neutres. On ne peut plus les rattacher à un référentiel originaire à l'aune duquel on pourrait évaluer leur validité. On « naturalise » l'état de fait.

SECONDE IDÉE : « LES ALGORITHMES SONT INTÉRESSANTS CAR ILS SONT DOTÉS D'UNE CURIOSITÉ AUTOMATIQUE MAIS ILS SONT DÉNUÉS D'OBJECTIVITÉ »

Cela s'explique par le fait qu'ils n'ont pas de corps. Nous, nous partons toujours d'un point de vue qui détermine ce qui nous est perceptible intelligible et interprétable. L'algorithme, par contre, a un mode d'existence tout autre. Il peut identifier de façon simultanée, mettre en corrélation et interpréter des points de données qui sont très éloignées dans l'espace.

65 Le Mouton numérique, Rencontre avec Antoinette Rouvroy : gouvernementalité algorithmique et idéologie des big data, ***You Tube****, 6 mars 2018*

66 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, ***Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet****, Larcier, 2018, pp. 436*

« Des algorithmes peuvent être très intéressants dans la mesure où ils font parler le monde y compris le monde social ».[67] Ils redent visibles certaines portions du réel auxquelles nous n'aurions autrement pas accès. En mettant en lien des points très éloignés (dans l'espace et/ou dans le temps), ils nous font voir des choses qu'on aurait pas vues autrement.

Peut-on en déduire que cette absence de « point de vue » confère à l'algorithme une objectivité totale ? Non. Il est vrai que cela promet des gains d'efficacité et de compétitivité, de productivité, d'ouverture et de transparence. Mais, les algorithmes ne peuvent saisir la vie dans ce qu'elle a de vivant.[68] Les algorithme « défont toutes les formes » car, pour le numériser, ils découpent le réel en une multitude de données indépendantes. Dès lors, ils ne peuvent rendre compte de toutes les formes de la vie (surtout en ce que ces formes peuvent « altérer le projet de la vie qui la constitue »). Ils sont même, en un sens, le contraire de la vie (dans ce qu'elle a de spontané et d'innovant). Pourquoi ? Parce qu'ils visent à contrecarrer toute altération du monde qu'ils prennent en charge. Ils veulent déjouer toute menace et supprimer tout ce qui potentiellement pourrait changer l'ordre établi. Dès lors, ils empêchent à tout changement de survenir. Or, la tendance de tout ce qui vit est précisément l'altération.

Ce qu'il faut critiquer, c'est l'idéologie des big data. Tout n'est pas numérisable d'ailleurs : nos rêves, aspirations, etc.

« L'enjeu, c'est la possibilité de décider. Décider implique l'incertitude et la non prédiction de l'avenir Décider en situation de certitude, ce n'est pas décider. C'est suivre aveuglément un système de recommandations. Décider présuppose l'incertitude, la demande et la réclame ».[69]

TROISIÈME IDÉE : « LES ALGORITHMES SONT IMPARTIAUX MAIS PEUVENT CONDUIRE À DES CONCLUSIONS ERRONÉES ».

Plutôt que de gouverner les gens en raison de leur appartenance, on les catégorise en fonction de leur profil (par exemple « fraudeur potentiel », « terroriste potentiel ») et, pour ce faire, les algorithmes se basent sur des comportements passés de la personne ou de personnes qui lui ressemblent. Cette forme de gouvernementalité porte non pas sur des actes mais sur des potentialités.

Les enseignements que fournissent ces données appréhendées jouissent d'une aura d'objectivité et d'impartialité car les algorithmes n'appréhendent plus les personnes en passant par des catégories idéologiquement marquées (femme, immigré, étranger, pauvre, riche, jeune, vieux, etc.). Les algorithmes ne recourent plus aux catégories socialement éprouvées. Les algorithmes ne font pas de politique. Ils n'ont pas de préjugés sociaux. Ils se désintéressent de ces catégories mais aussi des individus singuliers. Ils transpercent, traversent, transcendent les personnes. Ils se contentent de déterminer des profils. Des profils très détaillés mais impersonnels et prédictifs. Ils ne s'intéressent plus à ce que pense, dit ou fait la personne mais à ce qu'il désire ou plutôt à ce que les algorithmes, qu'elle allait désirer. Et cela avant même qu'elle n'en prenne conscience elle-même. A la rigueur, nos désirs nous précèdent.

Cette impartialité, cette objectivité leur confère un caractère d'incontestabilité. Cette incontestabilité procède aussi de l'exhaustivité des bases de données dont on peut dégager régularités et corrélations. En effet, contrairement à la statistique traditionnelle qui s'appuie sur des échantillons, l'algorithme s'appuie sur la totalité des données, en ce compris les valeurs que la statistique écarte car jugées non significatives (les points trop éloignés de la moyenne : ce qu'on appelait « le bruit »).

On ne peut pas accuser l'algorithme, comme on le faisait naguère pour le statisticien, d'avoir pris un échantillon trop réduit (ou trop large et donc non pertinent). Il y a une prétention à l'exhaustivité et à la non-sélectivité.

Il est donc difficile de remettre en cause cette incontestabilité. Pourtant, un algorithme est toujours conçu par des êtres humains qui choisissent de hiérarchiser certaines informations par rapport à d'autres, de pondérer certaines, etc. Il ne faut pas se laisser abuser par cette prétendue neutralité. Par ailleurs, des données anonymes qui sont corrélables le sont parfois par pur effet du hasard. Cela risque d'arriver d'autant plus fréquemment que la taille de ces bases de donnée devient gigantesque. Notons aussi que des erreurs d'interprétation restent possibles relativement à ces corrélations. Si A et B sont corrélés, c'est peut-être A qui cause B. Mais, cela peut aussi être l'inverse. Cela peut être aussi C qui n'a pas été aperçu et qui est la cause des deux.

La logique d'optimisation fondée sur l'examen de corrélations dégagées de l'examen des données massives peut certes s'avérer très précieuse. Mais, il est des cas où se satisfaire des corrélations est insuffisant voire dangereux. Par exemple, la « black box » médecine qui est en train de se développer dans les pays anglo-saxons. En négligeant la causalité, c'est-à-dire, en l'occurrence, l'explicitation des causes de la maladie ou de l'état qu'on prétend soigner, on risque de faire des erreurs de prescriptions aux conséquences désastreuses voire fatales.

Autre « erreur » des algorithmes limités aux probabilités comportementales ? Antoinette Rouvroy, citant un exemple tiré d'un livre de Grégoire Chamayou (Théorie du drone), évoque l'hypothèse d'une personne vivant au Moyen Orient qui recevrait dans sa boîte mail des messages non sollicités d'une association terroriste et qui, en outre, se serait rendu plusieurs fois en Syrie pour des voyages d'affaires. Cela pourrait suffire à l'algorithme à le désigner comme cible pour une attaque de drones.[70]

---

67 Ibidem

68 A. Rouvroy, La gouvernementalité algorithmique ou l'art de ne pas changer le monde, You Tube, ***Ecole Normale Supérieure*** *(PSL), 29'11", 1er décembre 2016*

69 A. Rouvroy, Big data et anticipation : vers une gouvernementalité algorithmique ?, You Tube, ***Séminaire, Espace Ethique/IDF****, 36'33", 16 avril 2015*

70 A. Rouvroy, La gouvernementalité algorithmique ou l'art de ne pas changer le monde, You Tube, ***Ecole Normale Supérieure*** *(PSL), 1er décembre 2016*

Il faut laisser une place à la délibération humaine une fois les corrélations détectées. Il y a des éléments non numérisables et il est intéressant que les gens prennent conscience des corrélations existant entre leurs comportements.

QUATRIÈME IDÉE : « POUR UN ALGORITHME, UNE CHOSE QUI EXISTE SOUS FORME DE POTENTIALITÉ EST UNE CHOSE QUI EXISTE DÉJÀ ».

Il n'y a pas de futur pour les algorithmes. Ce qui est possible est réel. L'algorithme agira alors par avance sur cette potentialité pour l'empêcher de survenir si elle est jugée indésirable. C'est ce qu'on appelle la préemption. Imaginons, dit Antoinette Rouvroy, que des compagnies d'assurance listent les personnes qui ont l'habitude de fréquenter les forums relatifs à la violence conjugale. Elles décideront probablement de ne pas les assurer sur la vie car elles appartiennent à une catégorie de personnes menacées qui, à leurs yeux, sont déjà mortes. On fait « comme si » ce qui existe sous la forme de la potentialité est considéré comme déjà réalisé par la gouvernementalité algorithmique.

L'algorithme propose d'agir avant que l'évènement indésirable ne se réalise. Et cela en agissant sur le contexte, parfois même à l'insu du potentiel délinquant.

La gouvernementalité algorithmique diffère fondamentalement de la gouvernementalité par le droit. En effet, le droit contient en lui la possibilité de désobéissance. Et le fait que la loi puisse être désobéie est essentiel au métabolisme juridique. Pourquoi ? Parce que la personne qui désobéit le fait peut-être parce qu'elle trouve la loi injuste, inefficiente, etc. Dès lors, la personne pourra expliquer devant la juridiction les raisons de son action. Et celles-ci seront peut-être reprises dans la jurisprudence qui, à son tour, en tant que source du droit, pourra peut-être inspirer le législateur pour faire évoluer voire changer la norme.

A contrario, en empêchant l'évènement indésirable de se produire, la gouvernementalité algorithmique empêche l'évolution du droit. Elle est profondément conservatrice. « Les algorithmes considèrent comme déjà advenu ce qui n'existe que sous une forme potentielle ».[71] Cette gouvernance va tenter de faire avorter la potentialité indésirable. Imaginons une personne détectée, erronément ou à juste titre, comme potentiellement terroriste. On la surveillera. Mais les algorithmes vont tenter d'empêcher la potentialité de s'actualiser. Par exemple en changeant l'environnement de la personne, en triant tout à quoi elle pourra avoir accès, en écartant des résultats de ses requêtes internet tout ce qui pourrait nourrir cette appétence au terrorisme, en compliquant son accès aux boîtes de nuit, etc.

Les big data peuvent devenir de précieux outils d'aide à la décision. Mais, cela peut être aussi un piège. En théorie, les algorithmes nous permettent de tirer des conclusions et conduire à des recommandations. On peut toujours décider de suivre ou non ces recommandations. Mais, en réalité, ces systèmes d'aide à la décision peuvent intimider le décideur.[72] Pourquoi ? Parce qu'en s'écartant de la recommandation, il prend le risque, si la probabilité à laquelle la recommandation proposait de remédier se réalise, d'être tenu responsable personnellement de ne pas avoir pu éviter ce risque alors qu'on lui avait recommandé une mesure qui lui permettait de l'éviter. C'est le cas, par exemple, du fonctionnaire qui doit décider de la remise en liberté anticipée ou conditionnelle d'un prisonnier.[73] Même s'il connaît la personne concernée, même s'il a pu se forger une intime conviction sur base de son expérience, il hésitera à désobéir à la recommandation car si la personne récidive, il en sera jugé personnellement responsable. Du coup, ces systèmes d'aide à la décision peuvent devenir des systèmes qui se substituent à la décision.

CINQUIÈME IDÉE : « LE NUMÉRIQUE EST EXACTEMENT LE CONTRAIRE D'UNE RÉVOLUTION : C'EST LA CONSERVATION ABSOLUE DE L'ÉTAT DE FAIT ».

Cette thèse est assez radicale. Pour Antoinette Rouvroy, « les données ne sont jamais des faits mais toujours des effets, c'est-à-dire des effets de rapports de force ».

Les algorithmes ne cherchent pas les causes. Ils ne visent pas à expliquer. Ils établissent juste des corrélations statistiques. Plus les données sont massives, plus ils sont à même d'établir des prédictions. Le but est d'anticiper ce qui peut survenir peu importe pourquoi cela survient.

Les données ne sont pas le langage des choses elles-mêmes. Le monde ne parle pas spontanément à travers les données. Il ne s'agit pas de données « brutes ». Les données sont toujours produites. Les données ne sont pas des faits des effets de rapports de force, de domination, de situations qu'on n'a pas changées alors qu'on aurait pu les changer. Il y a une naturalisation, une neutralisation des faits.

En réalité, selon Antoinette Rouvroy, « les big data ne sont que le reflet partial et passif de l'état de fait qui ensuite est rendu amnésique des conditions de sa production ».[74] Dès lors, faute de connaître le contexte dans lequel elles sont nées, les gens les utilisent comme un référent qui peut inspirer leur action. L'exemple que donne l'auteur, c'est celui de données témoignant de discriminations de femmes dans l'accès aux fonctions de cadres. Certaines entreprises pratiquent cette discrimination. Imaginons un chef d'entreprise qui ne désire pas discriminer mais qui désire juste procéder de manière identique aux autres entreprises et qui se réfère à cet état de fait. Il va reproduire cette discrimination non pas parce qu'il préfère les employés masculins aux employées féminines mais parce qu'il désire juste s'aligner sur les standards.

---

71 A. Rouvroy, Big data et anticipation : vers une gouvernementalité algorithmique ?, You Tube, ***Séminaire, Espace Ethique/IDF****, 26'13", 16 avril 2015*

72 A. Rouvroy, La gouvernementalité algorithmique ou l'art de ne pas changer le monde, You Tube, ***Ecole Normale Supérieure*** *(PSL), 1h37'05", 1er décembre 2016*

73 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, ***Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet****, Larcier, 2018, pp. 433*

74 A. Rouvroy, La gouvernementalité algorithmique ou l'art de ne pas changer le monde, You Tube, ***Ecole Normale Supérieure*** *(PSL), 1h06'05", 1er décembre 2016*

Contrairement à la gouvernementalité algorithmique, le droit est une dogmatique consciente d'elle-même et c'est la raison pour laquelle il organiserait lui-même les conditions et les processus de remis en question de ses propres productions. Cette contestabilité de la norme juridique est précisément une vertu du droit. De manière fort poppérienne, Antoinette Rouvroy propose d'ailleurs que le caractère contestable d'un algorithme soit un critère pour juger de sa valeur.

SIXIÈME IDÉE : « IL FAUDRAIT POUVOIR AUDITER LES ALGORITHMES »

Les algorithmes sont mus par des métriques. Ces dernières pondèrent l'importance de certaines données par rapport à d'autres données d'une autre nature. Ces métriques ont bel et bien été décidées par les programmateurs et cela en raison d'objectifs et ces derniers ne sont pas nécessairement élucidés. Il faudrait pouvoir rendre les intentions transparentes. En bref, on devrait pouvoir auditer, interpréter ces algorithmes. C'est possible techniquement dans une certaine mesure. Mais l'auteur précise que c'est plus difficile quand il s'agit d'algorithmes auto-apprenants. Ils apprennent en fonction des données qu'ils recueillent. Les données « enseignent » à l'algorithme. Ces derniers deviennent alors « aussi imprévisibles que la sauvagerie des faits ». Ces algorithmes auto-apprenants deviennent alors très opaques.

SEPTIÈME IDÉE : « LA GOUVERNEMENTALITÉ ALGORITHMIQUE EST UN SYMPTÔME ET UN ACCÉLÉRATEUR DU CAPITALISME »

Antoinette Rouvroy[75] reprend la définition que donnent du capitalisme Deleuze et Guattari : « libération des flux dans un champ déterritorialisé ». Le flux, c'est notamment l'argent, les marchandises, les hommes, etc., bref tout ce dont la circulation est facilitée par la liberté de circulation. Elle estime qu'il y a une continuité évidente entre capitalisme en ce sens-là et la gouvernementalité algorithmique. Le flux, c'est évidemment celui des données. Et cette circulation est dénuée de tout sens selon elle. Tout ce qui « coule » est a-signifiant. La mécanique est bien en place. Et elle asservit. Ou plutôt elle perpétue l'antique domination inhérente au capitalisme.

En réalité, cette gouvernementalité ne gouverne pas des hommes mais des réseaux de données agrégées sous forme de modèle. Il gouverne un « immense corps statistique numérique qui est supra-individuel et qui est composé d'une infinité de données infra-individuelles tirées de leur contexte, réagrégées avec d'autres données qui ont proliféré dans d'autres contextes et à d'autres moment ». Ce grand corps statistique est le seul sujet qui demeure encore.

RECUL CRITIQUE PAR RAPPORT AUX THÈSES D'ANTOINETTE ROUVROY

Les thèses du professeur Rouvroy sont très stimulantes. Au moyen de concepts féconds empruntés à la tradition foucaldienne et deleuzienne, elle nous offre une grille de lecture passionnante du phénomène des big data. Et nous souscrivons pour une large part à son constat. Non, nous n'avons pas accès directement au réel. Non, nous ne pouvons-nous dispenser de l'usage de représentations et de théories pour appréhender le réel. Non, ces données ne sont pas « brutes » : elles sont produites. Oui, les pratiques commerciales actuelles peuvent menacer notre vie privée. Oui, il est vrai que « les algorithmes ne « gouvernent » que dans la mesure où nous renonçons à (nous) gouverner nous-mêmes, dans la mesure où nous sous-traitons à des machines la charge de prendre les décisions qui nous reviennent ».[76] Oui, il faudrait pouvoir auditer les algorithmes.

Pour autant, ces données sont-elles le reflet de rapports de force et de dominations, de normes sociales et de préjugés ?[77] C'est une considération nettement moins pertinente. Elle est d'origine marxiste. **L'ordre social et la place que chacun et chacune occupe dans la société ne sont pas, selon nous, l'expression d'une domination.** Certes, nul ne peut prétendre que l'égalité des chances est pleinement garantie et opérationnelle dans notre pays. Mais on ne peut pas nier non plus que le travail, le mérite, le talent, la persévérance, la créativité, la capacité à répondre aux demandes du public, etc. sont des qualités qui entrent pour une large part dans la constitution de l'ordre social.

Les algorithmes sont-ils conservateurs ? Visent-ils à perpétuer le statu quo en désamorçant tout ce qui, dans les potentialités, **pourrait conduire au changement ?** On serait tenté de répondre : tout dépend des algorithmes. Il semble qu'Antoinette Rouvroy confonde ici deux choses : les algorithmes entendent anticiper la survenue d'accidents. Cela ne signifie qu'ils seraient hostiles au changement.

On ne peut pourtant pas nier que des changements considérables, révolutionnaires accompagnent l'irruption des big data dans notre vie de tous les jours. Les algorithmes changent réellement l'organisation sociale de notre société. Ils nous obligent à revoir les missions de l'enseignement et à réinventer la sécurité sociale. Nous avons d'ailleurs consacré une étude à la robotisation de l'économie.[78] Ils permettent aussi l'émergence d'une catégorie inconnue jusqu'alors, à mi-chemin de l'employé et de l'indépendant : le travailleur autonome. Les algorithmes révolutionnent la médecine qui devient préventive et personnalisée. Ils révolutionnent la génomique, la neurologie, l'épidémiologie, etc. Ils révolutionnent le secteur de l'énergie avec la consommation intelligente. Avec la mise en place des « Smart Cities », ils pourront probablement une réelle solution à la mobilité et, notamment, les problèmes de parking et de congestion. Etc. Le terme n'est pas usurpé : la révolution numérique est bel et bien une révolution.

---

75 A. Rouvroy, La gouvernementalité algorithmique : radicalisation et stratégie immunitaire du capitalisme et du néolibéralisme ?, ***La Deleuziana*** *– Revue en ligne philosophique – ISSN 2421)309 N.3/2016*

76 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, ***Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet****, Larcier, 2018, pp. 441*

77 A. Rouvroy, Homo juridicus est-il soluble dans les données ?, ***Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet****, Larcier, 2018, pp. 440*

78 C. de Salle & alii, ***Accompagner la robotisation de l'****économie, Etude du Centre Jean Gol, 2017*

Ce qui est vrai, c'est que cette révolution ne remet pas en cause l'économie de marché et c'est peut-être en ce sens-là qu'Antoinette Rouvroy l'accuse de préserver l'ordre établi. Mais, encore une fois, la réalité aujourd'hui, c'est une économie mixte, d'essence sociale-démocrate. Pas une économie de marché. Laquelle constitue une véritable force de progrès. Pas une force conservatrice. Et cela dans tous les domaines et à tous les niveaux. Y compris au niveau social car ce système est celui qui permet d'assurer la plus forte rotation des élites.

Par ailleurs, **ces algorithmes sont-ils l'expression la plus dépouillée de la logique capitaliste ?** Le capitalisme vise effectivement à faire circuler les « flux ». C'est entendu. C'est effectivement l'esprit de la liberté de circulation (des biens, des services, des capitaux et des hommes). Cela signifie-t-il que ces opérations sont dénuées de sens ? En soi, elles n'ont aucun sens mais elles auront le sens que chacun veut bien leur donner. Le sens qu'il leur est utile de lui donner. **Chaque acteur économique poursuit un ou plusieurs projets qui fait ou font sens pour lui. Il est faux de prétendre,** à l'instar d'auteurs tels que Badiou, **que le libéralisme (ou le « néo-libéralisme ») serait « nihiliste »**. Ce qui différencie le libéralisme d'autres idéologies, c'est que les finalités ne sont pas collectives. Il ne revient pas à l'Etat d'assigner une finalité à tous. Ce sont les individus qui déterminent et poursuivent chacun leurs propres finalités.

Si on adopte le point de vue de Sirius, tout, évidemment paraît dérisoire. L'argent n'est rien d'autre que du papier. Les dealers dans une salle de marché sont de petits insectes vibrionnants. Etc. Pourtant, ces ordres de marché sont l'expression agrégée de l'offre et de la demande en provenance de plusieurs milliards d'humains qui, à travers ces flux qui apparaissent comme de pures abstractions, font avancer leurs propres projets qui donnent sens à leur vie. De la même façon, les big data, vus de l'extérieur, constituent le flux et le reflux d'une masse proliférante et obscène d'informations décontextualisées qui parcourent un grand corps statistique et entre lesquelles des algorithmes s'emploient à découvrir des corrélations. **Mais ce flux numérique n'est pas, comme le prétend Rouvroy, un « stratégie immunitaire » du capitalisme. Il s'agit plutôt de son système sanguin.**

En d'autres termes, ces données peuvent aussi être produites, captées, traitées et analysées par quantité d'acteurs et servir d'outils à leurs projets. Elles ne sont pas le monopole des GAFAM. Ce sont des ressources qui peuvent être exploitées par chacun. Il appartient à chacun de créer ses outils de mesure, ses algorithmes ou d'utiliser les données produites par ceux des autres afin de réaliser son projet. Antoinette Rouvroy - et nous disons cela sans lui ôter en rien le mérite de son analyse originale et stimulante de la problématique qui nous occupe - s'inscrit dans une tradition foucaldienne, derridienne, deleuzienne, guattarienne, etc. plutôt hostile au « néo-libéralisme » (notion que, par ailleurs,[79] nous récusons). A ce titre, elle estime que ce modèle est dominant (ce qui est faux parce que ce qui domine aujourd'hui, ce sont les recettes de la social-démocratie) et que les algorithmes sont au service de cette idéologie prétendument dominante. Raison pour laquelle elle pense que ces outils favorisent et renforcent une logique d'exclusion.

Un exemple qu'elle utilise souvent dans ses conférences, est, nous l'avons dit, celui de la femme qui va consulter un forum sur les violences conjugale et qui, de ce fait, en raison du risque de décès prématuré auquel sont exposées les femmes battues (dans la groupe duquel l'algorithme l'a rangée instantanément du simple fait qu'elle ait consulté ce forum) est déjà tenue pour morte par l'algorithme prédictif qui est programmé pour ne pas faire de différence entre le potentiel et l'actuel. Dès lors, dit-elle, la compagnie d'assurance va évidemment l'éjecter ou augmenter sa prime d'assurance à un niveau tellement absurde que cela équivaudra à une exclusion. Mais, c'est là méconnaître la dynamique capitaliste et la loi fondamentale de l'offre et de la demande. Si une compagnie commence, sur base du calcul prédictif des algorithmes, à exclure tous ses affiliés qui présentent un risque important, les affiliés qui demeurent inscrits pourront se dire, à très juste titre, que s'ils demeurent assurés, c'est que l'algorithme a déterminé qu'ils ne couraient aucun risque sérieux. Dès lors, ils pourront exiger une réduction de prime voire quitter cette compagnie pour une autre. Au point que l'économie réalisée par l'assurance en éjectant des personnes qu'elle estime posséder un profil trop risqué sera largement annulée par la perte d'une bonne partie de ses affiliés au profil classique.

De manière plus fondamentale, imaginons que les algorithmes développent une puissance prédictive exceptionnelle. Une compagnie d'assurance qui exclurait tous les profils présentant le moindre risque se condamnerait elle-même à mort. En effet, procéder ainsi serait admettre elle-même qu'elle ne sert strictement à rien. Ce serait la fin du système assurantiel. Et il est possible que l'évolution amènera ces compagnies à devoir se repenser et se réinventer, à proposer des services d'une autre nature. Imaginons que, en raison des progrès de la médecine préventive et prédictive, elle découvre que l'un de ses affiliés va développer une maladie nécessitant de lourds traitements dans 15 ans. La logique capitaliste la poussera peut-être à devenir bancassureur et à remplacer les primes par une sorte d'épargne permettant à la personne de faire face, le moment venu, à cette maladie. Si elle ne le fait pas, d'autres compagnies concurrentes le feront.

L'analyse d'Antoinette Rouvroy à ce niveau se focalise exclusivement sur les dangers en négligeant les opportunités. Or, c'est le propre de la mécanique capitaliste de se réinventer à toute moment, de se « métamorphoser ». On retrouve aussi, chez Antoinette Rouvroy, une propension à personnaliser la révolution algorithmique, à en faire un sujet, une instance, un « grand corps statistique ». Elle déplore l'inanité de cette gouvernementalité sans sujet, cette machine absurde. En réalité, cela n'a pas de sens, pensons-nous, de faire du fétichisme : le marché, ce sont des hommes et des femmes qui constituent en même temps une demande et une offre de biens et de services. Algorithmes et big data ne sont pas des fins en soi mais uniquement des outils permettant, entre autres, d'ajuster de fluidifier, cette offre et cette demande. Ce qui est vrai, c'est qu'il faut veiller à ce qu'ils restent des outils.

79 C. de Salle, Le néo-libéralisme est une mystification intellectuelle, ***La Libre Belgique****, 14 février 2018 www.lalibre.be/debats/opinions/le-neoliberalisme-une-mystification-intellectuelle-opinion-5a83132bcd70f924c8024697*

Raison pour laquelle, il faudrait, dans le respect des droits de propriété intellectuelle, pouvoir exiger des GAFAM une **meilleure transparence des algorithmes** afin de connaître les détails des modes de calcul et des partis-pris qui sont adaptés. Au-delà, il faudrait s'atteler à la **rédaction d'une « charte des algorithmes »** reprenant un certain nombre de principes permettant de garantir l'accès à l'information pour tous les internautes et la **préservation d'un « monde commun ».**

Enfin, les algorithmes clôturent-t-elles la « personne » ? Empêchent-ils celle-ci de poursuivre son processus toujours inachevé de subjectivation en l'enfermant irrémédiablement dans un profil ? Il est certain que les algorithmes sont particulièrement avides de cette matière première (qui est en réalité produite) que sont les big data et qu'ils s'emploient à constituer des profils de tout un chacun. Il est certain aussi que les algorithmes ne s'intéressent pas à la personne en tant que telle mais aux comportements que celle-ci aura dans le futur. Cela dit, cette indifférence des algorithmes à la personne, à sa singularité indépassable ne signifie pas que la personne disparaisse. Pour le dire simplement, ce n'est pas parce que les algorithmes ne s'intéressent pas au sujet que ce dernier cesse d'exister. Croire cela, c'est succomber à une illusion contre laquelle Antoinette Rouvroy elle-même met en garde : croire que les algorithmes définissent, décrivent la réalité. En réalité, ils ne font que ponctionner des signaux de la réalité. Ils les traitent et en dégagent des corrélations qui sont autant d'opportunités pour agir dans tel ou tel sens. Les algorithmes ne disent pas le Réel. Ils l'instrumentalisent.

Certes, cette opération n'est pas sans danger : à force de traiter les gens comme des objets, on peut les transformer en objets. Mais, si ce danger est réel, ce n'est pas à cause des algorithmes et des gens qui les mettent en œuvre mais en raison de la propension des gens à autoriser les machines à prélever leurs données les plus intimes, à abdiquer de leur esprit critique et à s'en remettre aux algorithmes quand il s'agit de décider, etc.

Comme le disait La Boétie, la servitude est toujours volontaire. De la même façon que les gens font preuve d'une désinvolture coupable en étalant leur vie privée sur les réseaux sociaux, la propension de ces derniers à s'en remettre aux algorithmes parce qu'ils nous facilitent la vie, parce qu'ils nous guident, nous conseillent et nous soulagent de bien des peines et démarches, porte en elle le germe d'un asservissement futur.

Il importe que les algorithmes restent des outils. Pour cela, il faut prendre des mesures pour qu'ils ne deviennent pas des boîtes noires. Il faut pouvoir les auditer, les évaluer et en débattre sur la place publique. Il faut assurer un pluralisme des algorithmes et les mettre en concurrence entre eux. De la même façon qu'il existe quantité de forums à propos de tel ou tel produit ou innovation (le dernier IPhone, le dernier Mac, etc.), il faudrait des forums sur les différents algorithmes.

VI. PENSER LA DONNÉE

Laurent Hublet

On entend souvent dire que les données sont le carburant des algorithmes, et par extension de la révolution numérique actuellement en cours. Le volume de données produites chaque jour donne d'ailleurs le tournis : pour l'année 2017, il a dépassé le volume total de données générées dans les 5000 années précédentes.

Commençons par le commencement, avec une question simple et fondamentale. Qu'est-ce que « la donnée » ? Le terme « donnée » semble aller de soi, et pourtant sa signification nous joue beaucoup de tours.

MAL DONNE SUR LA DONNÉE

Jusqu'il y a peu, le terme « donnée » avait deux sens en français. D'une part, « la donnée » a signifié une forme particulière de don : la distribution d'argent aux pauvres. Cette utilisation s'est perdue aujourd'hui, mais on la retrouvait encore chez Saint-Simon à la fin du 18e siècle : « *Plus la donnée avait été nombreuse, plus la charcutière était aise* ».

D'autre part, « une donnée » signifie un point de départ irréfutable d'un raisonnement, et plus particulièrement ces dernières années, le point de départ d'un raisonnement ou d'une analyse faisant appel à des outils statistiques.

On pourrait donc dire qu'une donnée est « une mesure d'un état de choses à un moment donné ». Cette définition préliminaire a l'air toute simple en apparence, mais elle cache en réalité plusieurs malentendus.

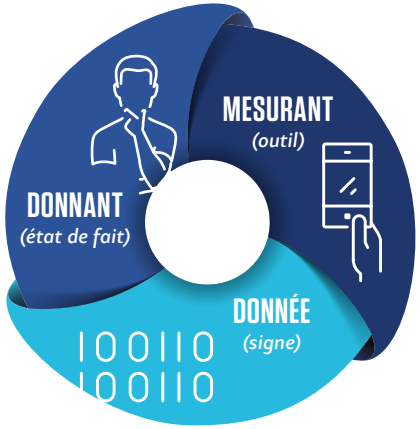
UN PREMIER MALENTENDU

Une donnée n'existe pas en soi. Elle requiert un **système** reposant sur deux conditions :

1. Pour qu'il y ait « donnée », il faut un « donnant » (l'« état de choses » dans notre définition)

2. Il n'y a pas de relation directe donné/donnant : il faut un « agent mesurant ». Ce mesurant, c'est l'instance technologique ; il faut nécessairement un outil (technè en grec ancien) pour capturer une donnée

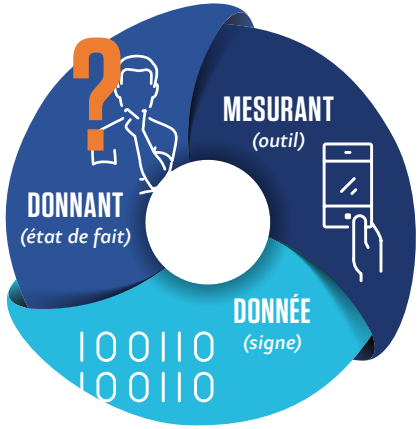
Derrière la donnée se cache donc un triangle donnée-mesurant-donnant



La langue française nous joue un premier tour, qui fait émerger un paradoxe. **La donnée ne se donne pas. La donnée n'est pas donnée. *La donnée se produit.*** Plus précisément, la donnée est le produit d'un système qui nécessite un donnant et un mesurant. Et qui dit production dit coût de production. Ce n'est pas donné de produire une donnée.

UN DEUXIÈME MALENTENDU

Attardons-nous quelques instants sur l'un des participants de notre triangle, le « donnant ». Que peut-il être ? N'importe quel élément de la réalité ? Un objet (technologique ou naturel), un animal, un humain,... ? Absolument !



Pour autant, est-il facile d'identifier précisément le donnant d'une donnée ? Prenons l'exemple suivant : vous marchez en rue, avec votre téléphone dans votre poche. Toutes les quelques secondes, votre téléphone émet un signal GPS qui est mesuré par une application de géolocalisation. Qui est le donnant ? « Je suis le donnant » me répondrez-vous ! Certes, mais êtes-vous le seul donnant ? Vous marchez probablement sur un trottoir : le propriétaire de ce trottoir (la municipalité par exemple) est également « donnant » au système de données « vous marchant en rue ». La réponse est donc moins simple qu'il n'y parait à première vue : il peut y avoir pluralité de donnants. Lorsque le donnant est un être vivant, l'unicité de donnant est plutôt l'exception que la règle[80].

Le donnant est souvent pluriel, et cela pose un problème très concret. On parle beaucoup de « données à caractère personnel ». C'est d'ailleurs le socle de la législation européenne sur les données[81]. Mais que signifie « personnel » dans les cas où il y a plusieurs donnants ? Et particulièrement, que se passe-t-il lorsque l'un des donnants est un bien public ?

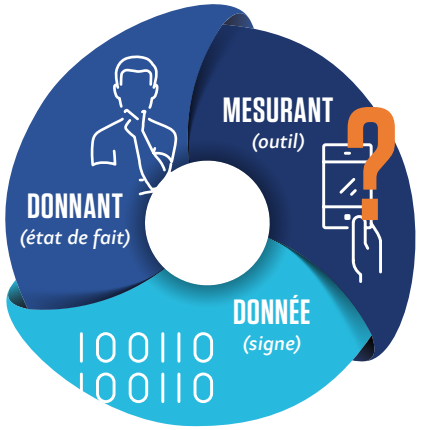
Selon la perspective que l'on prend, une donnée peut être à la fois personnelle et non-personnelle. La pluralité de donnants doit donc nous inviter à la prudence quant à l'utilisation du concept de « donnée personnelle ». Cette prudence est indispensable : **rendre une donnée personnelle alors qu'il y a pluralité de donnants revient à privatiser une ressource collective voire un bien commun**.

Non seulement le donnant peut être pluriel, mais il peut également être équivoque. Prenons le cas de votre voiture à l'arrêt qui émet une donnée sur sa localisation. Qui se cache derrière le donnant ? Le propriétaire de la voiture ou le constructeur ? Ou les deux ?

UN TROISIÈME MALENTENDU

Reprenons notre exemple du marcheur avec son GSM en poche. Qui est le mesurant ? Le téléphone me répondrez-vous !

Certes, mais est-ce le fabricant du téléphone (hardware), de son système d'exploitation ou d'une application mobile installée sur le téléphone (software) ? Est-ce le gestionnaire du réseau de télécommunication (l'opérateur téléphonique), ou du réseau satellite ? Bien souvent, plusieurs acteurs sont impliqués ici aussi... Caramba !



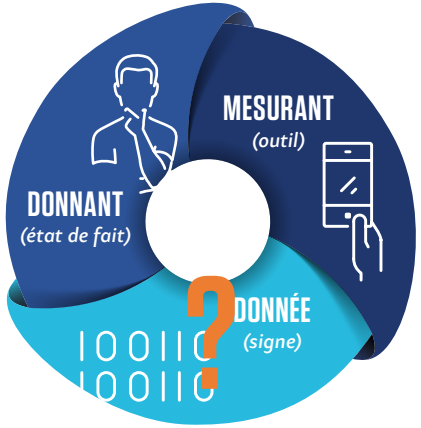
Et pour compliquer encore les choses : une même donnée est susceptible d'être mesurée par différents mesurants. La production d'une donnée par un système n'empêche pas nécessairement la production de cette même donnée par un autre système... Caramba (bis) !

UN QUATRIÈME MALENTENDU

Continuons notre investigation du (ou des) agent(s) mesurant(s). Que se cache-t-il derrière lui ? Pourquoi mesure-t-il une donnée ?

Une donnée n'est pas une fin en soi – elle est nécessairement utilisée en vue de quelque chose : piloter un avion automatiquement, réaliser des études statistiques, proposer une publicité en ligne, etc.

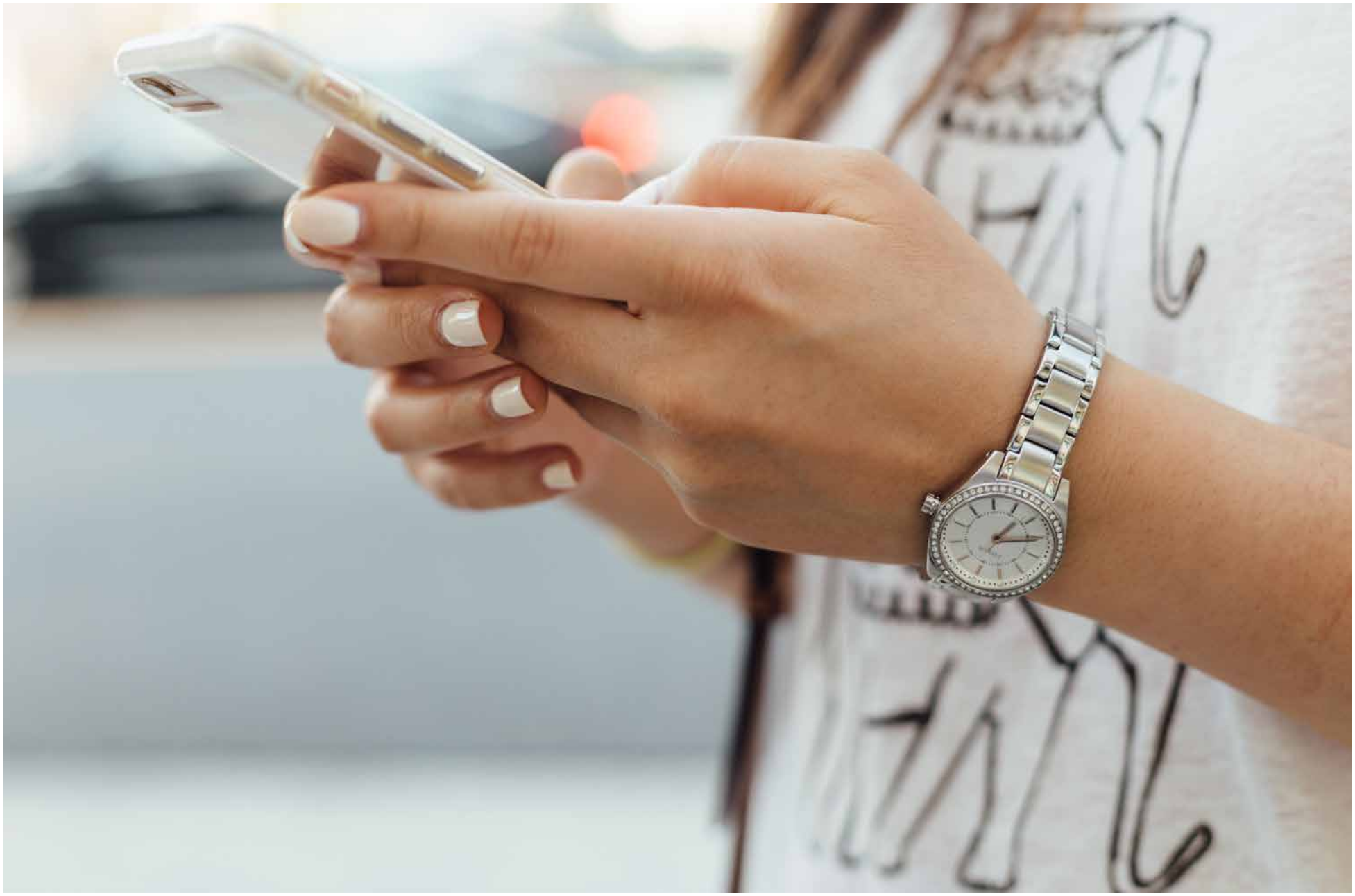
Bien souvent, une donnée isolée n'est pas susceptible d'une utilisation pertinente. Il faut un nécessairement un ensemble de données pour que celles-ci fassent sens. Dès lors, il convient sans doute d'utiliser le pluriel pour parler de données, comme le fait d'ailleurs l'anglais avec « data » qui est toujours au pluriel[82].



80 Un exemple de donnant vivant unique serait par exemple une donnée concernant le patrimoine génétique d'un être vivant. L'unicité du donnant touche donc à une question philosophique très profonde (et qu'on n'approfondira pas ici) : ce qu'un être ou une chose est en dehors de tout autre être ou chose.

81 La régulation sur la protection générale des données mais également la libre circulation des données non personnelles (considérée par certains comme la 5e liberté du marché unique)

82 Les Anglais utilisent le terme "data point" (point de données) pour nommer la donnée au singulier.



A QUI APPARTIENNENT LES DONNÉES ?

Les données sont donc le produit de systèmes donnants-mesurants-données, chaque terme pouvant être pluriel et équivoque. Ceci nous amène à la conclusion suivante : la question, si simple en apparence, « qui est le propriétaire des données ? » est en fait une question extraordinairement complexe. Les données sont susceptibles d'avoir de multiples prétendants légitimes à leur propriété ; chercher à déterminer « le » propriétaire, c'est souvent prendre le risque de se perdre dans le triangle des Bermudes.

Revenons à l'exemple de notre marcheur avec son téléphone en poche : qui est propriétaire légitime des données ? Le marcheur ? La municipalité ? Le gestionnaire de réseau télécom ? Le fabricant du téléphone ? Le développeur d'une application mobile ?

Pour sortir de cette impasse, il faut repenser la notion de « propriétaires de données » et la remplacer par celle de **(co-)-producteur de triangles de données**.

Insistons sur un point : le donnant est bien un co-producteur dans ce système. Il serait trompeur de le cantonner à un simple rôle de consommateur (approche que l'on retrouve dans certaines législations).

Il faut donc poser la question de la relation entre ces producteurs de données, et c'est ce que nous allons faire par l'entremise de la notion de « valeur ».

EN RÉSUMÉ

Les malentendus soulevés ci-dessous nous conduisent à quelques points de départ pour repenser le big data :

- La donnée n'est pas donnée – elle est co-produite
- Ce n'est pas donné de co-produire des données
- Une même donnée peut avoir de multiples donnants
- Différents mesurants peuvent co-produire une même donnée
- La donnée fait généralement sens quand elle est multiple
- La valeur de la donnée est extrinsèque

Ces points de départs ouvrent la porte de nombreuses questions nécessitant un vrai débat public, dont entre autres :

- Quelle sont les données à caractère « non équivoquement » personnel ? Quelles doivent être les forteresses imprenables de l'intimité numérique ?
- Quel participant à la production de données peut décider de quoi dans l'utilisation faite de celles ?
- Qui détermine la valeur de la donnée ? Selon quel mécanisme ?
- Comment peut-elle être équitablement répartie entre ses co-producteurs ?

VII. LES DONNÉES PERSONNELLES : MYTHES ET RÉALITÉS

Jérôme De Cooman[83] et Nicolas Petit[84]

INTRODUCTION

Comme tout phénomène technologique émergent, les données massives interpellent. Une fois n'est pas coutume, les projections économiques sensationnalistes font les manchettes de la presse. Nous avons sélectionné celles qui nous semblent les plus fréquentes pour les passer au crible d'une contre-expertise logique et empirique. Elles sont au nombre de trois :

1. les données sont le nouvel or noir ;
2. les données personnelles ayant une valeur marchande, les individus devraient pouvoir les vendre ;
3. les réglementations spécifiques sur la protection des données (comme le « RGPD »[85]) favorisent les monopoles au détriment des petites et moyennes entreprises.

1. PREMIER MYTHE : « LES DONNÉES SONT LE NOUVEL OR NOIR »

La presse aime présenter les données massives comme un nouvel or noir[86]. La métaphore est intuitivement forte. Comme l'or au XIX*ème* siècle ou le pétrole au XX*ème* siècle, une véritable course aux données, connue sous le nom de « digitalisation », semble s'être ouverte dans l'économie du XXI*ème* siècle. Dans une perspective économique, l'analogie est pourtant fallacieuse.

Les données n'ont d'abord rien à voir avec les *ressources naturelles* que sont l'or ou le pétrole. Présentes en quantités finies sur la planète, l'or et le pétrole sont des ressources non renouvelables qui s'épuisent par l'usage. En résulte une rareté croissante qui s'accompagne de problèmes bien connus d'allocation à court terme et de conservation à long terme. Ces dernières décennies, le progrès technologique a réduit le rythme d'épuisement des ressources naturelles par l'identification de substituts ou l'adoption de techniques moins voraces en ressources naturelles[87]. De toute évidence, les données massives ne peuvent être qualifiée de la sorte[88], elles qui se renouvellent instantanément, exponentiellement et perpétuellement. Si Malthus était des nôtres, il parlerait sans doute d'une augmentation géométrique, et non arithmétique, des données massives, en tous points corollaire de la croissance démographique.

---

83 Assistant à la Faculté de Droit de l'Université de Liège (ULiège).

84 Professeur ordinaire à la Faculté de Droit de l'Université de Liège (ULiège) ; Research Professor, University of South Australia (UniSA) ; Visting Fellow, Stanford University Hoover Institution.

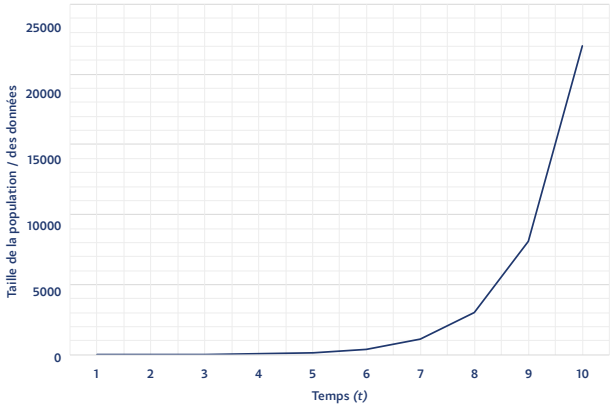
85 Règlement (EU) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), JO, L119/1, disponible sur www.eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32016R0679&from=FR (ci-après, RGPD).

86 ***Voy****. en guise d'illustrations : Kompass, « Les données, le nouvel or noir de votre business (livre blanc) », disponible sur www.prokcssblog.blob.core.windows.net/edito-fra/FRA/2018/01/livre_blanc_kompass_La-donnee-le-nouvel-or-noir-de-votre-business.pdf ; J.-F. Marie, « Data, le nouvel or noir », **NetApp**, s.d., disponible sur www.netapp.com/fr/company/news/press-releases/news-rel-20171120-365886.aspx ; D. Mezinis, « Les données personnelles, le nouvel or noir… », **Boursier.com**, 9 avril 2018, disponible sur www.boursier.com/actualites/macroeconomie/les-donnees-personnelles-le-nouvel-or-noir-761834.html. On notera que cette métaphore est elle-même une réutilisation d'une première comparaison, à savoir celle du pétrole au regard de l'or. Si le pétrole est aujourd'hui qualifié d'« or noir », c'est justement en raison de sa grande valeur économique et des similitudes entre la course à la recherche des principaux gisements pétrolifères et la ruée vers l'or.*

87 Certaines ressources naturelles sont imparfaitement renouvelables, c'est-à-dire qu'il faut plus de temps aux divers processus naturels concernés pour régénérer la ressource qu'il ne faut de temps pour l'épuiser.

88 L. Determann, « No One Owns Data », in 70 ***Hastings Law Journal****, Research Paper No. 265, 23 février 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3123957 (ci-après, Determann, « No One Owns Data »).*

REPRÉSENTATION SIMPLIFIÉE DE LA CROISSANCE EXPONENTIELLE DE LA POPULATION / DES DONNÉES



De plus, les économistes estiment qu'il y a une rivalité dans la consommation des ressources naturelles, absente dans le cas des données[89]. Si A consomme un baril de pétrole, cette quantité doit être retranchée des ressources pétrolières totales disponibles pour B, C, D ... Z. De nouveau, la rivalité dans la consommation des ressources naturelles fait défaut dans le cas des données massives. La communication à B de la date d'anniversaire de A n'épuise pas la disponibilité de cette information pour C, D ... Z. En conséquence, les données massives ne sont pas, en raison de leurs propriétés endogènes, des sources de barrières à l'entrée infranchissables ou de positions de monopole.

Enfin, les chiffres démentent toute analogie entre les données massives et l'or noir. Selon International Data Corporation (ci-après, « IDC »), la valeur du marché des « big data » et des « *business analytics* » avoisinerait 133,046 milliards de dollars[90]. A cela, il faut ajouter le marché de la publicité en ligne, d'une valeur de 200 milliards de dollars en 2017[91].

Au total, le marché des données représenterait une valeur annualisée de 333,046 milliards de dollars[92].

Impressionnants en valeur absolue, ces chiffres demeurent modestes lorsque mis en regard avec l'industrie du pétrole. En 2017, la production annuelle de pétrole s'élevait à environ 34 milliards de barils[93]. Le cours moyen annuel du baril en 2017 étant de 54,25 dollars[94], la valeur du marché pétrolier en 2017 avoisinait les 1 834,566 milliards de dollars[95]. A cette aune, l'industrie des données massives ne vaudrait « que » 18,15 % de l'industrie du pétrole. Certes, le marché des données massives est en pleine croissance. Rien ne permet d'exclure qu'il ne dépasse un jour l'industrie pétrolière. Notre estimation est pourtant généreuse. D'autres méthodes, plus conservatrices, aboutissent à des ordres de grandeur encore inférieurs.[96]

89 Cette absence de rivalité n'est pas complètement exacte : on peut retrouver une certaine forme de rivalité dans la temporalité d'utilisation des données : une donnée qui a déjà été utilisée a moins de valeur qu'une autre.

90 IDC, « Revenues for Big Data and Business Analytics Solutions Forecast to Reach $260 Billion in 2022, Led by the Banking and Manufacturing Industries, According to IDC », 15 août 2018, disponible sur www.idc.com/getdoc.jsp?containerId=prUS44215218 (ci-après, IDC, « Revenue for Big Data »).

91 N. Hashai, « Platform End Users as Free 'Data Labor' – Re-distributing the Value Created in Double Sided Markets », 9 février 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3121160 (ci-après Hashai, « Platform End Users as Free 'Data Labor' »).

92 Les prévisions pour 2018 s'élèvent à 166 milliards ; IDC, « Revenue for Big Data ».

93 Plus exactement, 92 649 milliers de barils furent quotidiennement produits en 2017, soit 33 816 885 000 barils à supposer une année de 365 jours ; « BP Statistical Review of World Energy », 67e Ed., juin 2018, disponible sur www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-oil.pdf.

94 Le cours du Brent et les statistiques liées à ses évolutions sont disponibles sur www.cnr.fr/Indices-Statistiques/Tous-les-indicateurs/Cours-du-Brent-date#haut.

95 Plus exactement 1 834 566 011 250 $.

96 En effet, cette première approche n'est pas exempte de critiques : d'une part, elle est basée sur des montants et moyennes arbitrairement choisis et d'autre part, elle vient comparer une valeur de marché en production avec une valeur de marché en vente. Il serait en effet possible, en nous basant sur les informations mises à disposition par Forbes dans le cadre de son « Global 2000 : The World's Largest Public Companies » , de consolider les données relatives aux ventes des entreprises pétrolières du classement, ce qui nous amènerait à la modique somme de 3 810,694 milliards de dollars . Dès lors, on devrait en conclure que l'industrie des données massives vaudrait 8,74 % de l'industrie du pétrole ; Forbes, « Global 2000: the world's largest public companies », disponible sur www.forbes.com/global2000/#374e3d343335d ; Les données que nous avons agrégées sont disponibles sur www.forbes.com/global2000/list/#industry:Oil%20%26%20Gas%20Operations.
Néanmoins, on peut s'étonner de voir la valeur des ventes de l'industrie du pétrole supérieure à la valeur de production de cette industrie. Cela s'explique par le fait que les entreprises sélectionnées ne se limitent pas qu'à l'exploitation de l'or noir, elles ont également étendu leurs activités notamment au gaz et aux énergies renouvelables. Sans ventilation du chiffre d'affaires, il nous est impossible de convenablement déterminer quelle part est à attribuer à la vente des ressources pétrolières, ce qui biaise également cette méthodologie. C'est la raison pour laquelle notre préférence va au premier calcul, selon lequel l'industrie des données au sens large vaut 18,15 % de celle du pétrole.

A tout cela, il faut encore ajouter que la métaphore se trompe de point focal. Pour reprendre une analogie plus opérante, lors de la ruée vers l'or, ce sont les fabricants de pelles et de jeans qui se sont enrichis (dont Levi Strauss).[97] Comme pour l'or noir, **la valeur économique de l'industrie de la donnée ne se localise peut-être pas tant au niveau de la ressource qu'au niveau des infrastructures permettant de les « extraire, exploiter, collecter, raffiner, transporter »**.[98] A cet égard, on voudra bien noter que les activités des agents économiques actifs dans le secteur de la collecte et du traitement des données personnelles requièrent de lourds investissements fixes, mais aussi d'importantes charges variables. Et, contrairement à une idée répandue, leur coût marginal de production n'est ni constant, ni nul. Au contraire, les investissements en sécurité, confidentialité et intégrité des données augmentent avec l'échelle des données considérée. Dès lors, les rendements d'échelle qui avantagent prétendument les grandes plateformes digitales ne sont peut-être pas si grands qu'il n'y parait.

2. SECOND MYTHE : « LES DONNÉES PERSONNELLES SONT NOTRE PROPRIÉTÉ, ET LEUR VENTE DOIT ÊTRE INDEMNISÉE »

Il est fréquent d'entendre que les données personnelles que nous répandons bon gré mal gré relèvent de notre propriété et, le cas échéant, que leur usage par des tiers constitue une vente. Cette affirmation est fausse.

1. Un droit de propriété sur les données personnelles ?

En droit, aucun instrument juridique n'organise un droit de propriété des données personnelles. Le droit de la propriété intellectuelle, et notamment le droit d'auteur, tout d'abord, prévoit des conditions d'originalité et d'inventivité qui excluent les données personnelles de sa protection[99]. *Idem* pour le droit de propriété *sui generis* applicable aux bases de données. Si la structure d'une base de données peut effectivement être protégée par un droit d'auteur[100], cette protection ne s'étend pas aux données elles-mêmes[101]. Enfin, les législations relatives à la protection de la vie privée – le RGPD est une illustration – organisent au mieux un commencement de droit de propriété sur les données personnelles. Tout au plus, les dispositions légales permettent-elles aux usagers d'empêcher certaines catégories de tiers, à savoir les pouvoir publics et les entreprises, d'utiliser leurs données personnelles.[102] Avec le Professeur Strowel, il est plus juste de considérer que « l'extension des prérogatives des individus sur leurs données révèle l'essor du modèle propriétaire »[103].

---

97 P. Yang, « Miners vs. Merchants: How Global Trade Made Men Wealthy during the California Gold Rush », 5 mars 2016, disponible sur www.flexport.com/blog/trade-merchants-rich-california-gold-rush/.

*98 A. Strowel, « Les mutations des droits de propriété intellectuelle sous l'effet du numérique (data ownership, text and data mining, hyperlinking) », in **Les enjeux de l'innovation : quelles politiques ? Quelles gouvernances ?**, B. van Pottelsberghe et al. (Dir.), 22ème Congrès des économistes, Ed. Université Ouverte de la Fédération Wallonie-Bruxelles, Charleroi, 2017, p. 122 (ci-après : Strowel, « Les mutations des droits de propriété intellectuelle sous l'effet du numérique »).*

99 C'est avec justesse que le professeur Hugenholtz relève que les données sont « libres comme l'air » et que seules la créativité et l'innovation sont protégées par le droit de la propriété intellectuelle. Créer un droit de propriété particulier aux données iraient à l'encontre de ce grand principe ; voy. P.-B. Hugenholtz, « Data Property: Unwelcome Guest in the House of IP », 2017, disponible sur www.ivir.nl/publicaties/download/Data_property_Muenster.pdf (ci-après, Hugenholtz, « Data Property ») : « A 'data producer's right' in machine-generated data would ride roughshod over the existing system of intellectual property. It would violate one of the IP system's main maxims that data per se are "free as the air for common use", and that only creative, innovative or other meritorious investment is protected ».

100 Aux termes de l'article XI.186, alinéa 1er, CDE, « les bases de données qui, par le choix ou la disposition des matières, constituent une création intellectuelle propre à leur auteur sont protégées comme telle par le droit d'auteur ».

*101 Aux termes de l'article XI.186, alinéa 2, CDE, « la protection des bases de données par le droit d'auteur ne s'étend pas aux œuvres, aux données ou éléments eux-mêmes et est sans préjudice de tout droit existant sur les œuvres, les données ou autres éléments contenus dans la base de données ». L'article XI.306 du Code de droit économique précise toutefois que la base de donnée sera protégée lorsque sa création a requis un « investissement qualitativement ou quantitativement substantiel ». À cette condition, le producteur de la base de données pourra interdire la réutilisation substantielle du contenu de celle-ci (**voy.** article XI.307, alinéa 1er, CDE). Néanmoins, à strictement parler, cette protection ne peut nous conduire à la conclusion d'un droit de propriété sur les données ; Determann, « No One Owns Data ».*

*102 En effet, ces législations protectrices n'ont en fait comme but que de s'assurer du respect de la dignité humaine, pas de conférer un droit de propriété ; Determann, « No One Owns Data ». Les différents attributs du droit de propriété ne se retrouvent d'ailleurs pas dans un droit permettant aux utilisateurs de protéger leurs données ; « Privacy laws do not incentivize or reward creation or investment, do not regulate the acquisition or transfer of ownership rights to other, and do not apply against everyone. Instead EU data protection laws confer exclusion rights against governments and businesses, but not against individuals acting for personal or household purposes » ; Determann, « No One Owns Data » ; **Voy**. également l'article 2, alinéa 2, du RGPD. Dès lors, vu l'état actuel du droit, il nous est permis de dire qu'il n'existe aucun droit de propriété sur les données, **de lege lata** ; Pour une analyse détaillée des raisons pour lesquelles aucun instrument de droit positif ne protège d'un droit de propriété les données, **voy**. Determann, « No One Owns Data ».*

*103 Strowel, « Les mutations des droits de propriété intellectuelle sous l'effet numérique », p. 128 ; L'ensemble de ce système de réglementations, sur lequel trône désormais le RGPD, ne confère pas aux individus un droit de propriété à strictement parler sur leurs données, mais bien un ensemble de droits dont le faisceau entraine un contrôle de plus en plus étendu sur lesdites données. Ces textes sont notamment la directive 95/46/CE du Parlement européen et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (cette directive a été abrogée par le RGPD), la directive 2002/58/CE du Parlement européen et du Conseil du 12 juillet 2002 concernant le traitement des données à caractère personnel et la protection de la vie privée dans le secteur des communications électroniques (directive vie privée et communications électroniques) et le règlement EU 2016/79 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE. Si l'on ne peut parler d'un regroupement des attributs de la propriété que sont l'**usus**, le **fructus** et l'**abusus**, il n'empêche que, en tant que générateur de données qui nous sont personnelles, nous avons droit à poser un consentement éclairé quant à l'utilisation de celles-ci, à être informé sur leur traitement et, le cas échéant, à s'y opposer, à y avoir accès ainsi qu'à les rectifier, voire les effacer ; Strowel, « Les mutations des droits de propriété intellectuelle sous l'effet numérique », p. 129.*

Mais serait-il opportun d'aménager un droit de propriété sur les données personnelles ? De nouveau, la science économique apporte des éléments de réponse. La création de droits de propriété sur « l'information » est justifiée en présence de phénomènes de « passager clandestin »[104]. De quoi parle-t-on ? Il s'agit d'une déficience du marché qui surgit lorsqu'un individu peut profiter d'une « information » – par exemple, une invention industrielle, un procédé technique, une œuvre artistique – sans avoir participé à son financement. Ainsi en est-il d'une entreprise qui reproduit l'innovation d'une concurrente sans avoir dépensé le moindre euro en recherche et développement[105]. En pareil cas, les économistes prédisent un investissement suboptimal en information[106]. Deux solutions, non exclusives, sont envisageables pour inciter les agents économiques à la production d'informations : les subsides et les droits de propriété.

De toute évidence, les données personnelles ne constituent pas des informations sujettes à de telles contraintes d'incitations. Comme le CO_2, les données personnelles sont un sous-produit, un produit dérivé, un accessoire de notre existence. Elles ne soulèvent donc aucun problème de production suboptimale. Au vrai, les données sont produites, et aujourd'hui extraites, en abondance. Le problème est en vérité inverse : y-a-il aujourd'hui sur production, extraction et circulation de données personnelles, nécessitant qu'y soient introduites des restrictions ? Nous revenons sur ce point à un stade ultérieur.

2. Une rémunération des données personnelles ?

Toute aussi inepte est la suggestion selon laquelle le droit doit aménager un régime permettant aux utilisateurs d'être indemnisés pour les données personnelles qu'ils fournissent à des tiers[107]. Cette idée, populaire, relève davantage de considérations de justice sociale, que d'efficacité économique. Simplissime, sa trame narrative emprunte à la théorie du complot : les grandes plateformes du numérique s'enrichiraient sur le dos de leurs utilisateurs, en monétisant leurs données personnelles à leur insu.

En vérité, l'idée d'une rémunération des utilisateurs en contrepartie de leurs données personnelles n'est économiquement admissible que si (i) l'activité de production de données dégage un coût marginal de production ; et (ii) ce coût est supérieur au bénéfice marginal retiré du service presté par le tiers. En pareil cas, le jeu du marché libre devrait conduire à l'émergence d'opérations de vente des données personnelles, dans lesquelles l'utilisateur est l'offreur et le fournisseur de services est l'acheteur.

Ni à un titre, ni à l'autre, cette condition ne semble toutefois satisfaite. Premièrement, comme nous l'avons vu, la production de données personnelles n'entraîne *a priori* aucun coût économique pour les utilisateurs.

Deuxièmement, les utilisateurs retirent des avantages marginaux de l'utilisation de services de plateformes. Généralement, ces avantages tiennent au fait que les plateformes réduisent les frictions liées aux transactions économiques et favorisent les échanges sur les marchés[108]. Plus spécifiquement, les plateformes réduisent les coûts de transaction dans de nombreux secteurs, qu'il s'agisse de la mobilité (par exemple, le covoiturage), des communications électroniques (par exemple la communication vidéo), des informations personnalisées (par exemple, les réseaux sociaux), du commerce (par exemple, le commerce en ligne) ou encore des loisirs (par exemple, les jeux en ligne).

Si, à l'heure actuelle, les utilisateurs transfèrent leurs données *« gratuitement » aux plateformes, c'est sans doute que le « prix virtuel »*[109] payé reste inférieur aux bénéfices que les utilisateurs retirent de l'utilisation du service.[110] Ce modèle économique, appelé « two-way platform », n'a rien de spoliateur.[111] Il n'y a là que le résultat du jeu libre du marché. Du reste, les ressources spoliées ne sont pas si significatives qu'il n'y parait.

---

*104 Ce problème est probablement plus connu sous son appellation anglophone de « free rider problem » ; **voy**. M. Battaglini, S. Nunnari, T. R. Palfrey, « The Free Rider Problem: A Dynamic Analysis », 25 mars 2012, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=2028261.*

105 Investopedia, v° « Free Rider Problem », disponible sur www.investopedia.com/terms/f/free_rider_problem.asp.

*106 M.-C. Villeval, « quand le marché ne suffit plus : biens publics et coopération conditionnelle », in **Idées** économiques **et sociales**, n°161, 2010/3, disponible sur www.cairn.info/revue-idees-economiques-et-sociales-2010-3-page-6.htm.*

*107 C'est la position soutenue par Nicholas Economides qui a déclaré que « users should be allowed to sell their data to the company » ; N. Economides, « Commentary: Facebook Can't Be Trusted. It's Time to Regulate It », in **Fortune**, 22 mars 2018, disponible sur www.fortune.com/2018/03/22/facebook-cambridge-analytica-data-privacy-scandal/.*

*108 **Voy**. D. S. Evans, R. Schmalensee, **Matchmakers: The New Economics of Multisided Platforms**, Harvard Business Review Press, Boston, 2016 ; **voy**. également D. Slocum, « 5 Questions With David S. Evans And Richard Schmalensee on Matchmaking », in **Forbes**, 25 mai 2016, disponible sur www.forbes.com/sites/berlinschoolofcreativeleadership/2016/05/25/5-questions-with-david-s-evans-and-richard-schmalensee-on-matchmaking/#6cbef4b54a67.*

109 P. Belleflamme, « Modèle économique des données : une relation complexe entre offre et demande », in Enjeux numériques, n°2, juin 2018, disponible sur www.researchgate.net/publication/327057597_Modeles_economiques_des_donnees_une_relation_complexe_entre_demande_et_offre?enrichId=rgreq-5c4b23ef0686d5798cf2823ec070f181-XXX&enrichSource=Y292ZXJQYWdlOzMyNzA1NzU5NztBUzo2NjAyNjk3MDk2NzY1NDVAMTUzNDQzMTk3NDMxOQ%3D%3D&el=1_x_2&_esc=publicationCoverPdf.

110 J. Laitenberger, « empowering and protecting European citizens in an evolving media landscape », Keynote speech at the 2018 Jevons Colloquium on "Future Perspectives on Media Markets", Rome, 22 mai 2018, disponible sur www.ec.europa.eu/competition/speeches/text/sp2018_08_en.pdf.

111 Hashai, « Platform End Users as Free 'Data Labor' ».

Dans son troisième rapport trimestriel pour 2018, Facebook révélait générer un revenu moyen par utilisateur annuel de 8,82 $ en Europe[112]... On est bien loin de la manne financière dont parlent parfois les médias lorsqu'ils évoquent notre patrimoine numérique individuel.

Enfin, toute obligation réglementaire de rétrocession des revenus générés par les plateformes au moyen des données personnelles risque, ensemble avec les coûts de transaction qu'elle induit, d'inciter les plateformes gratuites à renoncer aux modèles de monétisation de type publicité (ad-based), et à migrer vers des modèles alternatifs de type abonnement (subscription-based). Cette conjecture est celle que formule le cabinet Ernst & Young, qui estime que le business model basé sur l'accès gratuit en échange des données devrait progressivement perdre de son importance[113]. Il n'est pas certain qu'une telle évolution œuvre à l'amélioration du bien-être du consommateur.

3. TROISIÈME MYTHE : « LES RÉGLEMENTATIONS SUR LA PROTECTION DES DONNÉES PERSONNELLES FAVORISENT LES GRANDES ENTREPRISES ET NUISENT AUX PETITES ENTREPRISES »

L'adoption de règlementations spécifiques protectrices des données personnelles est dans l'air du temps. En Europe, le RGDP entrait en vigueur le 25 mai 2018. Aux Etats-Unis, la Californie vient d'adopter en juin 2018 le California Consumer Privacy Act (CCPA) qui entrera en vigueur le 1er janvier 2020.

Les justifications économiques sous-tendant les réglementations sur la protection des données personnelles sont bien connues. Le respect de la vie privée est un « bien public » soumis à un problème d'action collective[114]: bien qu'il soit dans l'intérêt collectif de protéger la vie privée, les décisions individuelles des agents économiques n'aboutissent pas à un niveau socialement optimal de protection de la vie privée. Ceci, à son tour, tient au fait qu'en négation de leurs préférences exprimées[115], les individus sous-estiment systématiquement les coûts de leurs décisions sur leur vie privée et celle des tiers[116].

Face à ces préoccupations légitimes, d'aucuns - dont Mark Zuckerberg[117] - se sont émus des coûts de « compliance » élevés imposés par ces réglementations. Ils mettent en garde contre un risque de concentration oligopolistique accrue et de disparition des jeunes pousses innovantes. Ben Thompson, spécialiste des technologies de l'information, spéculait récemment en traçant un parallèle interpellant avec le secteur pharmaceutique, où les coûts des essais cliniques imposés par la réglementation ont élevé la concentration et les barrières à l'entrée.

A l'heure actuelle, quelques anecdotes tendent à indiquer que la réglementation a précipité non seulement la sortie de petites entreprises[118] mais aussi de grandes firmes[119]. Toutefois, il est encore trop tôt pour évaluer empiriquement l'effet du RGDP, faute d'application systématique de ses dispositions[120].

---

112 Voy. Facebook Q3 2018 Results, disponible sur www.investor.fb.com/investor-events/event-details/2018/Facebook-Q3-2018-Earnings/default.aspx.

113 EY, « The Big Data Backclash », 2013, disponible sur www.ey.com/Publication/vwLUAssets/EY-The-Big-Data-Backlash/$FILE/EY-The-Big-Data-Backlash.pdf.

*114 J. A. T. Fairfield, C. Engel, « Privacy as a public good », in **Duke Law Journal**, Vol. 65, n°3, décembre 2015, disponible sur www.scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3824&context=dlj.*

*115 À cet égard, il est intéressant de relever que, aux Etats-Unis, à la question de savoir s'ils avaient perdu le contrôle sur l'utilisation faite de leurs données personnelles, 91 % des personnes interrogées ont répondu par l'affirmative. Bien plus, 66 % se dise prêt à soutenir une plus grande régulation allant dans le sens de la protection de leur vie privée. L. Rainie, « Americans' complicated feelings about social media in an era of privacy concerns », in **Pew Research Center**, 27 mars 2018, disponible sur www.pewresearch.org/fact-tank/2018/03/27/americans-complicated-feelings-about-social-media-in-an-era-of-privacy-concerns/ ; Cette étude a été identifiée par J. Jia, G. Zhe Jin, L. Wagman, « The Short-Run Effects of GDPR on Technology Venture Investment », 5 novembre 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3278912 (ci-après, J. Jia, G. Zhe Jin, L. Wagman, « The Short-Run Effects of GDPR »).*

*116 Voy. S. Barth, M. D.T. de Jong, « The privacy paradox: investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review », **Elsevier**, 2017, disponible sur www.reader.elsevier.com/reader/sd/pii/S0736585317302022?token=61D75EE62B074419C50D0AEA248A099DFBAB7B490272E9EB390943B40438862801 40FA8B79590B6E1F3C249A0E9E5A26.*

*117 « A lot of times regulation by definition puts in place rules that a company that is larger, that has resources like ours, can easily comply with but that might be more difficult for a smaller startup », citation de Mark Zuckerberg telle que reprise par Bloomberg ; L. Bershidsky, « Europe's Privacy Rules Are Having Unintended Consequences », in **Bloomberg**, 14 novembre 2018, disponible sur www.bloomberg.com/opinion/articles/2018-11-14/facebook-and-google-aren-t-hurt-by-gdpr-but-smaller-firms-are ; Voy. Également S. Scechner, N. Kostov, « Google and Facebook Likely to Benefit From Europe's Privacy Crackdown », in **The Wall Street Journal**, 23 avril 2018, disponible sur www.wsj.com/articles/how-europes-new-privacy-rules-favor-google-and-facebook-1524536324.*

*118 Ainsi en est-il, par exemple, d'Instapaper, qui avait décidé dans un premier temps de se retirer de notre marché avant de récemment redevenir disponible. Certains jeux vidéo en ligne ont également préféré bloquer les joueurs européens plutôt que d'investir dans la compliance ; D. Castro, M. McLaughlin, « Why the GDPR Will Make Your Online Experience Worse », in **Fortune**, 23 mai 2018, disponible sur www.fortune.com/2018/05/23/gdpr-compliant-privacy-facebook-google-analytics-policy-deadline/ (ci-après, Castro et McLaughlin, « Why the GDPR ? »).*

119 Facebook, à titre d'exemple, a décidé de fermer son service de reconnaissance faciale par défaut en Europe afin d'éviter les coûts liés à la compliance ; Castro et McLaughlin, « Why the GDPR ? ».

120 Néanmoins, Jian Jia, Ginger Zhe Jin et Liad Wagman ont récemment mis en exergue que, comparativement à leurs équivalents américains, ces entreprises européennes de ce secteur ont reçu moins de financement depuis l'entrée en vigueur du RGPD "Our findings suggest a $3.38 million decrease in the aggregate dollars raised by EU ventures per state per crude industry category per week, a 17.6% reduction in the number of weekly venture deals, and a 39.6% decrease in the amount raised in an average deal following the rollout of GDPR" ; J. Jia, G. Zhe Jin, L. Wagman, « The Short-Run Effects of GDPR ».

A ce stade, seule la théorie économique est en mesure d'apporter des éclairages utiles. Que nous dit-elle d'utile ? De manière générale, le degré de concentration d'une industrie dépend sinon exclusivement du moins principalement de la structure des coûts de production. Or, les coûts de compliance introduits par le RGPD sont essentiellement liés *:*

1. à la réalisation « d'évaluations d'impact » sur la protection des données personnelles et de la vie privée au gré des projets de traitements de données envisagés, et notamment les traitements à grande échelle ;

2. au financement d'un délégué à la protection des données (et de collaborateurs) ; et

3. aux activités de notifications des violations aux autorités compétentes et aux personnes concernées.

De fait, nous ne sommes pas en présence de coûts invariants ou fixes, qui seraient de nature à avantager les grandes firmes.[121]Au contraire, les coûts de compliance risquent d'être directement corrélés à l'échelle des données traitées par les entreprises concernées, de sorte que le RGDP semble a priori neutre vis-à-vis de la concentration industrielle. Du reste, certaines obligations - comme la nomination d'un délégué - sont formellement inapplicables aux firmes de moins de 250 employés et donc aux jeunes pousses.

En revanche, le RGDP est de nature à augmenter les coûts de transaction, notamment par le biais de la distinction qu'il trace entre « contrôleurs » et « processeurs » de données. Dans certains secteurs, comme la publicité en ligne, les grandes plateformes et annonceurs ont traditionnellement délégué certaines de leurs activités à des intermédiaires techniques (dits « ad tech firms ») qui agissent pour leur compte en qualité de « processeurs ».

Les obligations de vérification qu'impose le RGDP aux contrôleurs sur les pratiques de leurs processeurs pourraient le cas échéant conduire à une augmentation des coûts de transaction, susceptible d'alimenter des stratégies d'intégration verticale, éliminant les intermédiaires opérant entre les plateformes et les annonceurs. C'est l'illustration du choix coasien bien connu entre la firme et le marché[122].

En bref, et contrairement aux propos de Mark Zuckerberg, il n'est pas évident d'associer des effets concentratifs aux réglementations protectrices de données personnelles.

CONCLUSION

En guise de conclusion, nous souhaiterions substituer une nouvelle analogie à celle discutée en exorde de cet essai : les données ne sont pas comparables au pétrole, elles sont comparables au capital déposé sur un compte bancaire.

Quand un agent économique place son argent auprès d'une banque, il crée des externalités positives vis-à-vis des tiers. Son capital peut être mobilisé pour satisfaire les besoins de financement d'autres agents économiques, sans qu'il ne subisse d'externalités négatives (dans une hypothèse excluant la faillite de la banque). Il peut même être rémunéré pour avoir confié son argent à une banque, via le paiement d'intérêts, qui peuvent être capitalisés.

Tout, évidemment, n'est pas optimal dans le fonctionnement de la désintermédiation bancaire. Loin s'en faut. Le dixième anniversaire de la crise financière nous rappelle avec acuité la nécessité d'une régulation étroite de certaines activités de marché ainsi que l'encadrement obligatoire de l'intervention de l'Etat lors de l'octroi de facilités de crédit aux professionnels comme aux particuliers. Ceci étant, personne n'a demandé l'abandon réglementaire du système du compte bancaire.

Une même logique devrait s'appliquer aux données personnelles. Elles sont comparables à l'argent placé sur un compte en banque : il est possible d'en faire profiter les tiers (externalités positives), tout en retirant des bénéfices marginaux, sous la forme d'effets de réseaux (en matière bancaire, les instruments de paiements). Et de toute évidence, la réglementation est à sa place lorsqu'il s'agit de garantir la sécurité, l'intégrité et la confidentialité des données, comme cela est le cas des obligations bancaires applicables à la garantie des dépôts...

---

*121 Une étude récente estimait que le coût de compliance n'est pas un **one-time cost** ; Castro, « Why the GDPR ».*

*122 R. H. Coase, « the nature of the firm », in **Economica**, Vol. 4, 1937, disponible sur www.onlinelibrary.wiley.com/toc/14680335/1937/4/16.*

VIII. LES BIG DATA SONT UN ELDORADO

Emmanuelle Duquenne

Le volume de données exponentiel et l'analyse de ces données en temps réel recèlent un pouvoir de création de connaissance et de valeur sans précédent. Leur utilisation promet des gains de croissance, de productivité, d'efficacité et de bien-être à travers tous les secteurs de l'économie et de la société.

Transversale par nature, l'exploitation des données crée de la valeur dans toute la chaine de production de nos entreprises et génère des économies d'échelle.

Cette évolution touche tous les secteurs. L'innovation dans le commerce, l'industrie, la santé, l'administration publique, l'énergie et le transport augure de retours positifs sur le plan social, environnemental et économique. A ce titre, ces secteurs sont en bonne position pour figurer parmi les premiers à capturer ces opportunités et à les transformer en nouveaux marchés et en services innovants pour leurs clients et leurs usagers. Les entreprises qui parviendront à en tirer avantage pourront diversifier leurs produits et leurs services et espérer franchir les frontières traditionnelles de leurs activités.

QUELQUES CHIFFRES SUR « L'ÉCONOMIE DE LA DONNÉE »

Selon une étude de l'OCDE publiée en 2015[123], le volume des données produites est évalué à 8 zettabytes (8 milliards de milliard de gigabytes), soit, chaque jour, 167000 fois l'information contenue dans la bibliothèque du Congrès américain. Certaines estimations considèrent que ce volume sera multiplié par 40 d'ici 2020.

Il est démontré que les entreprises pratiquant l'innovation par les données ont vu leur taux de productivité augmenter de 5 à 10% plus vite que les autres.

Les innovations par les données contribuent aussi au bien-être des citoyens même si cet impact est plus difficilement quantifiable en l'absence de véritables transactions. Les nombreuses applications qui font aujourd'hui partie de notre quotidien en témoignent largement et on mesure rarement le confort que celles-ci nous ont apporté et nous apporteront encore.

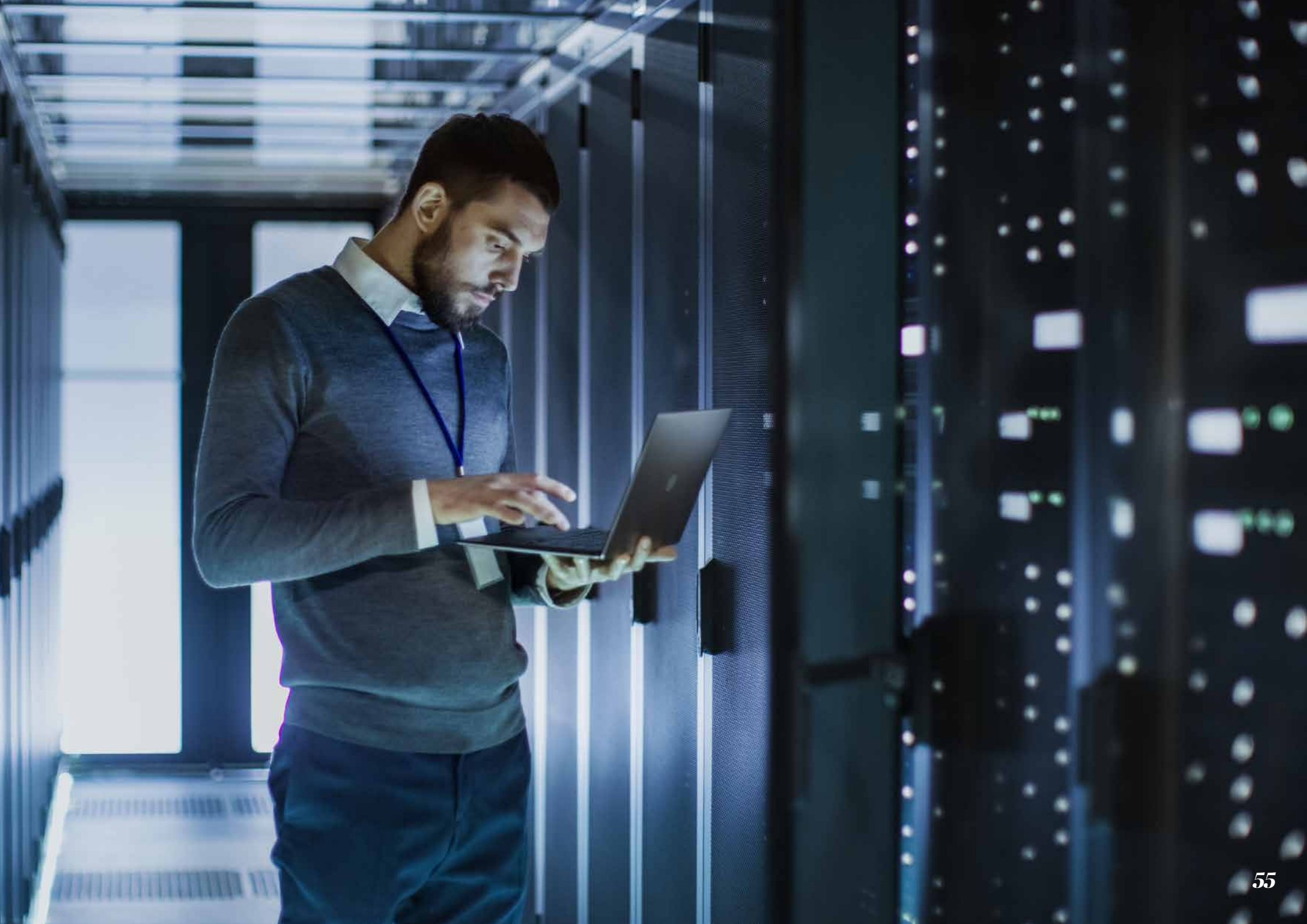
Le **marché européen des données ouvertes publiques** est un élément clé de l'économie des données dont la valeur devrait passer de 52 milliards d'euros en 2018 à 194 milliards en 2030. Selon McKinsey Global Institute, l'utilisation optimale de l'analyse de données par les 23 plus importants gouvernements européens permettrait de réduire les coûts administratifs de 15 à 20%, créant l'équivalent de 150 à 300 milliards d'euros d'économies[124]. Autant d'investissements qui trouveraient sans peine des projets essentiels susceptibles de faire la différence pour les citoyens.

L'internet des objets se développe principalement au sein de l'industrie, loin des applications plus marginales mais extrêmement populaires dans le domaine des maisons connectées. Ce marché en croissance de plus de 20% chaque année devrait bientôt supplanter les smartphones en tant que plus important consommateur du trafic internet. Nous n'en sommes pourtant qu'au début, avec le passage des projets pilotes aux applications rentables pour les entreprises. Une large part des milliards de capteurs qui seront déployés, par exemple, dans la sécurité, les soins de santé, l'environnement, l'énergie et les transports sont embarqués dans l'infrastructure urbaine. **Au niveau mondial, le marché des villes intelligentes devrait atteindre 400 milliards d'euros d'ici 2020, technologies et services confondus**. Abritant environ 60% de la population mondiale, les villes hébergeront une part proportionnelle des objets connectés.

Le revenu des entreprises de données européennes était évalué à 60 milliards en 2016 et pourrait atteindre 106 milliards d'euros en 2020 avec un taux de croissance de 15, 7 % par an.

L'économie des données dans son ensemble, tous secteurs confondus, était estimée à 285 milliards d'euros en 2015, soit 1,94% du PIB européen.[125] Avec un taux de croissance annuel d'environ 5% et accompagné de mesures politiques et législatives favorables aux investissements, cette évaluation pourrait monter à 739 milliards d'ici 2020, soit 4% du PIB européen.

123 OCDE, Data-Driven Innovation - Big Data for Growth and Well-Being, Editions OCDE, 2015
124 McKinsey Global Institute, Big data: The next frontier for innovation, competition, and productivity, mai 2011, p.17
125 IDC, Open Evidence, European Data Market – Final Report, Commission européenne, 2 mai 2017, p.201



L'EUROPE ET LE MONDE COMME CIBLES POUR NOS ENTREPRISES

Les comportements d'achat et les modes de consommation ont changé et nous exigeons aujourd'hui une expérience intégrée entre les différents canaux de communication et de vente de nos enseignes et de nos partenaires commerciaux. Que ce soit dans le commerce de détail ou dans l'industrie, une transformation profonde s'opère grâce aux données dans toute la chaîne de valeur, de la logistique à la relation avec les clients. Le rythme de développement de nouveaux marchés s'est extraordinairement accéléré. Bill Gross, le patron d'un incubateur aux Etats-Unis a calculé qu'il avait fallu 62 ans à la voiture pour s'imposer à 50 millions de consommateurs contre 19 jours pour Pokemon Go. Les données traversent l'ensemble des processus de nos entreprises et constituent le carburant essentiel à leur transformation dans tous les domaines, de l'excellence opérationnelle à atteindre au développement d'une relation client véritablement personnalisée.

C'est un défi majeur mais aussi une **formidable opportunité pour nos petites et moyennes entreprises** qui sont le moteur de notre économie. Elles représentent en effet plus de 99% de nos entreprises et fournissent 70% des emplois en Belgique. Elles sont 56% en Flandre, 11% à Bruxelles et 28% en Wallonie. Leur transformation grâce à l'exploitation des données et à la digitalisation en général est fondamentale pour notre économie et s'inscrit dans un contexte de croissance exponentiel du commerce en ligne.

La Belgique se classe aujourd'hui dans le top cinq des pays européens dans ce domaine. Une entreprise sur cinq dispose d'un magasin en ligne et **les dépenses en ligne atteignaient 5,29 milliards d'euros au premier semestre 2018, soit une augmentation de 8% par rapport à 2017**[126]. L'e-commerce représente ainsi 18% de parts de marché. 74% des utilisateurs d'internet et 63% de la population belge font des achats en ligne et on constate une augmentation importante de la régularité de ces achats. Le montant moyen dépensé mensuellement en ligne diminue légèrement en 2017 en passant de 199 € à 191 € mais cette diminution est compensée par l'accroissement de la fréquence de ces achats[127]. Les belges achètent principalement en ligne pour des raisons pratiques, de gain de temps, de confiance et de satisfaction générale, le prix n'étant plus le principal motif. Cette croissance est visible dans tous les secteurs avec une augmentation encore plus nette pour le divertissement, le bien-être et la santé.

Cette tendance est largement menée par la Flandre qui a formulé sa nouvelle politique industrielle dès 2011. En 2015, une étude publiée pour le compte du gouvernement wallon, "Regards sur l´Économie Wallonne : Économie par le Numérique" démontrait que les parts de marché de l'industrie dans le PIB régional ne cessaient de décroître. Cette étude montrait également que le degré d'adoption des technologies digitales par l'industrie wallonne était dramatiquement bas[128].

Afin de **sensibiliser nos entreprises aux opportunités de croissance qu'offrent leur transformation digitale** et l'exploitation des données générées dans le cadre de leurs activités, cinquante-cinq milliards d'euros d'investissements publics et privés seront mobilisés au niveau européen pour stimuler l'innovation, mettre en place des hubs régionaux, favoriser les partenariats, financer les premières lignes de production de composants électroniques de la prochaine génération ou encore développer un cloud européen[129].

Avec les efforts consentis pour la diminution du coût du travail par le gouvernement fédéral ces dernières années, la transformation de l'industrie grâce à la digitalisation des processus opérationnels, à la robotisation et à la formation adaptée du personnel participe à la réduction des coûts de production, qui génère à son tour des investissements et de nouveaux emplois.

Ce cercle vertueux participe également à la réalisation de nos engagements écologiques grâce à une utilisation plus rationnelle de l'énergie et à une logistique plus intelligente.

DES TECHNOLOGIES AU SERVICE DU BIEN-ÊTRE ET D'UNE MÉDECINE PERSONNALISÉE

Ces dix dernières années, 167 millions d'euros ont été investis dans les start-ups « digital health » de l'écosystème belge et ce chiffre n'inclut pas les subsides et les différentes formes de prêts qui ont aussi été accordés[130].

Le secteur de la santé est un des plus gros producteurs de données susceptibles d'impacter positivement le système et l'économie en général. Compte tenu de l'augmentation de l'espérance de vie, les données sont essentielles dans ce secteur afin d'améliorer les thérapies et les pratiques à tous les niveaux, au bénéfice des citoyens, des prestataires de soin et du système de santé.

En matière de prévention, les données ont montré leur efficacité grâce à une meilleure détermination des causes comportementales et environnementales des problèmes de santé. De nouvelles techniques d'imagerie structurelles telles que la résonnance magnétique fonctionnelle permettront notamment la détection précoce de maladies neuro-dégénératives comme Alzheimer ou Parkinson[131]. Dans de nombreux cas, l'amélioration de la prévention permet d'éviter des traitements plus lourds et plus longs. Des désagréments et des coûts inutiles sont ainsi épargnés.

---

126 BeCommerce, Nouveau record dans l'e-commerce belge : déjà 5,29 milliards d'euros de dépenses en ligne rien que pour le premier semestre 2018, Communiqué de presse, 21 septembre 2018

127 Comeos, E-Commerce Belgium 2017, p.16

128 Digital Transformation Monitor, Belgium: "Made Different", Commission européenne, mai 2017, p.3

129 Commission européenne, Commission sets out path to digitise European industry, Communiqué de presse, 19 avril 2016

130 Godard S., Les start-ups e-santé se heurtent à la complexité du système de soins belge, L'Echo, 28 mars 2018

131 OCDE, Data-Driven Innovation - Big Data for Growth and Well-Being, Editions OCDE, 2015, p.352

Les diagnostics peuvent être affinés grâce à l'exploitation des données de santé. Les médecins peuvent par exemple utiliser les données des patients pour évaluer leur performance par rapport aux meilleures pratiques. Les données peuvent fournir des indications pour améliorer les protocoles de soins ainsi que la coordination entre différentes spécialités.

Les futurs modèles de soins seront plus orientés sur le patient et sur son environnement domestique et social afin d'améliorer son bien-être général. Des soins personnalisés multidimensionnels prenant en considération la culture, le mode de vie, l'historique de soins ou encore le profil génétique du patient permettront d'apporter des réponses véritablement adaptées aux problèmes de santé.

Les technologies permettent déjà une **gestion plus facile et plus autonome de notre propre santé**, grâce par exemple à la prise de rendez-vous en ligne ou aux applications visant à encourager le sport et les modes de vie plus sains. Elles autorisent la prescription de traitements personnalisés, à la maison comme à l'hôpital. Les médecins peuvent, par exemple, être alertés quand les paramètres de santé d'un patient semblent anormaux. La surveillance en temps réel des paramètres de santé du patient ouvre de nouvelles perspectives pour le suivi et l'adaptation des traitements de même que pour la sensibilisation et la prévention à l'égard du patient. Des pratiques de « santé mobile » telles que le « nudging » qui aident le patient à prendre de meilleures décisions et à rester engagé dans son traitement ont démontré des améliorations mesurables en termes de sécurité et de qualité des soins.[132] Un large éventail d'appareils sont utilisés dans ce cadre : smartphone, tablette, GPS, équipement de soins ou de surveillance à distance.

Des produits et des services innovants se développent pour répondre aux problèmes liés au vieillissement et apporter des solutions à la dépendance pour les personnes âgées afin qu'elles puissent vivre plus longtemps chez elles. Des systèmes de capteurs peuvent fournir de précieuses informations sur des problèmes de santé physique ou mentale naissant, détecter des chutes et des problèmes de mobilité ou de socialisation, améliorer l'indépendance des personnes souffrant de problèmes de mémoire, d'organisation ou de réalisation de petites tâches de la vie quotidienne.

Les données permettent encore d'améliorer la gestion de la qualité, de la sécurité et de la performance du système de santé. On assiste au développement de système de santé auto-apprenant, auparavant réactif et concentré sur la maladie et désormais proactif, préventif et axé sur la qualité de vie. Les statistiques en temps réel offrent des projections précises des futurs besoins de soins de santé de la population afin d'améliorer la définition des politiques publiques et l'allocation des ressources. L'accès aux soins en est facilité et le temps d'attente diminué. La surveillance de la sécurité du matériel médical et pharmaceutique peut être gérée de façon plus sûre. L'efficacité des systèmes de santé, de la recherche clinique, des traitements et de leurs coûts d'un pays à l'autre peut être comparée. Plus largement, les données offrent enfin l'opportunité de s'attaquer aux défis de santé à l'échelle globale et d'améliorer la détection précoce des épidémies et de leur propagation[133].

DES DONNÉES POUR RESTAURER LA CONFIANCE POLITIQUE

Au sein du secteur public également, les données sont un moteur pour l'innovation, la croissance économique, la transparence et la participation citoyenne.[134] Elles ouvrent d'innombrables opportunités pour répondre aux défis climatiques et écologiques et pour assurer des services plus efficaces, centrés sur le bien-être des citoyens.

Si le service public est un important utilisateur de données, il est aussi une des principales sources de données dont l'ouverture et l'utilisation peuvent générer des bénéfices à travers tous les secteurs de l'économie, pour les institutions publiques elles-mêmes et pour les citoyens. Selon une étude européenne, **les avantages économiques cumulés découlant de la mise à disposition des données publiques s'élèveraient à 40 milliards d'euros par an** dans l'Union européenne.

Dans un contexte de contraintes budgétaires et de rationalisation, l'échange généralisé des données au sein et entre les institutions publiques offre des opportunités de décloisonnement, de réduction des coûts, de simplification et d'amélioration des performances. L'usage des données publiques permet d'augmenter et de préciser la production de statistiques dans tous les domaines et ainsi, d'**objectiver l'action publique.** La création de plateformes peut faciliter l'accès aux services publics mais aussi être à la source de nouvelles initiatives publiques ou citoyennes. L'analyse prédictive permet en effet de mieux identifier les besoins et d'y répondre par des services innovants et ciblés. Les investissements en infrastructures, la gestion opérationnelle et la maintenance peuvent gagner en efficacité grâce aux objets connectés et à l'engagement des citoyens rapportant via des applications l'existence de dépôts de déchets clandestins, de dégradations diverses, de nids de poule ou encore d'éclairages publics défectueux. Les applications sont nombreuses également dans le domaine de la sécurité, de la gestion des incidents ou des opérations de secours.

Les données sont aussi susceptibles de **restaurer la confiance envers les gouvernements et les institutions** grâce à une utilisation qui favorise l'ouverture, la transparence et la redevabilité de ces derniers vis-à-vis de leurs engagements. L'efficacité de l'action publique peut être objectivée, mesurée et rendue accessible au plus grand nombre. Cette accessibilité accrue peut aussi répondre aux revendications des citoyens d'être davantage impliqués dans la vie politique et sociale, par exemple au travers de processus participatifs et de consultations publiques.

132 OCDE, Data-Driven Innovation - Big Data for Growth and Well-Being, Editions OCDE, 2015, p.349

133 Gesundheit Österreich Forschungs- und Planungs GmbH, Study on Big Data in Public Health, Telemedine and Healthcare – Final report, Commission européenne, décembre 2016, p.46

134 OCDE, Data-Driven Innovation - Big Data for Growth and Well-Being, Editions OCDE, 2015, p.436

Au plan international, les applications se développent pour répondre aux **urgences humanitaires et aux défis du développement**. Des solutions se profilent également pour remplir nos engagements environnementaux et climatiques et offrent des opportunités économiques gigantesques comme on le voit avec le développement des smart cities.

L'INTELLIGENCE ARTIFICIELLE AU SERVICE DE LA PRODUCTION ÉNERGÉTIQUE

Dans le cadre de sa stratégie budgétaire à long terme pour la période 2021-2027, la Commission européenne investira 42 milliards d'euros dans les réseaux transeuropéens d'infrastructures, dont 8,7 milliards dans le secteur de l'énergie (30,6 milliards iront au transport et 3 milliards au numérique). Cela témoigne de la volonté d'accentuer la dimension environnementale de ses investissements, avec pour ambition de consacrer 60% du budget européen aux objectifs climatiques[135].

Dans la foulée des **engagements européens de réduire d'au moins 40% les émissions de gaz à effet de serre et de porter à 27% la part des énergies renouvelables d'ici à 2030**, ces dernières continueront leur forte croissance dans les prochaines années. Cette diversification des sources énergétiques constitue un défi majeur pour les infrastructures existantes et leur système de gestion inadapté à l'irrégularité de ces nouvelles sources d'énergie. Face à ces incertitudes, l'industrie se tourne vers les données et l'intelligence artificielle afin d'améliorer ses prévisions[136].

Bien que le secteur des infrastructures soit traditionnellement peu enclin à se digitaliser rapidement en raison de ses méthodes de travail conservatrices axées sur la sécurité et la réduction des risques, de son manque d'attractivité pour les profils de spécialistes des données et de la complexité de ses systèmes opérationnels, il s'agit d'un **secteur fertile pour les start-up où les acquisitions se multiplient à toute vitesse.**

Les objets connectés déployés autour des réseaux énergétiques dits « intelligents » tels que les « smart grids » et les « smart meters » transforment l'ensemble de la chaine de valeur de l'industrie de la production jusqu'à la relation client et permettent de répondre aux grands défis énergétiques. Ces technologies et leurs applications mobiles s'intègrent progressivement aux maisons intelligentes, aux immeubles connectés, aux nombreuses infrastructures publiques et aux « smart cities ». Elles génèrent ainsi de larges volumes de données sur les tendances de consommation qui peuvent être exploités pour augmenter l'efficacité énergétique, prévoir et gérer les pics de consommation, intégrer des sources d'énergie renouvelable dans la production et réduire les pertes dans la distribution et le transport.

Les données couplées à l'intelligence artificielle permettent ainsi d'**équilibrer la production entre les énergies renouvelables par nature intermittentes et les énergies fossiles** en diminuant la production de ces dernières, et par conséquent les émissions de gaz à effet de serre, lors des pics de production des premières.

Le pouvoir de prédiction des données facilite aussi la gestion et la maintenance des infrastructures, améliore le confort des immeubles et diminue les risques d'investissement dans l'efficacité énergétique.

Il s'agit d'un autre cercle vertueux qui, selon certaines estimations, augmente la profitabilité de l'industrie de 20 à 30% tout en participant à la réalisation de la stratégie européenne pour une économie neutre pour le climat en 2050. Enfin, ces défis écologiques et les opportunités économiques qui en découlent grâce à l'exploitation des données sont à l'origine de la création de nouveaux « emplois verts », porteurs de sens pour les citoyens.

UNE MOBILITÉ DURABLE GRÂCE À LA GÉOLOCALISATION

L'exploitation des données a un impact économique et social profond dans le secteur du transport et de la logistique. On estime que les données de géolocalisation sont susceptibles de générer 500 milliards de dollars de valeur dans le monde en économies de temps et de carburant. On évalue par ailleurs que 380 mégatonnes de CO2 seront épargnées grâce à ces technologies d'ici 2020. Rien qu'en Europe, un gain d'efficacité de 10% dans ce secteur pourrait en réduire les coûts de 100 milliards d'euros.

Malgré ces perspectives prometteuses, **peu d'entreprises ont déjà intégré les données en tant qu'élément créateur de valeur dans leurs processus et leurs services**. Comme dans bien d'autres secteurs, leur utilisation peut impacter positivement les opérations et l'expérience client et être à l'origine de la création de nouveaux modèles économiques. Ces innovations sont également fondamentales pour le développement de l'e-commerce dont le succès repose largement sur l'efficacité de la logistique et de la livraison. Avec des activités de fret supposées augmenter de 40% en 2030 et de 80% en 2050 par rapport à 2005, la transformation du secteur de la mobilité et de la logistique est cruciale.

Les **services émergents** grâce à l'analyse des données dans ce secteur sont innombrables et se retrouvent tant dans la gestion des infrastructures que dans l'évolution de nos modes de transport: autoroutes intelligentes, infrastructures ferroviaires proactives, ports et aéroports intelligents, mobilité urbaine intégrée, logistique partagée, assurance « pay as you drive », voitures, vélos et parkings partagés, véhicules connectés et dispositifs de sécurité et d'aide à la conduite, les exemples ne manquent pas.

135 Commission européenne, EU Budget: Commission proposes increased funding to invest in connecting Europeans with high-performance infrastructure, Communiqué de presse, 6 juin 2018
136 BDO, Why big data, AI and renewables are the perfect M&A storm, octobre 2017, p.3

Outre ces nombreuses opportunités économiques, **l'adoption de ces technologies participe à la fluidité et à la sécurité du trafic** ainsi qu'au désengorgement de nos villes. Elles améliorent notre bien-être a[137]u quotidien et nous permettent de mieux gérer notre temps en optimalisant nos trajets, nos horaires et nos modes de transport.

Au volant, l'**analyse continuelle des données de notre véhicule et du trafic avoisinant augmentent notre sécurité**, prévient les accidents et diminue le nombre de tués sur nos routes en identifiant et en résorbant les points noirs grâce à l'analyse prédictive.

La transformation de ce secteur comme de celui de l'énergie est un élément central pour développement des smart cities axées sur la **qualité de vie, le respect de l'environnement et le développement durable**. Dans ce cadre, la part belle est faite à l'économie collaborative génératrice de revenus complémentaires pour qui met en commun sa voiture ou encore son emplacement de parking, par exemple. Outre l'utilisation responsable des ressources énergétiques, une étude menée en 2014 montre que la mobilité partagée dans les transports urbains pourrait répondre à la demande de mobilité dans une mégalopole comme Singapour avec seulement 20% des voitures en circulation. Selon une estimation plus conservatrice de l'International Transport Forum, ce volume pourrait être réduit de moitié.

LES DONNÉES, LE PRIX À PAYER ?

Au-delà de ces opportunités, l'ubiquité du réseau et la multiplication des objets connectés permettent à chacun de tirer largement profit de ces technologies. Les services de communication (email, messagerie instantanée, réseaux sociaux...), de recherche (moteur de recherche, wikis, comparateurs de prix...), de cartographie ou de divertissement (vidéos, musique, jeux, coachs sportifs...) sont payés par la publicité et nous facilitent grandement la vie.[138]

Les technologies émergentes telles que la réalité augmentée, la médecine personnalisée ou encore les assistants virtuels pour n'en citer que quelques-unes apporteront à leur tour de nouvelles pratiques qui viendront transformer notre quotidien.[139]

Loin de nous soumettre et de nous déshumaniser comme on l'entend souvent, les progrès technologiques sont largement et librement adoptés parce qu'ils comportent bien plus d'avantages. Les exemples contraires sont d'ailleurs légion. Ces bouleversements relativement récents imposent bien sûr la vigilance. La technologie doit évoluer au service de la société dans le respect des valeurs et de la tradition libérales. Des questions éthiques ou aussi cruciales que la propriété des données se posent encore et devront retenir notre attention. Toutes les révolutions industrielles ont amené une évolution des modes de vie et une réforme profonde de l'éducation. Des changements que nous redoutons, dont nous ne voyons parfois que les dangers potentiels et auxquels nous sommes tentés de résister. La somme des évolutions positives auxquelles nous assistons et dont la vitesse peut couper le souffle doit néanmoins nous inciter à l'optimisme et à saisir à bras ouverts les opportunités extraordinaires qu'elles nous offrent.

137 R. Castinera A. Metzger, The TransformingTransport project - Mobility meets big data, 7th Transport Research Arena (TRA 2018), Vienna, Austria, 16-19 avril 2018, p.7
138 J. Bughin, The Web's €100 billion surplus, McKinsey Global Institute, janvier 2011
139 O. Cann, Voici les 10 principales technologies émergentes de 2018, World Economic Forum, 12 octobre 2018

IX. QUELLES DONNÉES PEUVENT-ELLES ÊTRE PARTAGÉES PUBLIQUEMENT ?

Arnaud Lombardo

La « donnée » est généralement considérée comme le pétrole du XXIème siècle. Pour donner un ordre de grandeur, la valeur de l'économie européenne fondée sur les données représentait 300 milliards d'euros en 2016. Si les mesures législatives et politiques adéquates sont mises en place, cette valeur pourrait atteindre jusqu'à 739 milliards d'euros d'ici à 2020, soit 4% du PIB de l'UE.[140] Il s'agit donc d'un enjeu majeur. Mais, comme le précisent Jérôme de Cooman et Nicolas Petit dans le chapitre VII de la présente étude, ces chiffres, impressionnants dans l'absolu, restent néanmoins modestes par rapport à d'autres secteurs tels que le secteur pétrolier.

En Belgique comme partout dans l'Union européenne, le secteur public détient d'importantes quantités de données, allant des données géographiques et météorologiques aux données éducatives, économiques et sociales. Aujourd'hui, les nouvelles technologies permettent non seulement d'augmenter massivement le volume de données que peut récolter le secteur public mais également d'en faciliter le classement, l'agrégation, le croisement et donc l'utilisation et en faire un outil d'aide à la prise de décision.

OUVRIR LES DONNÉES PUBLIQUES POUR PERMETTRE LEUR RÉUTILISATION

L'accès à toutes ces données publiques permet aux citoyens, aux médias, à la société civile, aux entreprises et aux organisations d'acquérir de nouvelles connaissances, de mettre au point des innovations qui améliorent la qualité de vie de chacun et de contribuer à une meilleure diffusion de l'information à l'échelle d'un pays et entre les États.[141] Ces informations (cartes, images par satellite, législation et jurisprudence, statistiques, registre de société, population, brevets, etc.) constituent un potentiel de croissance important dans la mesure où d'autres acteurs (entreprises (du secteur du numérique ou non), associations, etc.) sont à même de les utiliser pour créer de nouveaux services et de nouveaux contenus au service des citoyens ou des autres organisations, contribuant ainsi au développement économique.

C'est le principe de la réutilisation des données publiques que l'Union européenne, par diverses mesures et règlement, entend encourager. Dans ce débat, Andrus Ansip, le Commissaire chargé du Marché numérique européen déclare :

> *« Le marché unique numérique prend rapidement forme ; mais, sans données disponibles, il nous sera impossible de tirer le meilleur parti de l'intelligence artificielle, du calcul à haute performance et d'autres avancées technologiques. Ces technologies peuvent contribuer à améliorer les soins de santé et l'éducation ainsi que les réseaux de transport et à réaliser des économies d'énergie : voilà tout l'enjeu de l'utilisation intelligente des données »*

C'est pourquoi la directive ISP[142] a instauré le droit à la réutilisation (faire usage des données à des fins autres que l'objectif initial pour lequel elles ont été créées ou recueillies), ce qui signifie que toutes les informations du secteur public qui peuvent être largement diffusées en application des réglementations nationales régissant l'accès aux informations devraient, en principe, être réutilisables.

Il y a d'ailleurs déjà de nombreux domaines où les données publiques, lorsqu'elles sont ouvertes, créent de la valeur : transparence et contrôle démocratique, performance énergétique, mobilité, santé, etc.

140 Source : Commission européenne

141 Charte G8 pour l'Ouverture des Données Publiques – 8 juin 2013

142 La Directive 2003/98/CE sur la réutilisation des informations du secteur public

DES DONNÉES TROP PEU EXPLOITÉES

Mais d'une manière générale, et comme l'a démontré une étude récente menée en Wallonie auprès des pouvoirs locaux[143], les données publiques demeurent (trop) peu exploitées par les pouvoirs publics.

Pour passer le cap et entrer de plein pied dans la dynamique d'open data, il faut que les décideurs comprennent que la donnée est un carburant, un ferment de création de valeur et que la publication des données publiques participe à cette création de valeur.

Si la taille de l'organisation ou de la commune détermine dans une large mesure le niveau de maturité vis-à-vis de la data, c'est globalement une absence de la culture de la donnée et un manque d'adhésion aux valeurs de l'open data qui caractérisent aujourd'hui une grande partie des pouvoirs locaux ou des services publics. En effet, le nouveau paradigme créé par la digitalisation de la société induit que les acteurs publics n'ont plus le monopole de la gestion du territoire et de l'espace public, aujourd'hui partagée avec des acteurs privés. Cela signifie qu'ils n'ont par ailleurs plus le monopole sur la gestion des données collectée sur ce territoire, données qu'ils sont réticents à partager de peur d'en perdre le contrôle.

QUELLES DONNÉES OUVRIR ?

Pourtant, le mouvement est en route et les règlementations, au départ de l'Europe trace une trajectoire claire et des obligations tout aussi claires. Le principe de données ouvertes par défaut tel que défini au niveau européen suppose l'obligation de publier proactivement, sans attendre une quelconque sollicitation, toutes les données publiques produites par les pouvoirs publics, et qui ne sont pas des données à caractère personnel ou des données à caractère industriel ou commercial protégées.

L'absence de sollicitation nécessaire est importante car le contraire induirait un biais sélectif dans le choix des données publiées ou non. En effet, la nature même de l'innovation et des micro-services développés par l'économie numérique, c'est qu'ils vont provenir d'idées et de projets improbables. Il est impossible d'anticiper comment et où la valeur de l'ouverture et de la réutilisation des données sera créée et donc de prédire en amont les données qui ont un intérêt de celles qui sont superflues.

L'ouverture des données concerne les données statiques (cadastres, cartes, relevés, etc.) mais également les données dynamiques. Celles-ci, également dites « données en temps réel » font l'objet de demandes fortes des acteurs privés. Ces données dynamiques concernent généralement l'enregistrement continu de flux dans des secteurs stratégiques (informations météorologiques, mobilité ou énergie par exemple). Il s'agit d'une avancée importante pour le développement de services nouveaux mais également sensibles, notamment parce que ces données sont souvent considérées par les organismes qui les génèrent comme étant leur propriété et comme faisant partie intégrante du service public qu'il leur incombe d'offrir.

C'est par exemple le cas du secteur du transport au sein duquel les opérateurs publics de transport pourraient considérer que la création d'un service applicatif de mobilité, basé sur les données de mobilité dynamiques des trains, trams, bus opérés est de leur seule prérogative. Pourtant, dans l'optique du développement accéléré d'une mobilité intermodale, la création d'une application différente et spécifique pour chaque opérateur de transport (les TEC en Wallonie, De Lijn en Flandre, la STIB à Bruxelles ainsi que la SNCB pour prendre l'exemple belge) a peu de sens si chacune d'elle ne donne aux usagers que les informations dont il dispose sur son propre réseau. A l'inverse, si chaque opérateur ouvre les données en temps réel de son réseau de transport et que ces données peuvent être agrégées et partagées, il est alors possible de voir émerger des solutions applicatives qui offrent aux usagers de réelles solutions intermodales combinant les modes doux et l'utilisation de la voiture personnelle à l'ensemble des possibilités offertes par les sociétés publiques de transport voir à des solutions plus locales de véhicules partagées ou de vélos connectés en libre-service.

C'est d'ailleurs un des objectifs poursuivis par la Commission européenne dans le dernier train de mesures destinées à accroître la disponibilité des données : « *la réutilisation des informations du secteur public s'étend désormais aux données détenues par des entreprises publiques du secteur des transports et du secteur des services collectifs* »[144].

Pour encadrer et faciliter cette ouverture et cet échange des données au départ des acteurs publics soumis à cette obligation, les différents niveaux de pouvoirs doivent se fixer comme objectif le développement de plateformes open data, par exemples thématiques (mobilité, énergie, etc.) mais surtout interopérables les unes avec les autres (pour permettre le croisement de données provenant de thématiques différentes mais utiles au même objet) et facilement accessibles aux réutilisateurs (notamment les startups), au moyen d'interfaces par exemple.

N'Y A-T-IL PAS UN DANGER À PARTAGER LES DONNÉES ?

Parmi les risques liés à l'ouverture des données publiques figure bien souvent la protection de la vie privée. Il est important de préciser que l'obligation d'ouverture des données ne concernent pas les données à caractère personnel et qu'en respect du Règlement Général sur la Protection des Données (RGPD), seules les données agrégées et anonymisées sont appelées à être mise à disposition sur les plateformes open data. L'anonymisation des données permet de garantir la protection de la vie privée. Il est vrai néanmoins que l'anonymisation n'est jamais garantie de manière absolue.

143 Baromètre 2018 de FuturoCité : « Culture de la donnée dans les villes et communes wallonnes »

144 Les initiatives présentées aujourd'hui complètent le cadre pour la libre circulation des données à caractère non personnel dans l'UE

L'ÉTAT NE DEVRAIT-IL PAS VENDRE CES DONNÉES ?

L'open data n'a-t-il finalement pas une valeur que l'État pourrait monétiser ? Les villes et communes qui génèrent des données ne pourraient-elles les vendre au secteur public ? La question peut légitiment se poser si l'on considère que les pouvoirs publics disposent de données utiles au développement d'activités économiques et attractives pour le secteur privé.

Il faut d'abord rappeler que le cadre légal qui va être mis en œuvre et qui découle de la directive européenne est celui de l'utilisation par défaut des *open data* publiques qui seront gratuites :

« Les nouvelles dispositions limitent les exceptions autorisant les organismes publics à prélever une redevance pour la réutilisation de leurs données supérieure aux coûts marginaux de diffusion des données »[145].

Il faut ensuite avoir conscience que les acteurs privés dominants (les géants du web, les GAFAM) collectent et génèrent déjà eux-mêmes un volume de données très importants et suffisant à leur activité. Si les données publiques pourraient leur être utiles afin de compléter leurs services, elles ne sont pas indispensables. Les données relatives à l'organisation des travaux publics ou des manifestations sur la voie publique pourraient par exemple être utiles à Waze, l'application de mobilité détenue par Google, pour augmenter la granularité et la précision des informations données aux utilisateurs mais force est de constater que l'application remporte déjà un large succès, sans que ces données ne soient encore ouvertes, voire même qu'au travers du signalement participatif et instantané, renseigne parfois des événements avant même que les autorités n'aient l'information.

Il y a donc peu de chances que ces acteurs importants, qui disposent des moyens nécessaires, achètent des gisements de données publiques. A l'inverse, pour les startups et petites entreprises innovantes qui ont besoin de ces données comme carburant pour la création d'activités et de services, l'achat de données publiques représenterait sans doute un coût trop important ou les limiteraient à seulement certains jeux de données, freinant leur capacité à innover.

En générant, collectant, structurant et partageant des données publiques, non seulement les pouvoirs publics permettent à un écosystème local de startups et de PME innovantes de se développer, mais produire des données publiques de référence relève également de l'enjeu de la souveraineté. En n'ouvrant pas les données publiques articulées autour de standards légitimes, les États s'exposent à ce que l'économie s'organise autour d'autres données et d'autres standards, édictés par les géants du Web et répondants à un intérêt différent de celui des pouvoirs publics. Il faut donc ouvrir ses données pour mieux se protéger des géants du numérique, même si cela peut paraitre contre-intuitif.

EST-IL OPPORTUN DE FAIRE DES PARTENARIATS PUBLICS/PRIVÉS (PPP) ?

INTÉRÊTS MUTUELS MAIS AUSSI DIVERGENTS

Si les données générées par les autorités publiques peuvent intéresser les acteurs privés, l'inverse est vrai également : les données privées (produites par des acteurs privés dans le cadre d'une activité privée) peuvent, elles aussi, intéresser les collectivités.

Cet intérêt mutuel a déjà été illustré ci-dessus avec Waze et, en effet, un peu partout en Europe et notamment en France, on assiste de plus en plus fréquemment à des échanges entre des collectivités locales et des acteurs privés de la mobilité. En combinant les données statiques d'une agglomération (travaux, modification de sens de circulation, manifestations...) avec les données dynamiques, temps réel, de circulation collectées par les utilisateurs de Waze, il est possible de générer un état des lieux dynamique du trafic, de produire différents tableaux de bord sur les conditions en temps réel. Les données sont mises à la disposition de toute une série d'acteurs : services techniques communaux, services d'affichage urbain, d'aménagement urbain, services d'urgence, acteurs commerciaux...

Les acteurs commerciaux dans le domaine de la donnée, dont les GAFAM, ont des intérêts qui répondent à d'autres finalités et règles que les pouvoirs publics. Les acteurs publics doivent aussi être garants de l'intérêt général. Si le partenariat se révèle bénéfique pour toutes les parties, il ne faut pas perdre de vue qu'aujourd'hui il, est sans doute déséquilibré car au profit de l'acteur privé, notamment en termes d'utilisation des données.

MAITRISE PUBLIQUE DE LA DONNÉE ET DES ALGORITHMES

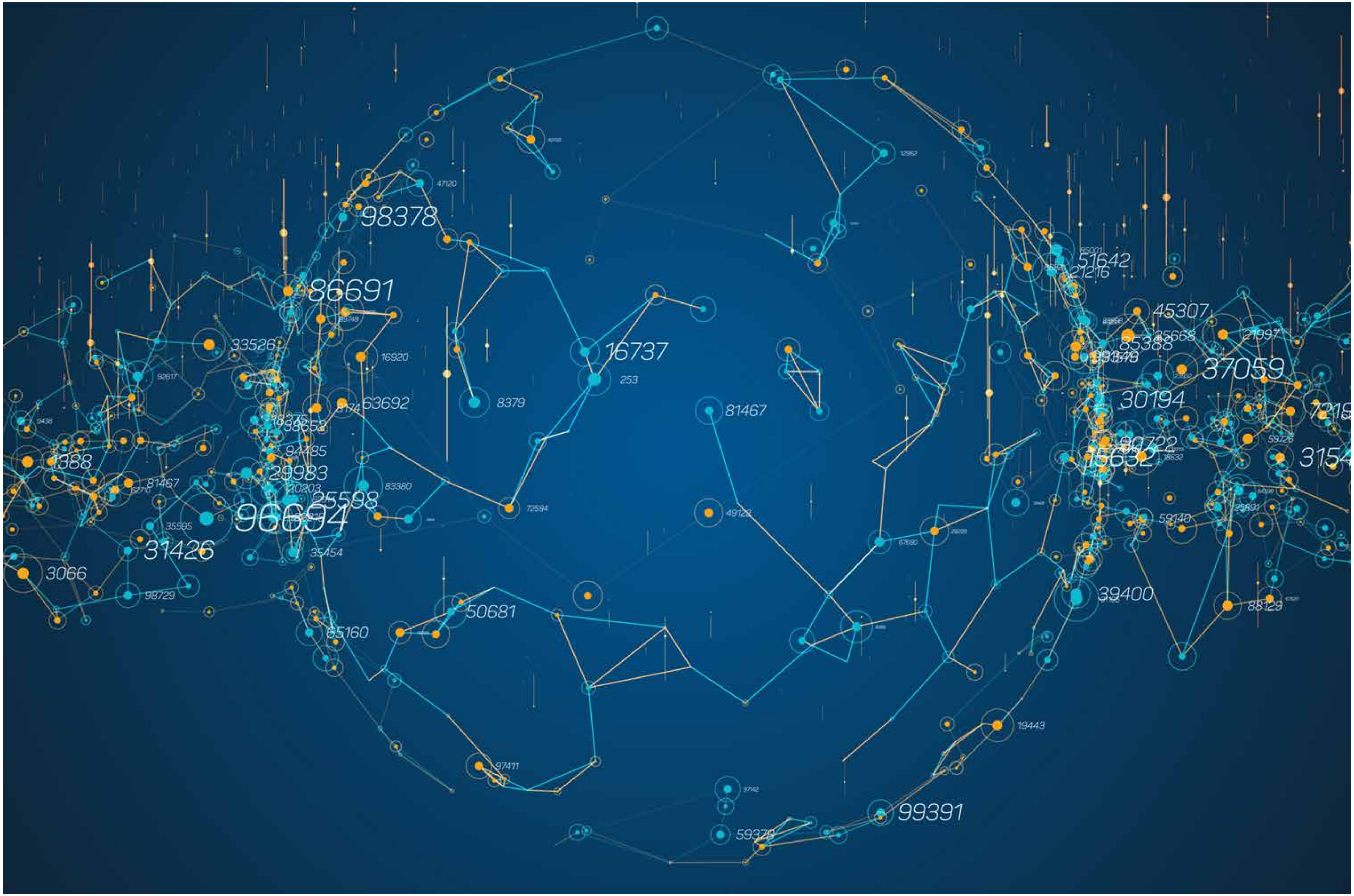
Dans l'intérêt collectif (des citoyens, du tissu économique local voire de la collectivité dans son ensemble), il est important de pouvoir rééquilibrer cette relation autour de la donnée entre acteurs publics et géants du web. Cela passe avant tout par l'acquisition d'une réelle culture de la donnée au sein des pouvoirs publics ainsi que d'une montée en gamme des compétences digitales au sein des administrations et entreprises publiques.

145 Proposition de directive du Parlement européen et du Conseil modifiant la directive 2003/98/CE concernant la réutilisation des informations du secteur public

Aujourd'hui, les données générées par un nouveau service dans le cadre de contrats entre un acteur public et un prestataire privé ne sont pas suffisamment valorisées. Des réels partenariats, basés sur un rapport de force rééquilibré devraient permettre aux pouvoirs publics de revendiquer les données générées sur leurs territoires au travers de prestations de ce type.

Dans le même ordre d'idée, lorsque l'algorithme d'un opérateur privé a un impact concret sur le quotidien des gens, ne peut-on considérer que cet algorithme doit être soumis à une dose de transparence afin que les autorités publiques soient en mesure d'en comprendre les mécanismes ? La gestion de la « cité », jadis monopole de l'État est aujourd'hui assurée en partie par des acteurs privés qui ont « préempté » le territoire grâce à leur avantage compétitif en termes de maitrise technologique.

Les partenariats entre ces géants du Net et l'État (au sens large) sont à même de générer une foule de nouveaux services utiles aux collectivités et aux citoyens mais, dans l'intérêt général, il est nécessaire de garantir la maitrise publique des données générées.


98378
86691
33526
16920
63692
16737
253
8379
81467
3066
31426
98729
65160
50681
97411
99391
19443
39400
88129
51642
45307
37059
30194

X. RECOMMANDATIONS

Corentin de Salle, Emmanuelle Duquenne & Arnaud Lombardo

I. SENSIBILISER AUX MENACES PESANT SUR LA VIE PRIVÉE ET AUX MOYENS D'Y FAIRE FACE

- Sensibiliser, dès l'école, les citoyens sur les **dangers d'une exposition excessive de sa vie privée** sur les réseaux sociaux, les messageries et sur le net.
- Sensibiliser, dès l'école, les citoyens sur tous les droits que leur offre la directive RGPD (droit de consultation, droit à l'oubli, etc.).
- Sensibiliser, dès l'école, les citoyens sur le fait qu'ils peuvent utiliser des **outils, moteurs de recherches et applications alternatives aux technologies actuelles** et qui ont pour point commun de respecter davantage leur vie privée et la confidentialité de leurs échanges. Par exemple :
  - Les moteurs de recherche Disconnect ou DuckDuck go.
  - La messagerie instantanée Signal
  - Le navigateur Tor Browser est le navigateur qui permet de protéger certaines informations identifiables
  - Etc.
- Insérer dans le Pacte d'Excellence la nécessité **d'enseigner des notions de statistiques et d'algorithmiques** dans les écoles afin de rendre les futurs citoyens capables de comprendre le fonctionnement d'un algorithme, de comprendre les finalités qu'il poursuit et les résultats qu'il recherche.
- Sensibiliser, dès l'école, les citoyens sur les **dangers de céder à la tentation du « guidage automatique ».** Les recommandations que formulent les algorithmes sont souvent pertinentes et utiles. Mais une chose est de les utiliser sur une base régulière, une autre est de s'en remettre à eux en s'imaginant qu'ils décrivent le réel, en estimant qu'ils nous connaîtraient mieux que nous nous connaissons.

II. PROTÉGER LA VIE PRIVÉE

- **Renforcer les moyens** matériels, humains et techniques de **l'Autorité de Protection des Données.**
- Veiller à ce que les **Cours et Tribunaux appliquent davantage les législations de protection de la vie privée et de protection des données personnelles.**
- Dans un contexte de menace terroriste, rester vigilant contre toute velléité de mettre en œuvre ou de réactivation de **programme de surveillance de masse** : de tels programmes sont attentatoires aux libertés et s'avèrent à l'expérience contreproductifs.
- Conclure un **véritable traité entre l'UE et les Etats-Unis** relativement à ce grand enjeu sociétal et juridique qu'est le **traitement des données à caractère personnel** plutôt que l'actuel et fragile accord entre deux visions et deux contextes trop différents.

III. DÉMOCRATISER LES ALGORITHMES

- Etablir une **charte des algorithmes** contenant certains principes tels que l'interdiction de limiter l'accès à l'information sur le net à certains internautes, l'interdiction de confiner des internautes dans de micro-communautés, l'obligation de maintenir une offre diversifiée, etc.
- Ouvrir, dans la foulée des travaux du professeur Frank Pasquale mais dans le respect des droits de propriété intellectuelle, une **réflexion sur** une exigence de **davantage de transparence et de responsabilité (accountability) des algorithmes** utilisés par les GAFAM et autres grandes entreprises privées afin de rééquilibrer la relation entre elles et leurs usagers. Aujourd'hui, la relation est profondément asymétrique : les GAFAM savent quasiment tous de leurs usagers et ceux-ci quasiment rien sur elles, leurs dirigeants et les algorithmes qu'elles utilisent.

IV. FAVORISER LE DÉVELOPPEMENT DE L'ÉCONOMIE DES BIG DATA

Jusqu'en 2013, la Belgique figurait encore parmi les 10 meilleurs pays exportateurs de services technologiques mais elle ne se démarque pas pour le développement de services basés sur les données. D'ici 2020, l'Europe prévoit de faire face à un manque d'environ 420.000 travailleurs dans le secteur des données, avec un nombre d'entreprises actives dans ce domaine passant de 255.000 en 2016 à 360.000 en 2020. Les leviers politiques à actionner pour saisir ces opportunités et rattraper notre retard sont nombreux et doivent se décliner à travers tout le spectre de l'action publique.

- **Pour une offre d'enseignement agile et adaptée aux besoins des entreprises du futur**
  - Sensibiliser dès le plus jeune âge à l'importance des matières scientifiques et à la valeur ajoutée de l'utilisation des données dans les domaines stratégiques et porteurs de sens pour l'avenir
  - Revaloriser l'esprit scientifique et critique, en décalage croissant avec d'autres formes de connaissances, à travers tous les programmes d'enseignement
  - Développer une offre de formation de pointe dans les sciences des données, l'intelligence artificielle, l'informatique quantique et les technologies émergentes
  - Développer les compétences personnelles granulaires susceptibles de faire la différence face aux machines telles que la dextérité ou les « soft skills » comme la créativité, la résolution de problème, la communication, la gestion des émotions ou la prise de parole en public
  - Encourager l'esprit d'entreprise par l'intégration de projets personnels à tous les niveaux d'enseignement
  - Favoriser le développement de laboratoires d'innovation ouverte (living labs) et de laboratoires de fabrication (fab labs) auprès des écoles et des universités
  - Mettre en place des incitants et des conditions propices au développement de relations fructueuses entre la recherche fondamentale et l'industrie
- **Pour des normes ouvertes en faveur du partage des données essentiel à la compétitivité**
  - Mettre en place des groupes d'experts pour la sensibilisation à l'adoption de standards ouverts techniques, sémantiques et organisationnels permettant d'assurer l'interopérabilité dans les domaines stratégiques
  - Définir des normes en matière de construction pour une gestion rationnelle de l'énergie
  - Définir des normes en matière de dispositifs de sécurité de série sur tous les véhicules automobiles
- **Pour des sources de données accessibles et sécurisées**
  - Ouvrir le débat sur la propriété des données en vue de la mise en place d'un système d'échange et de rétribution efficace
  - Renforcer la confiance de la société à l'égard du partage des données par une information claire sur les droits et les obligations qui s'y attachent
  - Soumettre les demandes de subside à l'obligation de partager les données des projets financés tout en veillant à la protection de nos intérêts économiques
  - Favoriser les partenariats public-privé en faveur des données ouvertes
- **Pour des financements diversifiés au bénéfice de l'innovation**
  - Maintenir la Belgique dans le haut du classement en matière de financement de la recherche-développement
  - Poursuivre la modernisation du droit des affaires et du droit des sociétés afin de favoriser les collaborations, les investissements, les appels publics de fonds et les fusions acquisitions des entreprises
  - Favoriser la multiplication des sources de financement et le partenariat public-privé pour diminuer les risques d'investissement et assurer un retour durable à toutes les parties prenantes sur les plans économiques, sociaux et environnementaux
  - Encourager les modes de financement alternatifs tels que la philanthropie et le crowdfunding
- **Pour un service public exemplaire en matière d'utilisation et de partage des données**
  - Recruter et retenir les talents pour assurer la transformation digitale rapide des administrations publiques
  - Mettre en place des groupes d'experts en vue de l'extension des sources de données publiques à valeur ajoutée existantes et du développement de nouvelles sources de qualité
  - Utiliser les technologies afin d'élaborer, de mettre en œuvre et d'assurer le suivi et l'ajustement rapide et flexible des politiques publiques, notamment en matière d'investissements
  - Décloisonner les services publics afin de permettre le partage des données, la rationalisation de leur fonctionnement et la simplification des procédures administratives

V. OUVRIR LES DONNÉES PUBLIQUES (OPEN DATA)

- **Publier proactivement, sans attendre une quelconque sollicitation, toutes les données publiques produites par les pouvoirs publics**, et qui ne sont pas des données à caractère personnel ou des données à caractère industriel ou commercial protégées.

- **Ouvrir davantage les données publiques pour permettre leur réutilisation** : l'accès à toutes ces données publiques permet aux citoyens, aux médias, à la société civile, aux entreprises et aux organisations d'acquérir de nouvelles connaissances, de mettre au point des innovations qui améliorent la qualité de vie de chacun et de contribuer à une meilleure diffusion de l'information à l'échelle d'un pays et entre les États.

- **Adopter des normes publiques en matière de création de bases de données facilitant l'interopérabilité des différents systèmes d'exploitation de données de l'administration publique et l'interconnectivité des bases de données entre elles et avec le secteur privé. Sur cette base, on pourrait, par exemple, créer un service applicatif de mobilité commun à tous les opérateurs publics de transport**, service basé sur les données de mobilité dynamiques des trains, trams, bus, etc. au lieu d'une application différente et spécifique pour chaque opérateur de transport (les TEC en Wallonie, De Lijn en Flandre, la STIB à Bruxelles ainsi que la SNCB pour prendre l'exemple belge) ou s'assurer du moins que chaque opérateur ouvre les données en temps réel de son réseau de transport et que ces données peuvent être agrégées et partagées.

- **Développer des plateformes open data communes aux différents niveaux de pouvoir** sur diverses thématiques (mobilité, énergie, etc.), interopérables les unes avec les autres (pour permettre le croisement de données provenant de thématiques différentes mais utiles au même objet) et facilement accessibles aux réutilisateurs (notamment les startups), au moyen d'interfaces par exemple.

- **Améliorer, dans la limite des technologies disponibles, la sécurisation des données lors de leur agrégation et leur anonymisation avant de les partager**

- **Faire des partenariats publics/privés (PPP) pour créer des plateformes open data thématiques (mobilité, énergie, etc.) et cela dans les deux sens** : certaines données générées par les autorités publiques intéressent les acteurs privés mais les données privées (produites par des acteurs privés dans le cadre d'une activité privée) peuvent, elles aussi, intéresser les collectivités. C'est le cas, par exemple, avec l'application Waze qui permet d'apporter en temps réel un état des lieux dynamique de trafic permettant de produire différents tableaux de bord sur les conditions en temps réel.

- **Garder la maitrise publique de la donnée et des algorithmes** : dans l'intérêt collectif (des citoyens, du tissu économique local voire de la collectivité dans son ensemble), il est important de pouvoir rééquilibrer la relation autour de la donnée entre acteurs publics et géants du web. Cela passe avant tout par l'acquisition d'une réelle culture de la donnée au sein des pouvoirs publics ainsi que d'une montée en gamme des compétences digitales au sein des administrations et entreprises publiques.

CONCLUSIONS

Corentin de Salle

Ces dernières années, le développement phénoménal des capacités de stockage et de la puissance de calculs des ordinateurs a donné naissance à la révolution numérique et au démarrage fulgurant auquel nous assistons aujourd'hui. Le principal aliment de cette économie, ce sont les données. Ces dernières ne valent pas grand-chose prises de manière isolée. Mais agrégées de manière massive, elles permettent aux algorithmes qui les traitent de dégager des corrélations qui sont autant d'opportunités pour agir et créer de la valeur. Ces nouvelles pratiques nécessitent une actualisation des normes protégeant la vie privée. **Dans le cadre de cette étude, nous avons abordé ces deux problématiques (comment libérer le potentiel de l'économie numérique et que faire pour protéger la vie privées) de manière complémentaire. Ces deux objectifs ne sont en rien antagonistes**.

Aujourd'hui, les Google, Apple, Facebook, Amazon et Microsoft détiennent ensemble près de 80% des informations personnelles numériques de l'humanité. A l'ère des big data et des algorithmes, il nous faut réinventer les outils normatifs protecteurs de la vie privée.

A cet égard, le Règlement Général de Protection des Données (RGPD) était attendu depuis longtemps. Car la directive de 1995 était devenue inadaptée à la réalité d'aujourd'hui. La conception universelle et égalitaire de la protection de la vie privée que consacre ce règlement explique pourquoi l'Europe va sans doute prendre le lead en la matière et imposer ses standards aux Etats-Unis et au reste du monde, car dès lors que tous les citoyens européens sont protégés par cette législation, aucune activité de traitement concernant ces citoyens ne peut être effectuée sans respecter le RGPD, même si les activités de traitement sont mises en œuvre en dehors du territoire de l'Union européenne.

En Belgique, il apparaît que les effectifs de l'Autorité de Protection de Données ne sont pas assez nombreux, formés et instrumentés pour mener correctement leurs missions, a fortiori depuis que ses missions ont été renforcées par le RGPD. Par ailleurs, les cours et tribunaux n'assurent pas une application satisfaisante de la législation en vigueur sur la protection de la vie privée et la protection des données privées.

La conjugaison des big data et des algorithmes sont à la base d'une accélération considérable du progrès dans quantité de domaines tels que la connaissance, les techniques médicales, la gouvernance, l'économie, la communication, la prospective, etc. C'est un réel eldorado économique.

Contrairement à une idée répandue, les citoyens ne possèdent pas un droit de « propriété » sur leurs données personnelles. Ils bénéficient certes de législations protectrices mais cela n'aurait pas de sens de monnayer ces données qui, en elles-mêmes, ne valent pas grand-chose. Si elles étaient payantes, quantité de services utilisés massivement aujourd'hui cesseraient d'être gratuits. De toute façon, il est rare que de prétendues données personnelles soient produites par une seule personne. Elles sont souvent coproduites.

Ce qui intéresse les grandes entreprises du web, ce n'est pas, à proprement parler, la vie privée des usagers car elles ne s'intéressent pas aux personnes mais à une myriade de données infra-personnelles que ces personnes produisent constamment et qui, agrégées, permettent de prédire le comportement d'achat. Par contre, ce qui pose problème et menace la notion même de sujet et la vie en société, c'est l'idéologie technicienne derrière les big data.

Le problème consiste à croire que les algorithmes nous diraient la vérité sur le réel alors qu'ils ne font que détecter, dans le champ inépuisable de la réalité, des corrélations sélectionnées précisément parce qu'elles offrent des opportunités pour agir.

Nous devons développer chez les citoyens un sens critique par rapport à ce que nous disent et nous conseillent les algorithmes. Ils sont souvent utiles et pertinents. Mais, ils ne sont ni infaillibles ni objectifs. Il faut apprendre à déceler les intentions qui président aux calculs qu'ils opèrent.

Aujourd'hui, deux personnes qui se connectent à internet et qui formulent des requêtes sur le même moteur de recherche recevront des résultats différents. Cela pose clairement un problème quant à notre droit à l'information. A l'heure de l'hyper-personnalisation de nos profils, de nos besoins et des offres et propositions qui nous sont adressées, nous devons exiger que soit préservé un « monde commun ». A travers, par exemple, une **charte des algorithmes** et une plus grande exigence de transparence. Les **algorithmes doivent être considérés non comme des calculateurs objectifs et incontestables mais comme des outils, voire des objets de consommation** en concurrence dont on devrait pouvoir débattre dans des forums d'usagers.

L'Etat a un rôle à jouer. En veillant au respect des règles protectrices de la vie privée et en adaptant ces dernières au besoin. En sanctionnant également efficacement ceux qui les violent. Mais aussi en facilitant l'économie des données. Notamment, en formant les travailleurs de demain qui seront amenés à travailler dans ce type d'économies.

Il doit aussi ouvrir ses propres données (open data), les anonymiser et les mettre à disposition gratuitement aux citoyens, entreprises et PME car elles peuvent constituer une véritable matière première pour la création de valeurs. Il doit passer des partenariats avec le privé mais en veillant à ce que cet échange ne soit pas asymétrique et en s'assurant de garder la maîtrise des données publiques et de ses algorithmes.

La révolution des big data est déstabilisatrice. Mais, elle ne doit pas être diabolisée. Elle contient en elle beaucoup de promesses. Elle est lourde également de menaces qu'il ne faut pas minimiser mais qui constituent aussi autant de défis pour garder et renforcer le contrôle de nos vies.

BIBLIOGRAPHIE

OUVRAGES

Cardon D., **A quoi rêvent les algorithmes. Nos vies à l'heure des big data**, Seuil, 2015

Docquir B., **Actualités du droit de la vie privée**, Bruxelles, Bruylant, 2008, pp. 7-9.

Dugain M. & Labbé Ch., **L'homme nu. La dictature invisible du numérique**, Plon, 2016

Floridi L., **The fourth revolution. How the infosphere is reshaping human reality**, Oxford University Press, 2014

Greenwald G., **Nulle part où se cacher. L'affaire Snowden par celui qui l'a dévoilée au monde**, Jean Claude CLattès, 2014

Rigaux F, **La protection de la vie privée et des autres biens de la personnalité**, Bruxelles, Bruylant, L.G.D.J., 1990.

Schmidt E. & Cohen J., **A nous d'écrire l'avenir. Comment les nouvelles technologies bouleversent le monde**, Denoël, 2013

ARTICLES SCIENTIFIQUES

Anderson, Ch., The End of Theory : the data deluge makes the scientific method obsolete, **Wired Magazine**, 2008 www.wired.com/2008/06/pb-theory/

Barth S., de Jong M. D.T., The privacy paradox: investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review, **Elsevier**, 2017, disponible sur www.sciencedirect.com/science/article/pii/S0736585317302022

Battaglini M., Nunnari S., Palfrey T.R., **The Free Rider Problem: A Dynamic Analysis**, 25 mars 2012, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=2028261.

Belleflamme P., Modèle économique des données : une relation complexe entre offre et demande, in **Enjeux numériques**, n°2, juin 2018

Cheffert, J.-M., Respect de la vie privée : quand les approches économiques et juridique se rejoignent, **Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet**, Larcier, 2018, pp. 505-524

Coase R.H., « the nature of the firm », in **Economica**, Vol. 4, 1937, disponible sur www.onlinelibrary.wiley.com/toc/14680335/1937/4/16.

Determann L. , « No One Owns Data », in **70Hastings Law Journal**, Research Paper No. 265, 23 février 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3123957

Evans D.S., Schmalensee R., **Matchmakers: The New Economics of Multisided Platforms**, Harvard Business Review Press, Boston, 2016

Fairfield J. A. T. & Engel C., Privacy as a public good, in **Duke Law Journal**, Vol. 65, n°3, décembre 2015, disponible sur www.scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3824&context=dlj.

Hashai N., **Platform End Users as Free 'Data Labor' – Re-distributing the Value Created in Double Sided Markets**, 9 février 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3121160

Hugenholtz P.-B., **Data Property: Unwelcome Guest in the House of IP**, 2017, disponible sur www.ivir.nl/publicaties/download/Data_property_Muenster.pdf

Jia J., Zhe Jin G., Wagman L., **The Short-Run Effects of GDPR on Technology Venture Investment,** 5 novembre 2018, disponible sur www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3278912

Laitenberger J., Empowering and protecting European citizens in an evolving media landscape, **Keynote speech at the 2018 Jevons Colloquium on "Future Perspectives on Media Markets**", Rome, 22 mai 2018, disponible sur www.ec.europa.eu/competition/speeches/text/sp2018_08_en.pdf.

Poullet Y., « La protection des données : entre libertés, droits subjectifs et intérêts légitimes », in **Liber Amicorum P. Martens**, Bruxelles, Larcier, 2007, p. 139.

Rainie L., « Americans' complicated feelings about social media in an era of privacy concerns », in **Pew Research Center**, 27 mars 2018, disponible sur www.pewresearch.org/fact-tank/2018/03/27/americans-complicated-feelings-about-social-media-in-an-era-of-privacy-concerns/ ;

Rouvroy A., Homo juridicus est-il soluble dans les données ?, **Droit, normes et libertés dans le cybermonde, Liber Amicorum Yves Poullet**, Larcier, 2018, pp. 415-442

Strowel A., Les mutations des droits de propriété intellectuelle sous l'effet du numérique (data ownership, text and data mining, hyperlinking), in **Les enjeux de l'innovation : quelles politiques ? Quelles gouvernances ?**, B. van Pottelsberghe et al. (Dir.), 22ème Congrès des économistes, Ed. Université Ouverte de la Fédération Wallonie-Bruxelles, Charleroi, 2017

Strowel A., Big Data and Data Appropriation in the EU, in T. Aplin (ed.), Research **Handbook on Intellectual Property and Digital Technologies**, Edward Elgar, 2018

Velu J. et Ergec R., « La convention européenne des droits de l'homme », **R.P.D.B.**, complément VII, Bruxelles, Bruylant, 1990, n°10 et n°100.

Villeval M.-C., Quand le marché ne suffit plus : biens publics et coopération conditionnelle, in **Idées économiques et sociales**, n°161, 2010/3, disponible sur www.cairn.info/revue-idees-economiques-et-sociales-2010-3-page-6.htm.

Waschmann P., « Le droit au secret de la vie privée », in F. Sudre, **Le droit au respect de la vie privée au sens de la convention européenne des droits de l'homme**, coll. Droit et Justice, n°63, Bruxelles, Bruylant, 2006, pp. 119-155.

Warren S. D. & Brandeis L. D., The right to Privacy, **Harvard Law Review**, Vol. 4, N°5, Dec.15, 1890

Yang P., **Miners vs. Merchants: How Global Trade Made Men Wealthy during the California Gold Rush**, 5 mars 2016, disponible sur www.flexport.com/blog/trade-merchants-rich-california-gold-rush/.

ARTICLES DE PRESSE

Bershidsky L., « Europe's Privacy Rules Are Having Unintended Consequences », in **Bloomberg**, 14 novembre 2018, disponible sur www.bloomberg.com/opinion/articles/2018-11-14/facebook-and-google-aren-t-hurt-by-gdpr-but-smaller-firms-are

Castro D., McLaughlin M., Why the GDPR Will Make Your Online Experience Worse, in **Fortune**, 23 mai 2018, disponible sur www.fortune.com/2018/05/23/gdpr-compliant-privacy-facebook-google-analytics-policy-deadline/

Economides N., Commentary: Facebook Can't Be Trusted. It's Time to Regulate It, in **Fortune**, 22 mars 2018, disponible sur www.fortune.com/2018/03/22/facebook-cambridge-analytica-data-privacy-scandal/.

Forbes, **Global 2000: the world's largest public companies**, disponible sur www.forbes.com/global2000/#374e3d34335d

Kompass, **Les données, le nouvel or noir de votre business (livre blanc)**, disponible sur www.prokcssblog.blob.core.windows.net/edito-fra/FRA/2018/01/livre_blanc_kompass_La-donnee-le-nouvel-or-noir-de-votre-business.pdf

RAPPORTS

Marie J.-F., « Data, le nouvel or noir », **NetApp**, s.d., disponible sur www.netapp.com/fr/company/news/press-releases/news-rel-20171120-365886.aspx ;

Mezinis D., « Les données personnelles, le nouvel or noir… », **Boursier.com**, 9 avril 2018, disponible sur www.boursier.com/actualites/macroeconomie/les-donnees-personnelles-le-nouvel-or-noir-761834.html.

Scechner S., Kostov N., Google and Facebook Likely to Benefit From Europe's Privacy Crackdown, in **The Wall Street Journal**, 23 avril 2018, disponible sur www.wsj.com/articles/how-europes-new-privacy-rules-favor-google-and-facebook-1524536324.

Slocum D., 5 Questions With David S. Evans And Richard Schmalensee on Matchmaking, in **Forbes**, 25 mai 2016, disponible sur www.forbes.com/sites/berlinschoolofcreativeleadership/2016/05/25/5-questions-with-david-s-evans-and-richard-schmalensee-on-matchmaking/#6cbef4b54a67.

Accenture, Digitizing Energy : Analytics-Powered Performance - Opportunities for oil and gas companies to improve business outcomes, 2013, 40 p.

BDO, Why big data, AI and renewables are the perfect M&A storm, octobre 2017, 8 p.

Castinera R. & Metzger A., The TransformingTransport project - Mobility meets big data, 7th Transport Research Arena (TRA 2018), Vienna, Austria, April 16-19, 2018

Digital Transformation Monitor, Belgium: "Made Different", Commission européenne, mai 2017, 8 p.

EY, « The Big Data Backclash », 2013, disponible sur www.ey.com/Publication/vwLUAssets/EY-The-Big-Data-Backlash/$FILE/EY-The-Big-Data-Backlash.pdf.

Facebook Q3 2018 Results, disponible sur www.investor.fb.com/investor-events/event-details/2018/Facebook-Q3-2018-Earnings/default.aspx.

FuturoCité, **Culture de la donnée dans les villes et communes wallonnes**, Baromètre 2018

Gesundheit Österreich Forschungs- und Planungs GmbH, Study on Big Data in Public Health, Telemedine and Healthcare – Final report, Commission européenne, décembre 2016, 117 p.

Comeos, E-Commerce Belgium 2017, 76 p.

IDC, Open Evidence, European Data Market – Final Report, Commission européenne, 2 mai 2017, 275 p.

IDC, « Revenues for Big Data and Business Analytics Solutions Forecast to Reach $260 Billion in 2022, Led by the Banking and Manufacturing Industries, According to IDC », 15 août 2018, disponible sur www.idc.com/getdoc.jsp?containerId=prUS44215218

Investopedia, v° « Free Rider Problem », disponible sur www.investopedia.com/terms/f/free_rider_problem.asp.

McKinsey Global Institute, The age of analytics: competing in a data-driven world, décembre 2016, 136 p.

McKinsey & Company, The Digital Utility: New challenges, capabilities, and opportunities, June 2018, 76 p.

OCDE, Exploring data-driven innovation as a new source of growth: Mapping the policy issues raised by "big data", in *Supporting Investment in Knowledge Capital, Growth and Innovation*, Editions OCDE, octobre 2013, 40 p.

OCDE, Science, technologie et innovation : Perspectives de l'OCDE 2018 (version abrégée) : **S'adapter aux bouleversements technologiques et sociétaux, Science, technologie et innovation : Perspectives de l'OCDE**, Editions OCDE, 2018, 54 p.

OCDE, **Data-Driven Innovation - Big Data for Growth and Well-Being**, Editions OCDE, 2015, 456 p.

Rapport de colloque du Sénat, **La vie privée des citoyens et la protection des données face aux nouvelles technologies : les enjeux**, Sénat de Belgique, 17 octobre 2016

Service public fédéral Economie, P.M.E., Classes moyennes et Energie, **Tableau de bord des PME et des entrepreneurs indépendants**, 29 novembre 2017, 79 p.

Villani C., **Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne**, 2018

Zampou, E., Milioti, C., Liapis, A., Rodrigalvarez, V., Flocke, F., Dimitrakopoulos, G. & Bravos, G., Big data analytics in e-commerce logistics: Findings from a systematic review and a case study, Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

LÉGISLATION

Règlement (EU) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), JO, L119/1, disponible sur www.eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32016R0679&from=FR

Editeur responsable : Olivier Chastel,
Président du Centre Jean Gol
Gestion et Action libérale asbl, 84-86
1060 Bruxelles

Retrouvez toutes nos études sur **cjg.be** *ou demandez-nous gratuitement un exemplaire par téléphone ou par mail*

Centre Jean Gol

Avenue de la Toison d'Or 84-86 1060 Bruxelles • ***02.500.50.40*** • ***info@cjg.be*** • ***facebook.com/centrejeangol*** • ***@CentreJeanGol***

www.cjg.be