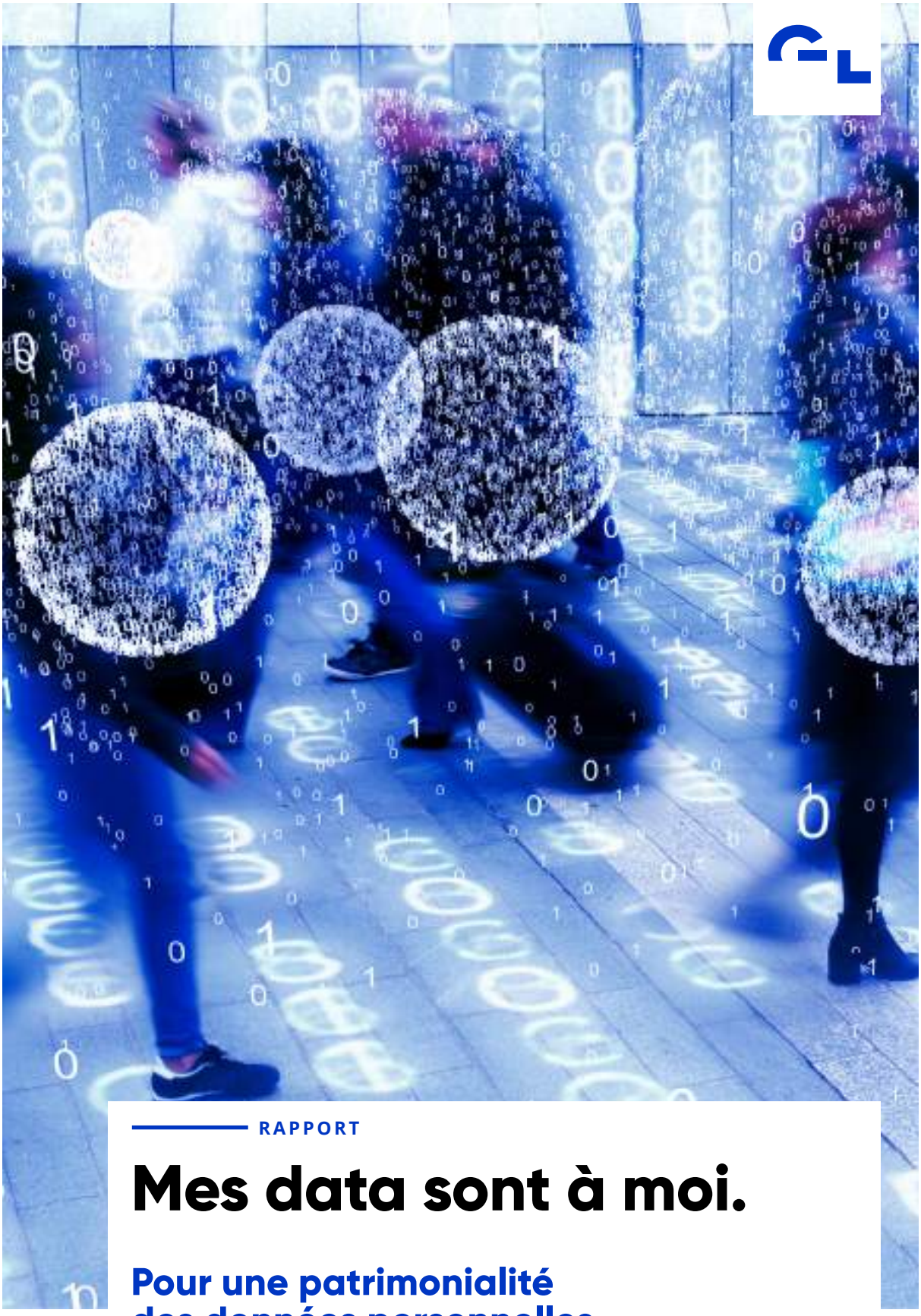
JANVIER 2018



RAPPORT

Mes data sont à moi.

Pour une patrimonialité des données personnelles.

Conception & réalisation : Laurent Darellis / Atelier 1
Impression : Mailedit
Janvier 2018

MES DATA SONT À MOI

PAR

ISABELLE LANDREAU

Avocat au Barreau de Paris, Docteur
en droit en propriété intellectuelle
et nouvelles technologies.

GÉRARD PELIKS

Ingénieur en cybersécurité,
expert en sécurité de l'information
et Président de CyberEdu.

NICOLAS BINCTIN

Professeur agrégé des facultés de droit,
professeur à l'Université de Poitiers.

VIRGINIE PEZ-PÉRARD

Enseignant-chercheur, maître
de conférences à l'Université Paris II
Panthéon-Assas, spécialiste des
questions de « privacy » et marketing.

SOUS LA DIRECTION DE

LUCAS LÉGER

Doctorant au CNAM, Directeur de
recherche du Think tank GenerationLibre.

Déclaration de responsabilité. Ce rapport est organisé en trois parties. Chaque auteur a contribué à une partie spécifique, dans son domaine de compétence. Son nom est indiqué en début de partie afin de distinguer le travail de chacun. GenerationLibre a piloté l'ensemble de ces contributions et mis en cohérence le texte final. Les auteurs ne sont pas réputés approuver l'ensemble des parties.

LE MOT DE GASPARD

« Droit fondamental »

Aujourd'hui, nos **données personnelles** sont dans la nature, *res nullius* appropriées et **revendues** par les grands acteurs du numérique.

Face à cette situation inique, chaque **école de pensée** tente d'imaginer une **réponse :** les sociaux-démocrates inventent des **droits et obligations,** les socialistes proposent des schémas de **taxe et redistribution,** les nationalistes imaginent une **souveraineté numérique…**

Il est important que les **libéraux** fassent également entendre leur voix en plaidant pour la **propriété privée** et la **rémunération.**

Proudhon considérait la propriété comme « la plus grande force révolutionnaire qui existe », en ce qu'elle confère à **l'individu** la souveraineté sur son domaine propre. Il n'y pas de maîtrise **sans possession.**

Ce **droit fondamental** doit aujourd'hui s'étendre aux **données,** prélude d'une véritable **propriété de soi sur soi.**

Gaspard Koenig

Président
Génération Libre

SOMMAIRE

Se repérer dans le rapport.

01

RÉSUMÉ DE L'ÉTUDE

Tous les jours, nous acceptons des dizaines de *cookies* sur nos ordinateurs et consentons à des conditions d'utilisation léonines qui nous dépossèdent de **nos données personnelles,** y compris les plus intimes. Les GAFA et autres plateformes se rémunèrent en grande partie par **la monétisation de ces données agrégées,** notamment via **la publicité.**

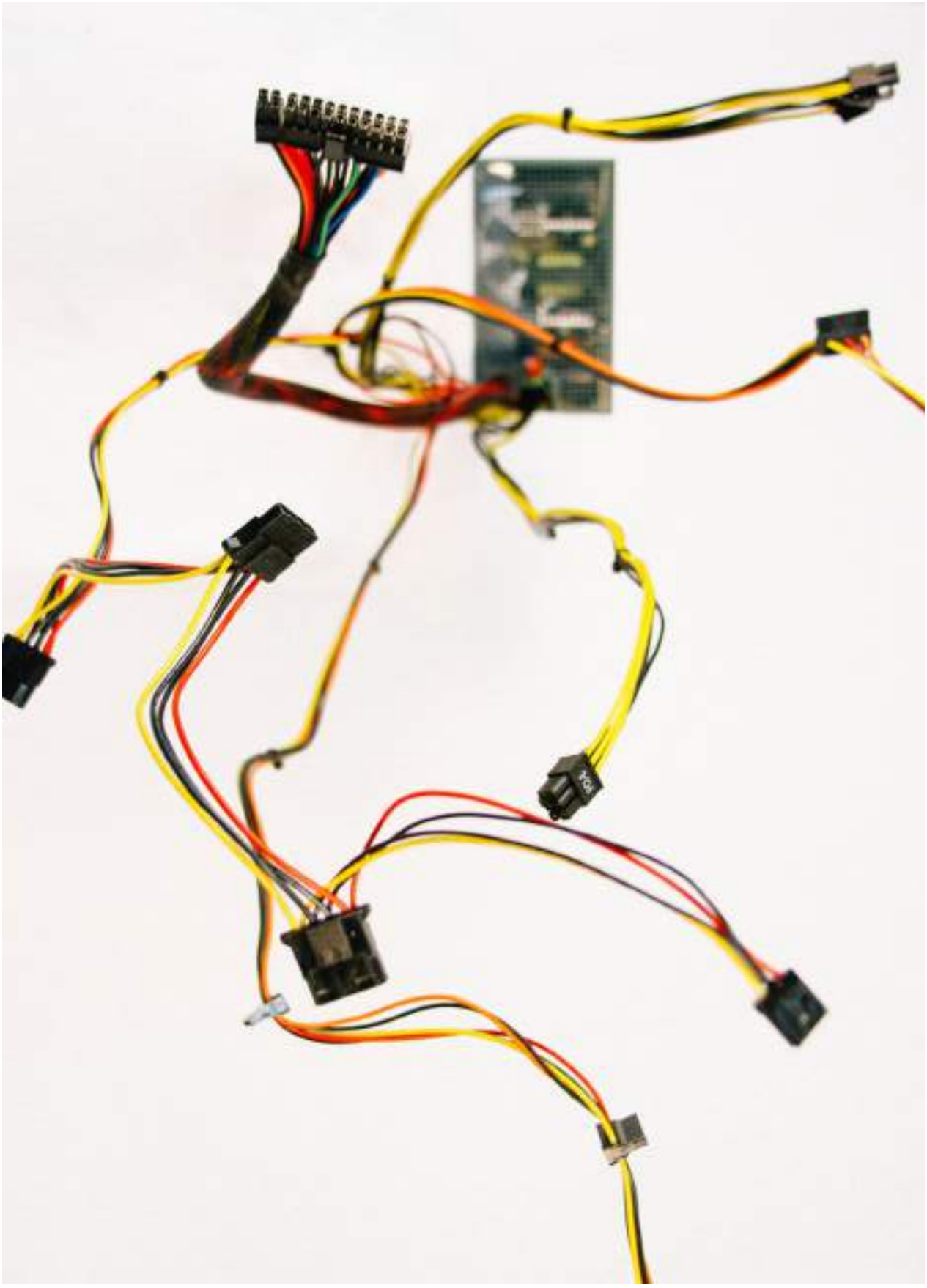
Pourtant, l'utilisateur ne retire aucune rémunération directe de la matière première qu'il fournit. Il se trouve *de facto* **exclu de la chaîne de valeur** de l'économie numérique et **prisonnier du ciblage publicitaire.** La gratuité des services masque un **pillage** en règle de nos données, c'est-à-dire de nous-mêmes.

Nous proposons dans cette étude d'instaurer une patrimonialité des données personnelles. De même que la révolution industrielle a rendu nécessaire le droit de propriété intellectuelle, la révolution numérique devrait créer un droit de propriété sur les données.

Cette innovation juridique viendrait bousculer le fonctionnement de l'écosystème numérique, en donnant aux utilisateurs-producteurs :

- La possibilité de **contractualiser** (éventuellement via des intermédiaires) l'usage de leurs données personnelles auprès des plateformes, afin de décider eux-mêmes de l'utilisation qu'ils souhaitent en faire.

- La possibilité de **monétiser** (ou non) ces données en fonction des



termes du **contrat** (vente, location...) préalablement conclu. N'est-il pas temps de se demander si Facebook, dont les profits dépassent les 4 milliards de dollars par trimestre, ne devrait pas nous payer pour l'utilisation de nos données ?

- La possibilité, à l'inverse, de **payer le prix** du service rendu par les plateformes sans leur céder nos données (le prix de la *privacy* ?).

Cette étude s'inscrit dans la logique développée aux Etats-Unis par **Jaron Lanier,** pionnier de la réalité virtuelle et auteur de *Who owns The Future ?*. Dans un papier de recherches récent intitulé *« Should We Treat Data as Labor? Moving Beyond 'Free' »*[1], Jaron Lanier propose, avec des universitaires de Stanford et de Columbia, que **les entreprises du Web nous rémunèrent** pour l'utilisation de nos données. Ce papier pose les principes du changement de paradigme que GenerationLibre appelle de ses voeux : rendre à **l'individu** producteur de données **la propriété de ses données** personnelles.

Seul **le droit de propriété** permettra de garantir une **maîtrise réelle** de nos données. Seule la création d'**un marché des data** pourra **rééquilibrer les rapports de pouvoir** entre les plateformes et leurs utilisateurs, en dotant chacun d'entre nous d'un **véritable capital.**

Avec l'appui d'un collectif composé d'universitaires (juristes, ingénieur, data scientist, économiste), GenerationLibre analyse dans cette étude l'ensemble des enjeux socio-économiques et éthiques liés aux données personnelles (I) et étudie comment introduire en droit une patrimonialité des données personnelles (II).

Après avoir montré que la donnée personnelle peut être caractérisée dans le cadre du **droit commun des biens,** nous analysons sous quelle forme juridique - comme propriété intellectuelle ou comme un contrat de licence - un droit de propriété peut voir le jour en droit français. Nous proposons de **soumettre les futures transactions à la TVA,** en considérant qu'il s'agit là d'une opération de troc (services contre données) entre un professionnel (le moteur de recherche ou le réseau social) et un consommateur (l'internaute). La TVA collectée viendrait nourrir les budgets des Etats de l'Union Européenne et assurerait un partage de la chaîne de valeur en lien avec l'intérêt général.

Le droit français comme le droit européen se prêtent particulièrement bien à l'inscription d'un tel droit de propriété.

Au niveau européen, le **Règlement général sur la protection des données personnelles** (RGPD), qui entrera en vigueur le 25 mai 2018, avance dans la bonne direction en attribuant aux entreprises un rôle de « gardiennes » de données et non pas de propriétaires, et en garantissant la **portabilité** des données personnelles. L'attribution d'un droit de **propriété** serait une suite logique de ces avancées réglementaires.

Nous explorons enfin comment la technologie nous permet aujourd'hui de mettre en œuvre ce nouveau droit de propriété (III). Nous analysons plusieurs méthodes possibles pour s'authentifier et mettre ses données à disposition, et proposons un modèle basé sur **une « blockchain »** gérant des **« contrats intelligents »** pour permettre à chacun de rassembler et éventuellement de **commercialiser** ses données.

Afin d'évaluer la valeur d'un tel « marché des données », cette étude sera complétée par une modélisation économétrique en cours de réalisation en partenariat avec des chercheurs de la *Toulouse School of Economics*. Cette modélisation permettra de se faire une meilleure idée des revenus que chacun pourrait attendre d'un flux quasi continu de nanopaiements.

Nous croyons en un Internet décentralisé où **l'identité individuelle** a encore une signification et où l'être humain n'est pas la proie de « serveurs sirènes » pour reprendre l'expression de Jaron Lanier. C'est l'occasion pour la **France** et l'**Europe** d'innover et d'imposer **leur modèle.**

02

INTRODUCTION

Le pacte softien.

PAR GASPARD KOENIG

Les surfeurs du net sont comme les voyageurs d'autrefois, détroussés en chemin. Les voleurs, ce sont les GAFA[1] et autres entreprises du Net. **Le larcin : nos données.**

Mais contrairement à nos prédécesseurs, nous semblons y prendre un certain plaisir. Tous les jours, nous acceptons des dizaines de *cookies* sur nos ordinateurs et cliquons « ok » sur des conditions d'utilisation (« terms and conditions ») qui nous dépossèdent de nos données personnelles. Selon une étude de Carnegie Mellon University, un Américain moyen en signe près de 1500 par an, ce qui correspondrait à 76 jours de lecture[2]. Les conditions d'utilisation de PayPal sont plus longues que Hamlet. Comment imaginer dans ces conditions que l'internaute donne son « consentement éclairé », comme le voudrait le droit ? Nous ne pouvons pas lire ces contrats, encore moins les négocier. Si nous y jetions un œil, nous pourrions avoir des remords. L'utilisateur de LinkedIn confie ainsi au réseau social un droit irrévocable, mondial, perpétuel, sans limite ni rémunération, de copier, d'utiliser et de revendre toutes les informations qui lui seront confiées. Lorsque Facebook a entrepris, à des fins de recherche interne, de manipuler les émotions de 700 000 utilisateurs en altérant les posts qui leur étaient donnés à voir, l'entreprise a pu se prévaloir de termes contractuels l'autorisant à effectuer des « recherches et analyses » à partir des données collectées[3]. Par facétie, GameStation avait inclus entre les lignes de ses contrats l'abandon de l'âme éternelle de l'utilisateur ; sans surprise, 700 utilisateurs avaient conclu ce pacte méphistophélique en une journée…

Naturellement, nous recevons en échange de ces données des services gratuits. Ce que souligne Jean Tirole : *« On entend souvent dire que les plateformes devraient payer pour les données que nous leur fournissons. En pratique cependant, certaines le font effectivement, non sous la forme d'un transfert financier mais sous la forme de services non tarifés. »*[4]

---

1 Acronyme usuel pour Google, Amazon, Facebook, Apple.

2 McDonald, A. M., & Cranor, L. F. (Carnegie Mellon University), The Cost of Reading Privacy Policies, A Journal of Law and Policy for the Information Society, 4(3), 2008, pp. 543–568.

3 Source : http://www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/.

4 TIROLE Jean, Economie du bien commun, PUF, 2016.

Les plateformes se rémunèrent en effet via l'utilisation et/ou la revente de ces données, principalement à des services de publicité qui peuvent ainsi mieux identifier les consommateurs[5]. La quasi-totalité des revenus de Facebook est ainsi générée par de la publicité ciblée (de manière croissante via l'application, plutôt que sur le site)[6], ce qui explique sa lutte acharnée contre les logiciels antipubs. Et ce n'est qu'un début : dans une lettre de mai 2014 à la *Securities and Exchange Commission,* le régulateur financier américain, Google expliquait que la publicité arriverait bientôt sur les frigos, les pare-brise de voiture, les lunettes et les montres…

Les pires dystopies sont possibles : dans une scène fameuse du film *Minority Report,* le héros se voit proposer des publicités personnalisées alors qu'il marche dans la rue (par reconnaissance rétinienne) : *« John Anderton ! Vous pourriez boire une Guinness » ; « Partez en voyage, John Anderton, oubliez vos soucis ».* Le deal est assez simple : contre la gratuité des sites, des réseaux sociaux, des moteurs de recherche ou des morceaux de musique, nous acceptons de confier nos données à des algorithmes qui nous proposent en retour des produits sur mesure. Appelons cela **le pacte softien** (comme software, et comme Faust). Arrangement contractuel entre parties consentantes, pour un bénéfice mutuel ? À voir. Le pacte softien comporte au moins cinq clauses léonines.

Une clause culturelle : loin de se voir offrir une infinité de choix possibles, chacun finit par voir ce qu'il veut voir, entendre ce qu'il veut entendre, lire ce qu'il veut lire. Prenant en compte nos recherches passées, Google nous livre en priorité les informations susceptibles de nous plaire, au risque d'amplifier nos biais subjectifs[7]. Au lieu de nous ouvrir au monde, l'Internet nous enferme dans notre bulle[8].

Une clause sociale : nous devenons dépendants des « serveurs sirènes » qui, après nous avoir attirés par le chant de la simplicité, nous imposent peu à peu leur propre système de normes. L'exclusion de toute image de nu par Facebook, quand bien même il s'agirait

---

5 Certains, comme Criteo, ont fait fortune dans les méthodes de « reciblage publicitaire ».

6 Source : https://investor.fb.com/investor-news/press-release-details/2017/facebook-Reports-Fourth-Quarter-and-Full-Year-2016-Results/default.aspx.

7 Il a été ainsi noté, après les attentats de Charlie Hebdo, que la diffusion des théories du complot a été accélérée par les algorithmes de Google (pour être clair, un habitué des sites pro-palestiniens sera automatiquement référé vers des contenus à tendance antisémite).

8 Ce que le cybermilitant Eli Pariser appelle la *« filter bubble »*.

d'un tableau de Courbet, en est un exemple fameux. Comme l'explique Jaron Lanier, l'un des pionniers de la réalité virtuelle : « *'Free' signifie inévitablement que votre vie va être décidée par quelqu'un d'autre.* »[9]

Une clause économique : la gratuité repose sur la fiction selon laquelle nous avons tous la même valeur vis-à-vis des plateformes (à savoir, zéro), alors que certains utilisateurs produisent d'importantes masses de données, et que d'autres se comportent en passagers clandestins, laissant le minimum de traces digitales.

Une clause politique : si nous restons aujourd'hui libres de nous déconnecter, rien n'assure que ce soit le cas demain. Les autorités auront la tentation de nous imposer le *Big Data* au nom du bien public. Pour réduire notre consommation d'électricité, rendre obligatoires les compteurs intelligents ; pour éviter embouteillages et accidents, rendre obligatoire l'interconnexion des GPS ; pour améliorer la santé publique, rendre obligatoires les bracelets connectés ; etc.

Une clause juridique, fondamentale : à quelle vie privée ai-je droit dans un monde où la moindre start-up peut racheter des fichiers contenant mes goûts, mes déplacements, mes amours ?
Il est remarquable qu'un chantre de la « fin de la vie privée » comme Mark Zuckerberg fasse construire un mur de deux mètres de haut autour de sa résidence hawaïenne, ou rachète toutes les maisons adjacentes à sa propriété de Palo Alto. La fin de la vie privée ne vaudrait-elle que pour les autres ?

En fait, le pacte softien est vicié par une anomalie fondamentale : aujourd'hui, nos données échappent à notre contrôle.

Elles sont éparpillées dans la nature sauvage de la toile, *res nullius* dont le premier occupant peut s'emparer, et convertir ensuite en bases de données protégées, elles, par la propriété intellectuelle[10]. Il s'agit

[9] LANIER Jaron, *Who Owns the Future ?*, Simon & Schuster, 2013.
[10] BENABOU Valérie-Laure, ROCHFELD Judith, *A qui profite le clic ? : Le partage de valeur à l'ère numérique,* Odile Jacob, 2015.

d'une captation de valeur phénoménale : selon le Boston Consulting Group, la valeur des données personnelles en Europe pourrait atteindre 1 000 milliards d'euros d'ici 2020, soit 8 % du PIB européen[11]. C'est la version contemporaine de la tragédie des communs, quand la coexistence de plusieurs troupeaux sur un pâturage public conduit à la surexploitation d'une ressource rare (l'herbe, en l'occurrence) : les acteurs digitaux broutent le même pâturage de données, sans se soucier des conséquences.

Face à ce constat largement partagé[12], trois logiques peuvent être mises en œuvre, reflétant trois options classiques de philosophie politique.

La première, c'est la nationalisation. L'Etat saisit le pâturage et alloue les parcelles. Les données sont alors considérées comme une *res communis*, au même titre que l'air ou l'eau de la mer. Il faudrait prévoir une sorte d'agence nationale des données, rassemblant, mutualisant et chiffrant l'ensemble des données de la population, pour les mettre ensuite à disposition, sous certaines conditions, des entreprises les mieux à même de les utiliser. Cette forme de communisme digital est assez populaire en France[13]. Une telle mainmise de l'Etat sur nos données créerait une bureaucratie diamétralement opposée à la culture de l'Internet, et donnerait au pouvoir central des moyens de contrôle extravagants.

La deuxième option repose sur les droits fondamentaux. L'Etat régule l'utilisation des pâturages et crée des obligations pour les bergers. Cette logique dite « personnaliste », car elle attache des droits à la personne, a été embrassée par la Commission européenne et les divers régulateurs nationaux (comme la CNIL en France). Elle est fondée sur le concept, dégagé par la Cour Constitutionnelle fédérale allemande, d'« autodétermination informationnelle » : chacun doit pouvoir décider de manière autonome de l'usage de ses données. Dans la lignée du droit à l'oubli, l'utilisateur se verrait octroyer des droits supplémentaires : droit d'autoriser la mise en circulation des données et d'en connaître la destination (y compris par exemple via des compte-rendus annuels), portabilité des données d'un outil à l'autre,

11 Boston Consulting Group and Liberty Global, The Value of our digital identity, novembre 2012.

12 Conseil d'Etat, *Etude annuelle 2014 du Conseil d'Etat - Le numérique et les droits fondamentaux,* La documentation française, Paris.

13 Particulièrement sous la plume de Pierre Bellanger, qui a mis en avant l'idée d'une « souveraineté numérique ».

renforcement des procédures de consentement, etc. Les plateformes, elles, devraient se soumettre à de nouvelles obligations : dévoiler les paramètres des algorithmes, offrir des services alternatifs non personnalisés, révéler les processus de traitement des informations obtenues, etc. Le risque de cette logique est d'aboutir à une judiciarisation exponentielle du monde digital, freinant l'innovation sans pour autant offrir aux utilisateurs de garanties réelles (et encore moins de rétribution).

C'est pourquoi il est urgent d'explorer une troisième option, qui complète et en un sens fonde la deuxième : celle de la patrimonialité des données. Autrement dit, rendre l'individu juridiquement propriétaire de ses données personnelles, ce qui n'est aujourd'hui le cas nulle part au monde. L'Etat garantit aux bergers la propriété de leur parcelle de pâturage — libre à eux de les échanger pour trouver le meilleur équilibre. Car si les données sont, selon la formule convenue, le pétrole du 21e siècle, il est temps de poser la question : à qui appartient le pétrole ? Au producteur primaire, qui le revend à d'autres pour le raffinage. C'est-à-dire à vous et à moi, producteurs de données, qui devrions être rémunérés pour la matière première que nous pourvoyons aux algorithmes du *Big Data*. De même que la révolution industrielle a rendu nécessaire le droit de propriété intellectuelle, la révolution numérique devrait créer un droit de propriété sur les données[14].

Cette option a été explicitement rejetée par le Conseil d'Etat dans son rapport sur de 2014 sur le numérique et les droits fondamentaux, au motif qu'elle impliquerait de *« renoncer à la logique de protection »* (de l'individu par l'Etat). Implicitement, le Conseil d'Etat fait écho à un argument familier des juristes, à savoir que les données personnelles ne pourraient pas être monétisées au nom de la protection des libertés fondamentales : la personne étant réputée indisponible et ne pouvant faire l'objet d'un commerce, les data qui en émanent devraient elles aussi être exclues du marché. Elles feraient partie de ce que *« l'argent ne peut acheter »*, pour reprendre l'expression du philosophe américain Michael Sandel.

[14] REES Christopher, « Who owns our data ? », *Computer Law & Security Review*, 30, 2015.

Cet argument moral et philosophique suggère au moins quatre niveaux de réponse :

• Le premier, **empirique,** constate l'existence aujourd'hui de profits colossaux engrangés par les agrégateurs de data : la justice la plus élémentaire ne consiste-t-elle pas à mieux distribuer la chaîne de valeur ?

Au nom de quelle « dignité humaine » refuser au consommateur-citoyen sa part légitime dans la production économique ?

• Le deuxième, juridique, rappelle que le droit de propriété est essentiellement conçu comme un outil de maîtrise : chacun peut ensuite en disposer comme il l'entend, y compris en refusant les mécanismes de marché. On ne peut véritablement choisir de donner que ce que l'on possède.

• Le troisième, **moral,** justifie de renoncer à la logique de la protection pour lui substituer une logique de responsabilité : dans une société mature, l'Etat doit abandonner la tutelle du citoyen-consommateur, et lui faire confiance pour utiliser ses données de manière éclairée.

• La quatrième, **philosophique,** revendique l'idée lockéenne d'une « propriété de soi[15]» comme idéal de la modernité, affranchissant l'individu de l'emprise de toute transcendance.

Cependant, il ne nous semble pas nécessaire d'ouvrir ce débat sensible. En effet, les data peuvent s'envisager dans le contexte du droit des biens, comme des choses que l'on peut s'approprier et contrôler. En ce sens, elles restent distinctes de la personne — de même que les « idées », qui elles aussi entretiennent un rapport intime avec l'individu qui les a produites, peuvent relever de la propriété intellectuelle.

Il faut également noter la crainte émise par le Conseil d'Etat que « *la reconnaissance du droit de propriété de l'individu sur ses données pose de sérieuses difficultés juridiques pour les pouvoirs publics*[16]», qui devraient dès lors justifier de la collecte et du traitement des données des citoyens au regard d'un but d'utilité publique. Cette crainte

---

15 LOCKE John, *Second Treatise of Government*, Chapter 5 : Property, 1690 : « *Though men as a whole own the earth and all inferior creatures, every individual man has a property in his own person ; this is something that nobody else has any right to.* » *Locke reprend en fait l'idée de « self-propriety »* déjà développée par Richard Overton quelques années plus tôt.

16 Conseil d'Etat, *Op. cit.*

représente plutôt pour nous un espoir : n'est-il pas souhaitable que l'Etat doive s'expliquer, au besoin devant le juge judiciaire, de la manière dont ses administrations (y compris les services de renseignement) s'emparent de nos data ? Cela ne permettrait-il pas de renforcer la nécessaire confiance entre l'Etat et les citoyens à l'âge digital ?

Le droit français et européen se prête particulièrement bien à l'inscription d'un tel droit de propriété, suite logique des nouvelles régulations autour de la protection des données. C'est l'occasion pour l'Europe d'innover et d'imposer son modèle. Pour que, demain, ce soit Facebook qui nous paie.

03

PARTIE 1

Les aspects socio-économiques et éthiques des données personnelles.

PAR VIRGINIE PEZ-PÉRARD - ISABELLE LANDREAU - LUCAS LÉGER

L'évolution technologique a régulièrement été à l'origine de transformations économiques majeures. L'émergence de l'industrie informatique a permis un accès et un traitement efficace de jeux de données de plus en plus importants. Les industries qui ont su se placer en intermédiaire sur ce nouveau marché sont aujourd'hui des entreprises surpuissantes, non seulement en termes financiers, mais également d'information. Dans les sociétés où cette dernière est devenue une commodité qui s'échange à prix d'or, il nous paraît crucial de s'interroger sur ce nouvel ordre économique. Cette partie introductive permet de replacer notre analyse dans un contexte historique, juridique, économique, et éthique.

L'objectif de cette partie est d'abord d'informer le lecteur. Le consommateur ne peut faire des choix que s'il détient suffisamment d'information pour être en capacité de mesurer l'impact de son choix, au moins de façon approximative, car nous passons un temps important sur la plupart des plateformes. En moyenne, nous consacrons 2,5 heures par jour en 2017 à parcourir les médias sociaux, contre 45 minutes en 2012[1]. Et plus nous passons du temps sur ces plateformes, plus elles collectent de données.

Ces capacités de collecte sont aujourd'hui très importantes. L'affaire Snowden a au moins eu le mérite de mettre à jour l'ampleur du phénomène : le programme PRISM lancé par la NSA pouvait intercepter, déjà en 2010, jusqu'à 1,7 milliard de courriels, appels téléphoniques et autres télécommunications[2]. Elle a également montré que les capacités techniques de collecte et de traitement de données sont aujourd'hui une réalité.

---

1 Source : http://blog.globalwebindex.net/chart-of-the-day/social-media-captures-30-of-online-time/

2 HARCOURT B.E., *Exposed*, Harvard University Press, 2015.

Une fois les données analysées, plateformes Internet et gouvernements tirent des informations qui vont bien au-delà de la simple interaction commerciale. Est-ce un problème ? Lisez ces quelques pages et vous pourrez le décider pour vous-même.

1. La société de l'information et l'explosion des données

1.1 AU COMMENCEMENT, LA RÉVOLUTION NUMÉRIQUE

Tout comme la Révolution industrielle, la Révolution numérique, et plus généralement les innovations radicales qui en découlent, ne transforme pas seulement notre façon de consommer et de produire, mais chamboule à son tour l'ensemble de notre paysage économique et institutionnel. Bien plus qu'en crise, nos économies sont en profonde mutation, caractérisée par l'avènement de la société que l'on pourrait qualifier de la connaissance[3], où les nouvelles technologies nous aident à surmonter nos limites cognitives. Ces révolutions technologiques successives ont profondément affecté l'humanité.
« Le changement technologique est en grande partie responsable de l'amélioration de la condition humaine, que ce soit par la taille de sa population, l'allongement de l'espérance de vie, le niveau d'éducation, le niveau de vie, l'évolution du travail, des télécommunications, des soins de santé, ainsi que les effets des activités humaines sur notre environnement. »[4]

Après la Révolution industrielle, nous assistons aujourd'hui à une véritable rupture entraînée par le changement technologique, où **l'information devient la variable créatrice de valeur**[5]**.** Les réseaux sont les instruments de transmission et d'échange de cette nouvelle

3 BELL Daniel, *The coming of post industrial society*, Basic Books, 1973 ; DRUCKER Peter, *Age of discontinuity*, Harper & Row, 1969 ; TOFFLER Alvin, *The third wave*, Bantam Books, 1980.

4 BOSTROM Nick, *Technological revolutions: Ethics and policy in the dark*, Published in Nanoscale : Issues and Perspectives for the Nano Century, eds. Nigel M. de S. Cameron and M. Ellen Mitchell, John Wiley, 2007, pp. 129-152. Bien que nous nous écartions de notre sujet, il faut noter que cette vision ne fait pas consensus. En effet, le rapport de l'homme à la technique a été vivement critiqué par Jacques Ellul, in *Le bluff technologique*, Hachette, 1988.

5 HIDALGO Cesar, *Why information grows : The evolution of order, from atoms to economies*, Basic Books, 2015.

forme de connaissance, et plus largement de l'innovation[6]. L'information[7] se transmet et se propage grâce au développement du logiciel.

Dans une déclaration écrite, l'investisseur et cofondateur de Napster, Marc Andreessen[8], avançait l'argument que, de plus en plus, **le logiciel** allait « *conquérir le monde* » et que tous les secteurs de l'économie seraient plus ou moins touchés par l'arrivée de nouveaux acteurs capables de réduire les coûts fixes pour fournir un service moins cher et plus efficace. À bien y réfléchir, la nouvelle économie de la connaissance a en effet permis l'émergence de ces logiciels, qui permettent à leurs utilisateurs d'accéder en quelques clics à un service performant : Amazon dans le commerce de détail, Google dans le marketing et la publicité, les drones dans le secteur de la défense, le super ordinateur Watson développé par IBM pour assister de nombreux médecins dans leurs diagnostics, la musique avec Spotify, iTunes ou Deezer, le paiement via PayPal, les MOOC (Massive Open Online Courses) dans l'éducation, la photographie, les films animés, etc.

Le développement du logiciel touche tous les secteurs de l'économie et contribue à l'automatisation dans l'industrie et les services. Toutes ces innovations se nourrissent d'une quantité de données toujours plus importante, qui concerne le plus souvent des comportements monétisables. Afin de vous fournir un service aussi performant, **ces nouvelles plateformes doivent collecter et retraiter des données personnelles,** que ce soit lors d'un achat sur Internet, ou d'une simple navigation sur le Web.

1.2 LA SURVEILLANCE PAR LES ETATS : « BUSINESS AS USUAL » ?

De manière classique, les Etats surveillent, et le *Big Data* est l'occasion de surveiller encore plus, au seul prétexte de garantir une protection à ses citoyens.

Chaque seconde, c'est un torrent de données personnelles et autres

6 VALENTE, T.W., « Social network thresholds in the diffusion of innovations », *Social Networks*, 18 (1), 1996, pp. 69-89; DEROIAN, F., « Formation of social networks and diffusion of innovations », Research policy, 31 (5), 2002, pp. 835-846.

7 Connaissance et savoir sont les deux composantes fondamentales à l'émergence de l'information

8 ANDREESSEN Marc, « Why sofware is eating the world », World Street Journal, August 20, 2011.

qui se déversent dans des « data centres[9] **»,** la plupart situés sur le sol américain. En toute conscience, nous sommes en quelque sorte **dépossédés de notre identité.** La frontière autrefois évidente entre la sphère privée et notre identité publique (notre nom, notre adresse, etc.) n'est plus. Si Google vous connaît mieux que vos proches[10], l'utilisation des données massives n'est pas l'apanage des seuls GAFA. Bien au contraire, un certain laxisme quant à la propriété des données personnelles permet également à l'Etat d'accroître sa surveillance et d'améliorer ses méthodes. En outre, ce renversement ne s'est pas fait de façon coercitive.

La Chine reste toujours avant-gardiste en termes de surveillance et de restriction de la liberté d'expression. Déjà en 2009, elle proposait la mise en œuvre du « Green Dam project ». Ce projet avait pour but d'installer un logiciel de contrôle parental sur tous les nouveaux ordinateurs produits en Chine. Si l'intention est louable, quelques tests suffisent cependant à démontrer que le programme allait bien au-delà du simple contrôle[11]. Au-delà des risques de sécurité qu'il posait, ce logiciel était parfaitement intrusif et permettait à la Chine de pratiquer un espionnage intensif de ses citoyens.

Plus récemment, le 14 juin 2014, le Conseil d'Etat chinois dévoile son Plan directeur pour la construction d'un système de crédit social (2014-2020)[12]. L'ambition serait de construire un système de réputation national, permettant l'évaluation des citoyens chinois sur la base de leurs données numériques personnelles. L'objectif d'un tel outil, pour le leadership chinois, serait de renforcer, à terme, l'intégrité et la *« sincérité dans les affaires publiques, le commerce, les questions sociales et la construction de crédibilité judiciaire*[13]*»*. Les moyens ? L'analyse et la compilation du comportement numérique de ses citoyens sous forme de scores individuels. Cet outil de pilotage, fondé sur un système d'incitations à bien agir, pourrait prendre dès aujourd'hui la forme d'une application smartphone, que l'on pourrait déployer en version synchronisée

---

9 Source : http://www.iflscience.com/technology/how-much-data-does-the-world-generate-every-minute/. Sur le sol américain, ce sont 2 657 700 de gigabits de données Internet qui sont générés chaque minute.

10 Source :http://www.journals.uchicago.edu/doi/abs/10.1086/680084. Voir également : STEPHENS-DAVIDOWITZ S., *Everybody Lies : Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*, Dey Street Books, 2017.

11 Source : https://opennet.net/chinas-green-dam-the-implications-government-control-encroaching-home-pc.

12 Source : http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion

13 Source : https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/

sur nos divers écrans, notre montre connectée, et la partie du *cloud* nous étant réservée. La mise en place de tout un système de réflexes pavloviens pourrait nous inciter à conduire nos affaires en toute intégrité, récompensés dans notre bien-agir et sanctionnés pour nos écarts. Car c'est bien de nous qu'il s'agit, de nos comportements, traduits en données partagées sur le *cloud* lors de nos pérégrinations numériques. C'est bien sur la base de nos propres données que de tels systèmes reposent. Sans elles, ils ne sont rien.

Selon une vidéo de *The Economist*[14], le gouvernement chinois détiendrait **un outil de reconnaissance faciale** puissant, grâce notamment à l'efficacité croissante des outils de « machine learning ». Cathay a ainsi constitué une base de données d'environ 700 millions de profils différents, soit la moitié de sa population. Les entreprises qui gèrent du contenu sensible, comme les banques ou même les Etats pourraient sauter le pas à plus ou moins long terme. Ainsi, on pourrait améliorer la sécurité des individus. La reconnaissance faciale est un outil efficace pour limiter l'usurpation d'identité, traquer les criminels et terroristes. À ce jour, la start-up *Face-six* se gargarise d'un taux de succès de 99 %. C'est encore insuffisant pour des applications dans la sécurité, qui retourneraient potentiellement trop de faux positifs à gérer administrativement, mais ce taux est en croissance.

On pourra toujours rétorquer qu'il ne s'agit pas d'une démocratie puisque ces quelques exemples illustrent la tendance en Chine. Cependant, l'affaire Snowden a révélé que la surveillance de masse n'était pas uniquement monopolisée par des gouvernements aux principes démocratiques douteux.

1.3 LA TRAQUE DES DONNÉES PERSONNELLES ÉRIGÉE EN MODÈLE D'AFFAIRES

Nous voyons également apparaître des applications commerciales à la reconnaissance faciale. L'un des potentiels bénéfices pour les individus serait la fin des mots de passe. L'iPhone X l'a supprimé et son accès est désormais facilité. Toujours en Chine, on l'utilise pour l'accès aux

[14] Source : https://www.youtube.com/watch?v=nT_PXjLol_8.

parcs d'attractions ou comme moyen de paiement dans la restauration rapide[15]. Les données de comportement sont ainsi stockées et traitées dans le but d'améliorer « l'expérience client » nous dit-on. En effet, le ciblage publicitaire en fonction de notre genre et nos caractéristiques d'achat se fait sans même que nous n'ayons besoin de remplir un quelconque formulaire ou de détenir une carte de fidélité.

Cela signifie-t-il, pour autant, que nous contrôlons nos données personnelles ? Nos traces numériques, si elles ne servent pas encore à la compilation de score de crédit social des citoyens, sont bel et bien utilisées dans la production de valeur de ceux qui savent les exploiter. L'université de Caroline du Nord estime à 426 millions de dollars[16] le chiffre d'affaires de 2012 des neuf courtiers de données personnelles les plus importants. Nous, producteurs primaires de ces données, ne participons pourtant pas à ce marché. Nous partageons nos données **volontairement** — du moins en apparence — le plus souvent car cela constitue **une condition d'utilisation** des services que nous utilisons sur Internet. Pourtant, certaines études montreraient que nous serions même prêts à payer pour que nos données ne soient pas partagées lors de nos interactions sur Internet[17]. Et si c'était possible ? Pour le savoir, revenons sur les raisons pour lesquelles nos données sont aujourd'hui au cœur des modèles d'affaires des entreprises.

Dans le modèle traditionnel, les entreprises financent les services digitaux qu'ils proposent aux consommateurs pas le biais de la publicité. C'est ainsi que la plupart des sites web ou applications gratuites (pour ne pas dire tous) jouent des pop-ups ou des bannières publicitaires qui leur permettent de générer des revenus publicitaires. Mais face à la résistance des consommateurs (réactions de rejet face à ces formes de publicité ; progression de +20 % des *adblocks* en 2016[18]), les audiences des publicités chutent. Il devient indispensable de proposer un meilleur ciblage des campagnes publicitaires pour augmenter la rentabilité des actions publicitaires pour les annonceurs. Par ailleurs, il est primordial de préserver l'attention des cibles en limitant la pression commerciale qui est exercée à leur égard. Pour cela, les entreprises misent tout sur les données. Celles-ci sont collectées et utilisées à la fois pour dresser un profil détaillé du client mais aussi pour les revendre à des « partenaires » (qui sont en fait des clients bien sûr !).

---

15 Toujours selon le reportage de The Economist, *Ibid.*

16 Source : https://onlinemba.unc.edu/blog/data-brokers-infographic/

17 *Idem.*

18 Baromètre Adblocks IAB France/ Ipsos – Novembre 2016, http://www.iabfrance.com/content/presentation-de-la-v2-de-letude-ipsos-realisee-pour-liab-france-sur-les-adblocks.

Noms, prénoms, adresses email, numéros de téléphone, mais aussi produits consultés, achetés, intérêts manifestés, habitudes… Toutes ces données sont aujourd'hui l'une des sources de valeur majeure pour les entreprises, en substitut ou en complément de recettes publicitaires réduites.

Mais qu'en retire le client en échange ? La collecte de données par les entreprises répond certes avant tout à une logique mercantile. Les données sont une source de revenus considérable, via l'augmentation de la valeur des espaces publicitaires vendus (du fait d'un meilleur ciblage possible), et la revente aux « partenaires ». Mais à la base, celles-ci sont indispensables pour que les technologies digitales fonctionnent de manière intelligente et pertinente. Grâce à elles, les entreprises peuvent personnaliser le service, comme conserver les produits mis au panier lors d'une précédente visite par exemple, ou recommander des produits en fonction du profil du consommateur ou de ses achats précédents. Elles peuvent ainsi proposer aux utilisateurs des expériences fluides, simples et pratiques. De fait, même s'ils n'en sont pas toujours conscients, les consommateurs retirent de l'utilisation de leurs informations personnelles des **bénéfices fonctionnels** (gain de temps, confort, bénéfices de commodité). Parfois, ils peuvent aussi en retirer des **bénéfices monétaires.** C'est le cas notamment dans le cadre des cartes de fidélité. En passant leur carte au moment de leurs achats, les clients obtiennent des réductions ou des avantages, directement ou indirectement, en échange de points qu'ils cumulent proportionnellement à leur consommation. Offrir ces bénéfices aux clients est nécessaire, car ils garantissent que le client « jouerait le jeu » du programme et passerait sa carte à chaque achat, et cela est indispensable pour avoir une vision complète et fidèle du client. Mais la **question de l'équité** de telles pratiques se pose. Même si les données permettent d'offrir des services améliorés aux citoyens-consommateurs et que cette amélioration est une forme de valeur qui répond à de vraies attentes, **la part de valeur restituée aux clients** sous forme de ces avantages est sans doute **relativement faible** par rapport à **la valeur totale qu'en retirent les distributeurs** via la revente des informations.

Restaurer **l'équilibre** paraît essentiel dans une perspective éthique, mais aussi pour préserver la croissance et ne pas arriver au point de non-retour qui ferait que les citoyens-consommateurs se détourneraient

de ces services. Pour avancer sur cette question, il est nécessaire de comprendre quand et comment les données sont recueillies, en pratique.

1.4 LA COLLECTE ET L'UTILISATION DES DONNÉES PERSONNELLES EN PRATIQUE

Les entreprises profitent de chaque occasion de contact pour collecter des données sur leurs clients, et les enrichissent en continu via de multiples démarches.

- **En magasin physique, la carte de fidélité ou le « fichier client »**

En magasin physique, il n'y a pas un passage en caisse qui ne soit désormais assorti de la sempiternelle question du vendeur « avez-vous la carte de fidélité du magasin ? » ou « êtes-vous enregistré dans notre fichier client ? » Même si ces informations ne sont pas utiles pour finaliser l'achat, les consommateurs se voient demander leur identité, leur adresse ou code postal, leur adresse email, leurs coordonnées téléphoniques, leur date de naissance, voire même la composition de leur foyer ou le prénom de leurs enfants ! Chaque achat y est minutieusement enregistré : produits, date, heure, magasin.
Ces données sont ensuite croisées pour dresser un profil le plus complet possible du client. Si une entreprise de puériculture voulait par exemple cibler de futures mamans, elle pourrait ainsi acheter des données de consommatrices qui, sur une même période, ont acheté en magasin des tests d'ovulation ou de grossesse et dont la consommation d'alcool a baissé dans les derniers mois. La marque aurait ainsi toutes les chances de toucher sa cible avec de tels critères et de maximiser par là même son retour sur investissement.

Ces pratiques posent question : au-delà des revenus générés par leurs données sans qu'ils en aient conscience, ces méthodes peuvent pousser les consommateurs à sur-consommer et à acheter des produits dont ils n'ont pas réellement besoin, en leur faisant miroiter une pseudo-promotion donnant l'illusion d'une « bonne affaire ». Dans un contexte où le taux d'endettement des ménages ne cesse de grimper (+1,1 point entre 2015 et 2016 en France[19]) soumettre les consommateurs-citoyens

[19] Chiffres communiqués par la Banque de France, consultables sur le lien https://www.banque-france.fr/statistiques/credit/endettement-et-titres/endettement-des-agents-non-financiers.

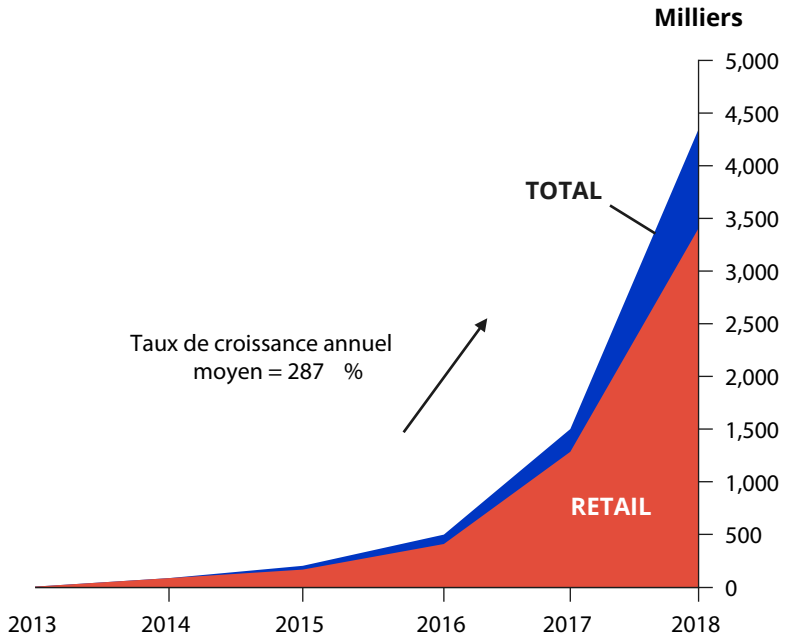
à de telles tentations est-il raisonnable ? La question mérite d'autant plus que l'on s'y intéresse que des recherches récentes ont montré que certaines pratiques marketing ciblant les consommateurs sur la base de leurs données personnelles pouvaient les inciter à faire des achats qu'ils regrettent ensuite. Cela peut générer un malaise dans leur relation avec l'entreprise et au final, être source de mal-être pour le consommateur[20].

• En magasin physique, les « beacons » et autres technologies connectées

Les « *beacons* » sont des bornes installées dans les magasins (à ses abords ou à l'intérieur des rayons) qui fonctionnent grâce à la technologie bluetooth. Elles permettent de capter qui sont les consommateurs proches des bornes, car l'entreprise sait rapprocher, grâce à son fichier client, un consommateur et son téléphone. Une fois l'individu repéré aux abords du magasin ou au sein d'un rayon particulier, **l'entreprise peut alors lui pousser en temps réel des offres promotionnelles personnalisées.** D'après le site Business Insider, il existera en 2018, 4,5 millions de boîtiers en activité aux Etats-Unis, dont 3,5 millions pour le secteur du Retail (cf. graphique)[21]. 50 % des principaux détaillants nord-américains ont déjà initié des phases d'expérimentation depuis 2014, et on estime qu'ils ont équipé 85 % de leurs points de vente à fin 2016. Pour les détaillants, il s'agit de la vitesse d'adaptation de technologie la plus rapide depuis l'équipement en lecteurs mobiles de cartes de crédit ! Comme souvent, le marché français devrait logiquement suivre le schéma d'adoption du marché américain en accélérant l'équipement de ses points de vente dans les prochaines années.

[20] BUTORI R., MIMOUNI A., PEZ V., « Le côté sombre de la pression exercée sur les consommateurs par les programmes de fidélité: enjeux éthiques et pratiques », Recherche et Applications En Marketing, Vol. 32, n°3, 2017, pp. 76-89.

[21] Source : http://www.businessinsider.fr/us/beacons-are-the-most-important-new-retail-tech-2014-7/.



Prévision du nombre de *beacons* en service sur le marché américain en milliers d'unités (*Business Intelligence Estimates*).

Les magasins physiques capitalisent également sur les technologies digitales pour proposer des services à valeur ajoutée aux consommateurs… tout en en profitant pour récupérer des informations personnelles bien sûr. En 2017, Facebook a par exemple testé une nouvelle fonctionnalité de fidélité au sein de son application mobile (« Rewards »). L'option proposait aux utilisateurs de scanner des QR codes déployés en magasins physiques pour bénéficier de promotions ciblées[22]. Cette technologie est un outil de promotion des ventes intéressant pour le distributeur, qui peut générer du trafic supplémentaire en magasin. **Mais c'est surtout pour Facebook l'occasion de collecter de nouvelles données sur ses utilisateurs et de tirer des revenus publicitaires supplémentaires.** Même chose les cartes de fidélité connectées, qui permettront à l'avenir de proposer des expériences de shopping personnalisées, avec recommandations de produits et réductions à la clé, et qui seront l'occasion pour les enseignes d'en savoir toujours plus sur leurs clients (exemple : l'enseigne de vêtements Kiabi a testé une carte de fidélité connectée en magasin en janvier 2017 [23]).

[22] Source : https://techcrunch.com/2017/05/01/facebook-rewards/.

[23] Source : http://www.e-marketing.fr/Thematique/general-1080/Breves/Kiabi-teste-carte-fidelite-connectee-magasin-313408.htm#EABi3g3hu0PSfY8d.97.

• **La collecte de données sur les parcours clients digitaux**

Lors de la navigation web, **chaque connexion est tracée.** Les pages visitées, le parcours suivi, le temps passé par page, les produits visualisés, les informations lues, les produits mis au panier, les produits réellement achetés, etc. sont récupérés par l'intermédiaire de *cookies* installés sur les appareils. Sur les smartphones, les localisations et les usages peuvent être tracés. Les smartphones collectent la position géographique de leur utilisateur plusieurs fois par minute[24], souvent via des applications en arrière-plan (c'est-à-dire sans que l'utilisateur en ait conscience) ce qui alimente les bases de données des géants de l'Internet.

• **Le scan des conversations privées**

Les services de messagerie gratuits, de type Gmail, se financent par le biais de la publicité et fonctionnent comme de véritables régies médias. L'objectif de ces services de messagerie est de **proposer un ciblage le plus fin et le plus précis possible, afin de permettre aux annonceurs d'optimiser leurs investissements publicitaires** en maximisant les taux de transformation (c'est-à-dire le ratio entre le nombre d'individus qui ont réellement adopté le comportement souhaité, comme acheter le produit par exemple, sur le nombre de publicités jouées/payées par l'annonceur). Afin d'offrir le ciblage le plus fin possible pour optimiser ce taux, les services de messagerie analysent les conversations privées de leurs internautes pour en extraire des mots clés, et ainsi savoir s'ils manifestent un intérêt pour un produit particulier, s'ils comptent partir en vacances ou se déplacer prochainement et où, s'ils participent à un événement qui nécessiterait un équipement particulier… Bref, de façon générale pour identifier des besoins ou appétences potentiels.

• **La collecte d'informations en échange d'un service gratuit**

De nombreux sites financièrement gratuits pour le consommateur (comparateurs d'assurance, comparateurs d'hôtels ou de billets d'avion, sites d'estimation des prix immobiliers, sites d'estimations de la cote des voitures…) ont un business model basé sur la donnée. Afin de bénéficier

[24] LEFILLIÂTRE J., « Où étiez-vous hier ? Google peut vous le montrer et c'est effrayant », 19/12/2013 http://www.challenges.fr/high-tech/big-data-comment-la-geolocalisation-de-google-traque-tous-vos-deplacements_10453.

du service proposé par le site, le consommateur doit y renseigner un certain nombre d'informations. Pour les sites de comparaison de vols par exemple, l'utilisateur doit renseigner ses dates de voyage, sa destination ou encore le nombre et l'âge des voyageurs…
Pour le site d'estimation immobilière, l'utilisateur doit indiquer s'il est locataire ou propriétaire, l'Etat des parties communes de l'immeuble, voire même des informations a priori hors de propos et sans intérêt pour réaliser le service, comme l'Etat marital et la fourchette des revenus du foyer. On comprend mieux **pourquoi ces informations très personnelles sont demandées lorsque l'on sait que ces sites se financent en grande partie par la revente de ces informations.** C'est ainsi que l'utilisateur recevra peu après l'utilisation de ces services des offres promotionnelles pour les hôtels de sa prochaine destination de vacances ou des appels intempestifs d'agents immobiliers en quête de nouveaux clients. Service gratuit, disaient-ils ?

• La veille sur les réseaux sociaux

Ces informations peuvent être complétées par des données issues des réseaux sociaux, de manière plus ou moins automatisée en fonction de la maturité relationnelle des entreprises. Même si cette pratique n'est pas généralisée aujourd'hui (60 % des entreprises n'intègrent pas de données issues des réseaux sociaux dans leurs analyses[25]), il est très probable que ces informations soient plus largement utilisées à l'avenir au fur et à mesure que les entreprises se doteront d'outils CRM (outils de gestion de la relation client) de dernière génération. **Ces informations sont très riches : elles émanent des individus eux-mêmes qui les partagent en pensant converser avec leur cercle relationnel proche.** Ils n'ont pas conscience que ces informations peuvent être récupérées par les marques. Cette mine d'or d'informations, parfois très intimes, est encore peu exploitée aujourd'hui faute de savoir en industrialiser le traitement. Mais leur utilisation devrait s'intensifier dans les années à venir.

• Les jeux-concours

[25] Etude Oracle, « *Can Virtual Experiences Replace Reality ?* », 2016, http://business.lesechos.fr/directions-marketing/marketing/marketing-digital/0211586548032-en-2020-un-marketing-en-realite-virtuelle-303621.php.

Les jeux-concours sont un moyen simple, rapide, et très rentable pour collecter des données sur les individus par les entreprises. Au motif de pouvoir vous identifier clairement et vous remettre votre lot en cas de victoire, la participation à des jeux-concours requiert souvent la communication de nombreuses informations personnelles, au premier rang desquelles les noms, prénoms, adresse complète, numéro de téléphone (pour prévenir de la victoire, bien sûr !) et date de naissance. **Parfois, d'autres informations sont demandées sous promesse de vous offrir un lot personnalisé, comme la composition du foyer ou des éléments liés aux préférences personnelles, goûts et habitudes.** Et enfin, il est quasi systématique de se voir proposer de multiplier ses chances de gain en communiquant des adresses email de proches susceptibles d'être intéressés par l'offre. En les parrainant ainsi, eux-mêmes deviendront la cible du jeu et communiqueront à leur tour toutes ces informations. Par effet « boule de neige », une campagne de jeux-concours peut ainsi rapporter très gros en matière d'informations personnelles. Les informations collectées sont réputées de bonne qualité, car les individus communiquent honnêtement les informations dans la perspective où il faudrait les contacter pour leur remettre leur gain personnalisé.

- **L'achat de données incrémentales via les fichiers clients**

Pour encore mieux qualifier les individus présents dans les bases de données des entreprises, celles-ci peuvent avoir recours à des achats de fichiers permettant d'enrichir des informations manquantes (ou de constituer de nouveaux fichiers, dans le cadre des politiques de prospection). **Ces fichiers, dont la qualité et la fraîcheur ne sont pas toujours au rendez-vous, s'achètent très facilement auprès de prestataires dont le nombre a incroyablement augmenté ces dernières années.** Ces fichiers permettent d'augmenter la valeur des informations dont dispose déjà l'entreprise, partant du principe qu'une information seule n'a pas de sens. Ce sont les données agrégées, riches, qui ont de la valeur. En proposant une vision de l'individu la plus complète possible, l'entreprise augmente de façon exponentielle la valeur à ses données.

- **… et à chaque occasion de contact**

Enfin, n'oublions pas que chaque contact avec l'entreprise est une occasion de collecter des données sur les individus. Les courriers, échanges téléphoniques, échanges avec les vendeurs en magasin, sont scannés, enregistrés, annotés, « historisés » de façon à compléter au mieux le profil des clients.

Ainsi, en combinant l'ensemble de ces sources d'information, les entreprises parviennent à obtenir **une vision très fine et complète des individus.** Elles parviennent à une véritable « vision 360° » des individus-consommateurs, et les spécialistes de la donnée prévoient une intensification de ces pratiques. Si la plupart des entreprises collectent en effet beaucoup de données, rares sont celles qui sont capables aujourd'hui de lui donner du sens. Selon une étude d'Oracle[26], 42 % des entreprises sont incapables à l'heure actuelle d'extraire des analyses exploitables des données collectées. Il y a donc une marge de progression importante pour les années à venir. Avec les nouveaux outils et métiers de l'ère du *Big Data*, les pratiques de personnalisation vont nécessairement s'intensifier.

[26] Etude Oracle, *Ibid.*

2. Redonner le contrôle au consommateur ?

Dans ce panorama, les citoyens-consommateurs sont de plus en plus inquiets à mesure qu'ils se familiarisent avec ces pratiques. Certains redoutent l'intensification de l'effet « big brother » et la génération de revenus à leurs dépens. Ils commencent à organiser la résistance.

Que fait notre citoyen dans ce monde 2.0 ? Est-il sujet de droit ou objet de toutes les convoitises ? Il est évident que deux visions s'opposent : la vision européenne tournée vers la protection des données à caractère personnel et la vision américaine tournée vers le potentiel d'affaires énorme.

L'avenir du *Big Data* passe par non seulement la sécurité mais aussi par la mise en place d'un nouveau *business model* qui remet **le citoyen au centre de l'exploitation de ses données,** appuyé par un modèle juridique simple, pratique et existant dont nous allons exposer les mécanismes et opportunités.

2.1 LA RÉSISTANCE DES « CITOYENS-CONSOMMATEURS » S'ORGANISE

Les consommateurs n'ont jamais été aussi aguerris et au fait des pratiques des entreprises. Ils en connaissent les outils, les techniques, les *business models*. Dans le même temps, ils se montrent de plus en plus méfiants vis-à-vis du système de consommation et des pratiques commerciales. Les réactions de résistance montent en puissance, prenant la forme de comportements d'évitement de certaines entreprises, d'abandon de certains produits, ou encore de formes de rébellions plus actives comme la réclamation ou le boycott. Ces dynamiques ont toujours existé, mais avec l'avènement du digital, elles peuvent s'organiser facilement, être rapidement visibles à grande échelle ou devenir virales en seulement quelques clics. Les outils tels les pages Facebook des marques, les forums, les sites d'avis consommateurs, voire même

les sites de pétition ou d'appels au boycott organisés sont autant d'armes avec lesquelles ils reprennent le pouvoir. C'est ainsi que des marques comme Petit Navire, Starbucks, Marineland, H&M ou LU par exemple sont aujourd'hui confrontées à la colère des consommateurs qui agissent en véritables contrepoids.

Sur le plan théorique, ces comportements résultent « d'un état motivationnel de résistance », c'est-à-dire d'une tension interne survenant chez le consommateur suite à la perception d'une situation perçue comme oppressive[27]. Plus précisément, ces manifestations sont la réponse de l'individu pour mettre en échec une tentative de pression ou d'influence perçue comme inacceptable en raison des représentations dissonantes et des émotions négatives qu'elle génère. Autrement dit, à trop mettre la pression, les marques pourraient obtenir l'effet inverse de ce qu'elles recherchent.

À cela s'ajoute le fait que nous arrivons à un stade où les consommateurs sont parfaitement familiers des pratiques des entreprises de l'économie digitale. Cette familiarité est de nature à créer des « métacognitions de marché [28]» , sortes d'idées préconçues que développent les consommateurs sur les entreprises et leurs outils, qui font qu'ils ont tendance à associer leurs idées préconçues à toutes les pratiques indépendamment de leurs caractéristiques réelles. En fait, les individus construisent un ensemble de connaissances à propos des techniques d'influence utilisées par les entreprises, au fil de leurs observations directes ou de leurs interactions avec des individus qui y ont été confrontés. Ces connaissances sont ensuite mobilisées pour interpréter les démarches commerciales des entreprises. Confrontés à ces dernières, les consommateurs cherchent à maintenir leur processus de décision classique, indépendamment de ce que l'entreprise tente d'obtenir par le biais de sa démarche.

En matière de données personnelles, c'est ainsi qu'un individu pourrait systématiquement estimer que l'on préempte ou « vole » ses données de façon injustifiée et illégitime (par exemple, pour les revendre), **alors qu'elles sont peut-être simplement indispensables**

[27] ROUX D., « La résistance du consommateur : proposition d'un cadre d'analyse », *Recherche et Applications en Marketing*, 22, 4, 2007, pp. 59-80.

[28] WRIGHT P., « Marketplace Metacognition and Social Intelligence », *Journal of Consumer Research*, 28, 4, 2002, pp. 677-83

à l'octroi du service. Pour couper court à de telles inférences, les entreprises émettent tous azimuts des chartes éthiques sur le respect de la vie privée de leurs clients. En informant sur la nature des données collectées et l'utilisation qui en est faite, elles espèrent éviter d'être taxées d'intentions plus ou moins malhonnêtes ou intéressées.

Ainsi, le citoyen-consommateur digital nourrit un véritable paradoxe : il veut que les services qui lui sont proposés soient les plus personnalisés possible, adaptés à son besoin spécifique que les parcours soient fluides et « sans couture », et tout cela, sans débourser un centime. Mais dans le même temps, il souhaite un relatif anonymat et un contrôle absolu sur ses données personnelles. Ces attentes sont-elles conciliables ?

2.2 VERS UNE RÉTRIBUTION DES DONNÉES PERSONNELLES ?

Face à la prise de conscience collective des pratiques des entreprises en matière de données personnelles et au danger potentiel que cela représente, les consommateurs entendent reprendre le pouvoir.
88 % se disent dérangés par l'exploitation de leurs données personnelles ; 88 % aussi se disent inquiets que leur navigation soit enregistrée par des entreprises privées[29]. Des initiatives indépendantes destinées à faire prendre conscience aux citoyens-consommateurs des revenus publicitaires qu'ils génèrent au profit des géants de l'Internet fleurissent, à l'image du Facebook *Data Valuation Tool*[30], petite extension développée par trois chercheurs espagnols qui permet de visualiser en temps réel la somme des profits générés par Facebook via les données personnelles de son utilisateur.

[29] Baromètre Adblocks IAB France/ Ipsos – Novembre 2016, http://www.iabfrance.com/content/presentation-de-la-v2-de-letude-ipsos-realisee-pour-liab-france-sur-les-adblocks.

[30] Source : http://www.fdvt.org/.

L'extension Facebook Data Valuation Tool permet de visualiser en temps réel les revenus générés par Facebook grâce aux données personnelles de l'utilisateur (http://www.fdvt.org/)

Une étude, publiée par le Ponemon Institute en 2015[31], montre que **les individus seraient prêts à partager leurs informations personnelles en échange d'une rémunération.** Leur enquête, réalisée sur un panel de consommateurs, demandait aux répondants à partir de quelle somme ils seraient prêts à partager telle ou telle information les concernant. Par cette méthodologie, les analystes ont montré qu'en moyenne les individus valorisaient une information personnelle à 19,60 $. Les informations les plus chères sont les mots de passe 75,80 $, les données santé 59,80 $, les informations de paiement $36, la situation crédits 29,20 $ et les habitudes de consommation 20,60 $. Les moins chères sont le genre 2,90 $, le nom 3,90 $ et le numéro de téléphone 5,90 $. Bien sûr la méthodologie est incomplète puisqu'elle se base uniquement sur les perceptions des individus, mais le rapport démontre une prise de conscience accrue des citoyens-consommateurs de la valeur des actifs qu'ils dilapident.

En 2016, les **« adblocks »** ont connu une progression inédite (+20 %), ce qui témoigne de la volonté des consommateurs de s'équiper de véritables « boucliers » destinés à faire barrage aux pratiques qui entravent, de leur point de vue, leurs principes ou leur liberté.
Le sentiment d'intrusivité est de plus en plus fortement ressenti par les citoyens mais paradoxalement ils comprennent l'importance de ces pratiques pour leur offrir les services qu'ils souhaitent. Comment gérer ce paradoxe ? Deux pistes de solution se profilent : la première est de **compenser les coûts psychologiques et sociaux** de la collecte

[31] Source : https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/internet-of-things-connected-life-security/.

de données par des bénéfices utilitaires ou fonctionnels ; la deuxième est de **rétribuer la donnée.**

Sur le premier point, des recherches académiques se sont déjà penchées sur la question. Elles ont tenté de mettre en évidence des pistes permettant de **réduire les sentiments d'intrusion** dans la vie privée des citoyens-consommateurs. Leur conclusion est que les individus-consommateurs sont prêts à partager leurs données personnelles et à accepter leurs multiples traitements et utilisations s'ils y trouvent un intérêt. L'obtention de certains bénéfices, et notamment des bénéfices utilitaires ou de facilité d'utilisation, légitime l'utilisation des données, même des plus personnelles. Une recherche récente[32] souligne que les entreprises disposent ainsi de deux leviers d'action : d'une part, elles peuvent proposer davantage de bénéfices aux clients pour « compenser » l'intrusion, comme une meilleure ergonomie, une plus grande fluidité, de la convivialité, et une meilleure personnalisation ; d'autre part, elles peuvent mieux gérer les problèmes de vie privée par une éducation numérique aussi bien pour les entreprises que pour les utilisateurs (chartes éthiques sur l'utilisation des informations personnelles, engagements de confiance, éducation, information et sensibilisation des citoyens-consommateurs dans un souci de transparence). Sur le second point, les réflexions sont en marche, à l'image de ce travail collaboratif, pour proposer **un modèle de rétribution du citoyen** sur ses données personnelles.

Les technologies à venir seront sans doute encore plus intrusives. En 2017, les spécialistes prévoient une utilisation accrue des outils d'intelligence artificielle, de la réalité virtuelle et des **« chatbots »,** sorte de logiciels-robots qui peuvent s'immiscer dans les conversations privées pour dialoguer avec un individu et lui proposer un service personnalisé (comme réserver ses billets de train par exemple, s'il discute d'un prochain voyage avec un ami). Une étude Oracle[33] révèle même que plus des trois quarts des marques s'appuieront sur les « chatbots » pour gérer l'expérience client d'ici à 2020 ! Ces perspectives doivent nous rappeler l'urgence d'apporter des réponses à la question de la valorisation des données des citoyens. Face à tous ces services qui projettent de générer des profits en exploitant massivement les données des citoyens-consommateurs, il est grand temps de partager équitablement les richesses et de rendre au citoyen-consommateur ce qui lui appartient.

32 BELVAUX B., HERAULT S., « *Privacy paradox et adoption de technologies intrusives. Le cas de la géolocalisation mobile* », Décisions Marketing, 74, 2014.

33 Etude Oracle, *Op. cit.*

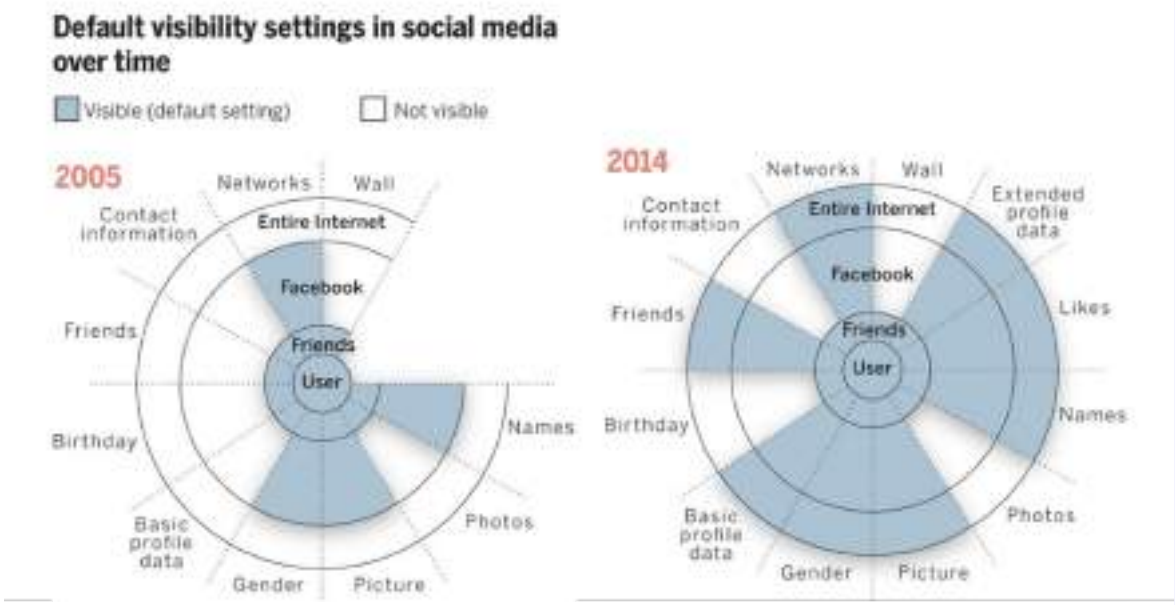
S'il apporte des réponses face à tels enjeux commerciaux et sociétaux, un modèle où la donnée est rétribuée n'est pas sans soulever de nouvelles questions. Le fait de rétribuer les citoyens-consommateurs pour leurs données pourrait en effet générer des comportements opportunistes, où ses derniers créeraient de fausses identités numériques (faux profils Facebook, adresses email « poubelle » multiples…) par appât du gain, dans le seul but de percevoir la rémunération. Ces comportements de « sabotage » pourraient alors dévaluer la donnée en rendant son exploitation moins efficace (car basée sur des éléments fabulés). Face à un tel scénario, le modèle se régulera sans doute de lui-même : l'apporteur de données est rémunéré en fonction de l'efficacité des campagnes jouées sur la base de ses données ; il ajustera en fonction l'indemnité versée au citoyen-consommateur qui aura cédé ses données. Les modèles se cherchent un temps, mais rapidement se régulent, s'ajustent, et trouvent un équilibre. Après tout, ce système poursuit un intérêt commun : protéger les modèles d'affaires. Rétribuer la donnée de façon juste, c'est préserver la croissance et poser le socle de relations saines, éthiques, et donc durables.

2.3 LE RESPECT DES CHOIX INDIVIDUELS, UN ENJEU ÉTHIQUE

Les données personnelles sont, nous venons de le voir, au cœur des enjeux gouvernementaux et commerciaux. L'application concrète va de la prolifération des drones militaires[34] au ciblage de plus en plus précis de la publicité en ligne. Dans ce but, Facebook a par exemple continuellement accru la visibilité des données disponibles dans son mode par défaut, comme le présente le graphique ci-contre[35].

[34] Science, *The End of Privacy*, 30 January 2015, p. 497.

[35] ACQUISTI A., BRANDIMARTE, L., LOEXENSTEIN G., « Privacy and human behavior in the age of information », Science, vol. 347 (6221), 2015.



Source : Acquisti, Brandimarte, Loewenstein (2015).

Au-delà de ces dimensions, se pose donc une question éthique. Quelle est la place de la confidentialité (« privacy ») dans le monde du *Big Data* [36] ? Le général américain Michael Hayden rappelait dans un entretien que l'Etat américain « tue sur la base de méta-données ». De leur côté, les données collectées par Google permettent de déduire un certain nombre de choses sur vos préférences individuelles[37].

Cette quantité d'information n'a plus rien de comparable, et ce d'autant plus qu'elle est commercialisée, avec une simple interaction sociale au sein d'un groupe. Il est vrai que chacun d'entre nous avons une identité publique sur la toile. Mais l'envergure des « connaissances » des GAFA sur nos comportements et préférences est sans commune mesure avec le fait qu'un groupe restreint d'individus et connu de soi ait connaissance de notre adresse, nom, prénom et quelques détails extérieurs que l'on veut bien transmettre à son interlocuteur direct.

Dans quelle mesure la vie privée est-elle un concept universel ? Doit-on la réduire sous le prétexte de la sécurité intérieure d'un pays, où la gratuité d'un service fourni à un utilisateur ?

---

[36] Pour mieux comprendre le débat d'un point de vue philosophique, on peut se référer à : https://plato.stanford.edu/entries/it-privacy/#ImpInfTecPri.

[37] Stephens-Davidowitz S., *Op. cit.*

Au-delà d'une vision utilitariste qui consisterait à mesurer le pour et le contre, peut-on encore espérer pouvoir protéger notre vie privée ?

Un argument de poids consiste à dire qu'il suffirait de ne tout simplement pas utiliser les services des géants d'Internet. Est-ce là la seule alternative alors qu'Internet est né d'une promesse de décentralisation et de diffusion de l'information ?

Ces questions éthiques sont importantes et sont aussi la source des réflexions qui ont poussé à l'écriture de ce texte, s'inspirant directement des travaux de **Jaron Lanier,** activiste et développeur américain. Nous pensons qu'un autre modèle est possible, plus éthique, en ce sens qu'il respecte au mieux les choix et libertés individuels.

- **Détails de l'approche utilisée et définitions**

Dans ce contexte, notre proposition consiste à déplacer la responsabilité de la donnée de la plateforme ou de l'entreprise vers son propriétaire. De cette responsabilité en découle des droits lui permettant de mieux protéger ses données personnelles, ou a minima, de renverser le pouvoir économique vers le consommateur des services et d'en faciliter sa protection.

Notre approche est essentiellement juridique. La question principale étant dans quelle mesure pouvons-nous associer la donnée personnelle à un droit de propriété ? Cette question ne peut être dissociée de trois autres aspects : I) pouvoir caractériser la donnée personnelle dans notre cadre juridique, II) établir la valeur de la donnée et un moyen de l'évaluer, et enfin, III) comment transférer cette donnée de manière sûre d'un acteur de la chaîne de valeur à un autre ?

Cela paraît d'autant plus important que le *Big Data* touche tous les secteurs : l'industrie, la santé, les transports, la ville, l'éducation, les services, les collectivités territoriales et même l'Etat.

Il existe un énorme enjeu non seulement économique mais aussi stratégique : le stockage massif de données et leur exploitation au centre de la nouvelle économie.

En France le *Big Data* peut représenter entre 3,6 et 7 % du PIB. Le vrai problème, comme le souligne Mme Antoinette Rouvroy, n'est pas tant le « *traitement inapproprié des données personnelles mais surtout la prolifération et la disponibilité même de données numériques*[38]. »

Le premier enjeu est bien de qualifier cette donnée, et donc de lui donner une définition suffisamment souple pour qu'elle s'intègre à un cadre d'analyse capable d'en intégrer toutes les caractéristiques.

• **Quelles données prendre en compte ?**

Nous donnons une définition large de la donnée :

> La donnée première est produite par le citoyen et recouvre ses **données d'identité** (nom, prénom, date et lieu de naissance, domicile, statut) et ses **données sensibles** (d'orientation sexuelle, de santé, d'appartenance religieuse ou à des groupes politiques).

On peut la définir ainsi : toute information stockée, enregistrée sur un support numérique, liée à une personne physique la permettant de la distinguer, issue de la personne elle-même ou de ses objets connectés ou de son humanoïde dont elle a le contrôle (définition de l'auteur). Ces données sont au cœur de l'économie de la donnée.

> La donnée générée est celle qui est collectée par diverses entités ayant une démarche lucrative ou non (sites web, fournisseurs d'accès à Internet, plateformes, entreprises du e-commerce, institutions, associations, ONG) par des traqueurs, des cookies, etc. sur la base de la donnée première. Cela concerne les **données de consommation** (habitudes d'achat), ou encore les données financières (moyens de paiement, Etat des prêts et financement).

[38] ROUVROY A., « *Big Data : l'enjeu est moins la donnée personnelle que la disparition de la personne* », Le Monde/The Conversation, 22 janvier 2016.

La donnée agrégée est celle qui est analysée selon un objectif précis et utilisée à des fins de Méga Data, sur la base de la donnée générée. Le citoyen seul ne peut pas le faire et cette tâche est laissée à des entreprises privées, grands groupes, parfois monopolistiques.
C'est la vitesse et la capacité de calcul de multiples sources de données qui fait la richesse du résultat.

La source est toujours la donnée première. Or c'est le citoyen qui la produit. Le citoyen est devenu le premier fournisseur gratuit de la richesse du XXIème siècle.

• La donnée et l'information : quelles différences de nos jours ?

Les lois, les règlements et directives européennes depuis 1978 ont englobé avec frénésie l'ensemble des informations et en ont fait un magma multiforme appelée donnée (Méga Data). Cette donnée rassemble en fait aussi bien des secrets de fabrique, des informations commerciales, des informations financières, des brevets, des marques, des dessins et modèles, des images, des écrits, des paroles, des informations personnelles, des liens familiaux, des croyances, des habitudes de consommation, etc.

Or le droit appréhende ces informations par des régimes différents. Qu'est-ce qui va caractériser la donnée ? Scientifiquement, on distingue :

- la donnée qualitative
- la donnée quantitative
- la donnée catégorielle
- la donnée dénombrable
- la donnée structurée
- la donnée non structurée

La donnée est toute information qui est stockée et lue par un ordinateur en format informatique, CSV par exemple. Or le langage informatique n'est pas encore à la hauteur du langage humain mais par contre, il en dépasse la force de stockage et la capacité de calcul. C'est pourquoi ce qui est intéressant économiquement est la donnée agrégée structurée.


(o.createElement("div")),f.
display:inline;zoom:1",t.
null,t}();var
(p[f]={},l||(p[f].to
acceptData(e)){var r,i,
}(n||(delete s[u].data,
(e){return e=e.node
!==e.nodeType)
rs"))){for(r=o.attr
(){b.data(this,e,n)}),
r?!0:"false"===r?!1:
e?(n=(n||"fx")+"queue",
i&&(i=n.shift(),r--,
(e,n)})})}}),b.fn.
queue:function(e){
r,i=1,o=b.Deferred(),
textarea|button|object)
(this,b.attr,e,t,arguments,
is.each(function(){try{
this[a],r=1===n.nodeType
this.each(function
e?b.trim(r):""}return
o,a=0,s=b(this),u=t,
Name=this.class

04

PARTIE 2

Créer une patrimonialité des données à droit constant.

La propriété et la régulation ne sont pas à opposer[1]. Le droit apporte un cadre légal qui va légitimer un véritable marché de la donnée personnelle. Cet encadrement garantira la bonne foi de la transaction. Nous revoyons ici le cadre juridique dans lequel s'inscrit ce nouveau marché et qui peut assurer la protection du consommateur.

Dans un premier temps, le droit français est passé en revue. Si la donnée personnelle peut être caractérisée dans le cadre du droit commun des biens, alors un droit de propriété peut émerger soit comme propriété intellectuelle, soit comme un contrat de licence.

Notre argumentation s'inscrit dans la continuité du RGPD, dans la mesure où son article 17 garantit la portabilité des données.

1. Retour sur un cadre juridique complexe et en pleine évolution

PAR NICOLAS BINCTIN

Voilà maintenant plus de trente ans que le statut juridique de l'information, de la donnée suivant le vocabulaire actuel, si les deux notions sont similaires, et son intégration dans le régime commun ou un régime spécial du droit des biens sont discutés. L'article fondateur de Catala[2] qui survenait après quelques premières ébauches[3] a nourri depuis une activité doctrinale intense[4]. Il ne semble plus nécessaire de

1 ACQUISTI A., TAYLOR C., and WAGMAN L., « The Economics of Privacy », *Journal of Economic Literature*, 54(2), 442–492, 2016.

2 CATALA P. « Ebauche d'une théorie juridique de l'information », *Rev. de droit prospectif* 1983, n°1, p. 185 ; D. 1984, chron p. 975 ; *Le droit à l'épreuve du numérique*, Puf 1998, p. 224 (seule cette dernière version est utilisée pour les développements suivants) ; même auteur, « La propriété de l'information », *Mélanges Raynaud*, Dalloz-Sirey 1985, p. 97.

3 Notamment, LECLERCQ P., « Essai sur le statut juridique des informations », Ministère de la Justice, 1980 ; *L'information sans frontière*, (1980), la Doc. Française, Paris ; CHAMOUX J.-P., *Impacts économiques et juridiques de l'informatisation*, Paradoxes 1982, p. 116.

4 HILTY R., « La privatisation de l'information par la propriété intellectuelle : problèmes et perspectives. Introduction », *Revue Internationale de droit économique*, 2006/4, p. 353 ; VIVANT M., « La privatisation de l'information par la propriété intellectuelle », *Revue Internationale de droit économique*, 2006/4, p. 361 ; même auteur, « A propos des biens informationnels », JCP éd. G 1984, I, n°3132 ; LUCAS DE LEYSSAC C., « Une information seule est-elle susceptible de vol ou d'une autre atteinte juridique aux biens ? », D. 1985, p. 43 ; DEVÈZE J., « Le vol de " biens informatiques" », JCP G 1985, I, 3210 ; A. Piédelièvre « Le matériel et l'immatériel. Essai d'une approche de la notion de bien », *Les aspects du*

reprendre et discuter ces différentes contributions[5]. Relevant le caractère fuyant de cette notion, tout comme son caractère central dans la société de l'information, on peut suivre Catala et retenir que *« l'information est d'abord expression, formulation destinée à rendre un message communicable (et) est ensuite communiquée ou peut l'être »,* un *« message quelconque », « intelligible »* et *« communicable »*[6]. L'information captée et communiquée prend la forme d'une chose[7] qui présente un intérêt juridique par sa capacité à circuler, à être exploitée sous toute forme[8]. L'hétérogénéité du sous-jacent est sans influence sur l'analyse que l'on peut porter à cette information, ce sera des résultats sportifs ou boursiers, des données comportementales de consommateurs, ou de toutes populations humaines ou non étudiées sur un plan scientifique, médical ou sociologique, telles les données collectées au travers des cartes de fidélité des consommateurs.

La donnée, mise en cohérence, en particulier au regard de sa source ou dans un traitement de masse dans la logique du *Big Data*, prend un sens pour celui qui la reçoit et présente un intérêt, dans le sens le plus large possible, non sans faire écho à la notion d'intérêt développée par la Cour européenne des droits de l'Homme pour analyser le domaine de la propriété du 1er protocole additionnel. Enfin, le droit pénal, dans toute sa rigueur, s'en empare[9].

L'approche juridique de la donnée est d'autant plus délicate que cette chose connaît, appréhendée au travers de la théorie de la valeur comme lecture de la chose appropriée, un intérêt évolutif. Savoir qu'une personne anonyme aime manger du chocolat est une donnée sans intérêt. En revanche, savoir que, dans une population donnée,

droit privé en fin du XXème siècle, Mélanges Michel de Juglart, Montchrestien 1986, p. 55 ; GEIGER C., « La privatisation de l'information par la propriété intellectuelle. Quels remèdes pour la propriété littéraire et artistique », *Revue Internationale de droit économique*, 2006/4, p. 389 ; LECLERCQ P., « L'information est-elle un bien ? », *Droit et informatique*. *L'hermine et la puce*, Masson, 1992 p 91 ; GALLOUX J.-C., « Ébauche d'une définition juridique de l'information », D., 1994, chr., p. 229 ; MALLET-PUJOL N., « Appropriation de l'information : l'éternelle chimère », D. 1997, Chron. 330 ; E. Daragon, *« Etude sur le statut de l'information »*, D. 1998, chron. p. 63 ; J. Passa, « La propriété de l'information : un malentendu ? », *Droit & Patrimoine*, mars 2001, p. 64.

5 Voir VIVANT M., « La privatisation de l'information par la propriété intellectuelle », *op. cit.*

6 CATALA P., « Ébauche d'une théorie juridique de l'information », *op. cit.*

7 GALLOUX J. -C., « Ébauche d'une définition juridique. », op. cit. ; contra, W. Dross, Droit civil – Les choses, LGDJ 2012, n° 483-1.

8 PASSA J., « La propriété de l'information », op. cit., l'information comme « action consistant à communiquer à un public des faits ou des opinions ».

9 Voir notamment, Cass. crim., 16 nov. 2011, n° 10-87.866, à paraître, D. 2012. 137, obs. M. Léna ; *AJ pénal* 2012. 163, obs. J. Lasserre Capdeville ; RSC 2012. 169, obs. J. Francillon ; *RTD com*. 2012. 203, obs. B. Bouloc ; JCP G 2012. II. 322 ; voir aussi, P. Berlioz, « Quelle protection pour les informations économiques secrètes de l'entreprise ? », *RTD com*. 2012.263.

X personnes aiment manger du chocolat constitue une information utile.

L'agglomération de données crée une nouvelle donnée — ou information — qui présente un intérêt important, sans pour autant voir sa nature juridique évoluer. L'emprise juridique sur cette chose, hors des régimes spéciaux du droit de propriété, dont la propriété intellectuelle, est délicate à mettre en œuvre, même si l'on suit là aussi l'analyse de Catala et une application du droit commun des biens tant civil que pénal.

Ainsi, l'appropriation de la donnée s'effectue essentiellement par le biais de son traitement, qui débute dès la collecte, et par sa mise en valeur, notamment au travers de son tri au sein de bases de données et son analyse algorithmique. Le concept « collecte-formulation »[10] reste d'actualité, il a même été clairement consacré par l'émergence du régime de droit voisin de producteur de bases de données par la directive de 1995[11]. La donnée libre d'accès est un *res communis* offerte à l'observation de tous, mais toutes les données n'ont pas cette nature, notamment celles collectées auprès de personnes physiques.

Le statut juridique des données signifie que l'on dépasse le simple stade du fait pour entrer dans le cadre d'une appréhension juridique de celles-ci. Qu'elle soit ou non appropriable en tant que telle, la donnée peut être contrôlée par celui qui la collecte. Si la collecte est libre et ouverte, l'utilisation de l'information collectée est, en revanche, soumise, pour tout ou partie, à la volonté du collecteur. Ce n'est pas le statut de la donnée qui change, c'est le stade de traitement de celle-ci qui évolue. Une même donnée, à des stades différents, peut juridiquement être traitée différemment, éventuellement au désavantage de son émetteur initial : par exemple, quand l'accès à la donnée n'est pas public et libre mais est attaché à une personne. Ainsi, on peut distinguer des données de flux, prises dans une globalité (telle la circulation routière, la consommation d'eau ou d'électricité, la fréquence des moyens de transport en commun), des données personnelles, relevant directement de l'activité d'une personne donnée.

Dans cette seconde catégorie, la collecte de la donnée suppose l'accès à la personne et l'accord de cette dernière pour collecter les données. La collecte s'effectue alors sur une base contractuelle. Elle doit aussi s'effectuer en conformité avec la protection des

[10] CATALA P., Ebauche d'une théorie juridique de l'information, *op. cit.*, notamment p. 234.

[11] Directive 96/9/CE du 11 mars 1996 concernant la protection juridique des bases de données, *JOUE n° L 077, 27 mars 1996 p. 20.*

personnes, et notamment le règlement 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[12].

Après avoir envisagé la nature des données (I), nous reviendrons sur leur appropriation (II) et l'influence du pouvoir du collecteur sur leur statut (III).

1.1 LE STATUT JURIDIQUE DE LA DONNÉE DANS LE CONTEXTE DU DROIT FRANÇAIS

Le cadre légal de la donnée permet de définir des éléments de statut pour cette dernière en fonction de la nature de la donnée en cause. Deux grandes catégories de données émergent pour déterminer le régime de celles-ci en fonction de leur nature. D'une part, la donnée dont l'importance va varier dans le temps, c'est-à-dire l'enjeu durable ou éphémère de la donnée (A) ; et d'autre part, le caractère sensible de la donnée (B), principalement au regard de l'Etat de la personne.

Ces deux catégories de données permettent de dégager des éléments de régimes important, tant pour les conditions de collecte que pour les conditions de communication de ces données.

A / Donnée éphémère ou durable

P. Catala l'avait largement évoqué, et le XXIème siècle le confirme chaque jour, la question de l'appropriation des données est le corollaire de l'enjeu économique que celles-ci représentent. La valeur de la donnée conduit à s'interroger sur sa réification. Toutes les données n'ont pas les mêmes caractéristiques et la même évolution de leur valeur dans le temps, ce qui influence leurs statuts juridiques. Une approche catégorielle des données est proposée par le législateur, souvent sous l'angle de lois de police afin de contrôler leur circulation. Au regard de l'appré-

[12] Règlement (UE) 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).

hension temporelle des données, certaines données ne présentent d'intérêts économiques que dans l'instant où elles sont émises et potentiellement collectées et diffusées, alors que d'autres peuvent présenter un intérêt beaucoup plus durable, même si l'idée de perpétuité de l'intérêt est largement absente.

La donnée éphémère se caractérise par un enjeu instantané et par le nécessaire contrôle de sa divulgation. La collecte est contrôlée et la divulgation maîtrisée, principalement par le biais de systèmes de responsabilité, mais aussi avec l'émergence de quasi-appropriation. Deux secteurs illustrent cette première catégorie : le secteur du jeu de hasard et des paris et celui de l'information sur les marchés financiers. Dans les deux cas, les données ne sont pas appropriées en tant que telles, mais le législateur intervient pour encadrer leurs conditions de collecte et de diffusion.

Pour le cas des jeux et paris, la loi relative à l'ouverture à la concurrence et à la régulation du secteur des jeux d'argent et de hasard en ligne[13] régule le sort des données car elles présentent de nombreux risques sociaux. Le législateur ne vise pas de façon unifiée les jeux d'argent mais distingue précisément les paris en ligne et les jeux de hasard. Pour s'en tenir à la première catégorie qui repose principalement sur les paris sportifs, on constate qu'une donnée banale, non appropriable en tant que telle, un résultat sportif, intègre un cadre légal spécifique en raison de sa nature et de l'enjeu économique qu'elle représente. La lutte contre la fraude occupe une place importante dans le dispositif légal.

Le point marquant de la loi de 2010 est qu'elle met en place un monopole de collecte et de diffusion de ces données en insérant dans le Code du sport une série de dispositions sur ces informations. Ainsi, l'article L333-1-1 étend le droit d'exploitation des fédérations sportives et organisateur d'événements sportifs en incluant le droit de consentir à l'organisation de paris sur les manifestations ou compétitions sportives.

---

13 Loi n° 2010-476 du 12 mai 2010 relative à l'ouverture à la concurrence et à la régulation du secteur des jeux d'argent et de hasard en ligne, JORF, 13 mai 2010, p. 8881.

C'est un mouvement de réification de la donnée afin d'en assurer un contrôle. La donnée ne présente aucune caractéristique pour être appropriée en tant que telle mais elle est porteuse d'enjeux économiques majeurs. Les outils de la propriété sont sollicités pour la mise en place d'un cadre légal. Les données sur les paris n'ont qu'un intérêt éphémère, le temps du dénouement du pari, c'est donc uniquement dans ce laps de temps que le législateur encadre la prise de paris et la diffusion des données sur le résultat. La réification de la donnée est d'ailleurs très limitée puisque l'article L. 333-1-2 du Code du sport dispose que *« les fédérations sportives et organisateurs de manifestations sportives ne peuvent ni attribuer à un opérateur le droit exclusif d'organiser des paris ni exercer une discrimination entre les opérateurs agréés pour une même catégorie de paris ».* L'absence de pouvoir d'exclusivité traduit que l'enjeu ne porte que sur le cadre de collecte et de diffusion de données en tant que telle, afin de répondre à des objectifs d'intérêt général.

Cette analyse se trouve confirmée par le dernier alinéa de cet article qui limite le caractère vénal de la donnée en tant que telle. Si l'autorisation de collecter et diffuser la donnée pour des paris sportifs peut ouvrir **le droit à une rémunération,** celle-ci doit être déterminée en tenant compte notamment des frais exposés pour la détection et la prévention de la fraude. Le législateur lutte contre la fraude que peut générer l'exploitation de ces données en encadrant les conditions de la collecte et de la diffusion de celle-ci par une catégorie d'opérateurs. La donnée sportive est libre[14], mais certaines exploitations de données sportives appellent un cadre légal spécifique. L'utilisation de données influence le régime juridique de celles-ci. En revanche, passé le dénouement de l'opération de jeu, le statut de la donnée sportive retrouve celui de la donnée en général, sans aucun cadre spécifique, le diptyque collecte/diffusion suffit.

On retrouve une même approche pour les données influençant le fonctionnement des marchés financiers. Il n'est pas nécessaire de reprendre dans le détail l'ensemble des dispositions devant assurer la lutte contre les délits d'initiés[15]. Comme dans le cas précédent, la collecte est contrôlée, les personnes informées limitées et supportant des obligations et une responsabilité forte, la diffusion est strictement encadrée. Ce n'est qu'une fois l'opération publiée que le régime spécial de ces données est écarté.

14 Cass civ 1, 6 fév. 1996, France 3 c. FOCA, *Bull. civ.* 1, n° 70, p. 46.

15 Voir art. L. 465-1 et sq. C. Mon. Et Fin.

Lorsque la donnée ne présente pas un intérêt pécuniaire instantané (et un risque fort de fraudes), la collecte de la donnée est libre, tout comme la diffusion, sous réserve du droit des personnes.

Ces données s'inscrivent plus durablement dans le temps et leur intérêt ne vise pas le sort d'un enjeu instantané. La diffusion de ces données sera toute aussi libre et ouverte. C'est le terrain idéal pour la mise en place de bases de données dont la valeur repose non pas sur la rareté des données mais sur leur collecte, leur accumulation, et leur analyse. Plus le volume de données est grand, plus le résultat de leur traitement gagne en pertinence, terrain de valorisation non pas d'une donnée mais d'une synthèse analytique des données. L'économie de la donnée, notamment celle des moteurs de recherche et des réseaux sociaux, se situe dans ce champ. Dans cette catégorie de données, on intègre les données personnelles collectées et parfois anonymisées qui permettent d'établir des profils comportementaux. Il s'agit des informations obtenues par le biais des cartes de fidélité ou collectées au travers des cookies des ordinateurs ou des téléphones.

Ces profilages sont le fruit de l'analyse d'une multitude de données souvent anonymes, sans valeur individuelle dans la plupart des cas. L'intérêt est de connaître le comportement d'un profil social pour ensuite exploiter ce profil auprès d'opérateurs de marché afin qu'ils adaptent leurs offres et leurs communications à ces catégories. Aucun dénouement particulier n'influence la valeur de ces données, à l'inverse du pari sportif ou de l'information boursière. La réforme de la charte de confidentialité[16] de *Google* ou de *Facebook* semble aller en ce sens, offrant à cette société une capacité de profilage accrue. La société de l'information et son écosystème permettent de prendre conscience, jour après jour, de l'enjeu économique que peuvent présenter des données *a priori* banales. Les réseaux sociaux sont à la frontière de ce mécanisme de collecte de l'information anonyme, par une utilisation par strates des données volontairement fournies par les membres du réseau.

16 http://www.google.fr/intl/fr/policies/privacy/

B / Données sensibles

Le statut spécifique des données sensibles[17] — en ce sens qu'elles visent l'intimité d'une personne physique — n'est plus une révolution, la directive de 1995 ayant largement organisé au sein de l'Union européenne un mécanisme de protection et de coopération, avant d'être remplacée par le règlement 2016/679. Toutefois, depuis 1995, la capacité technique à collecter des données sensibles s'est accrue et ont émergé de nouvelles catégories d'informations sensibles, tel le pistage par GPS ou la géolocalisation des téléphones mobiles.

Dans ce cas, le régime de la donnée est guidé, non pas par son enjeu ponctuel ou durable, mais par le contenu même de la donnée. Le dénouement d'une opération ne fait pas évoluer la qualification de la donnée à la différence de la première catégorie évoquée. On constate même, avec la question de la capacité à l'oubli sur Internet, que l'approche dans le temps de ces données sensibles constitue un nouveau point de friction.

Ces données personnelles connaissent une sensibilité économique importante et des modèles économiques émergent pour leur exploitation. Le cas des réseaux sociaux en ligne est significatif. Facebook a dû adapter sa pratique aux Etats-Unis et il est probable que cette société ait à le faire en Europe. Les données personnelles circulent sur les réseaux avec une incertitude : d'une part, de la perception de la conséquence de l'acte réalisé par celui qui les met en ligne ; et d'autre part, de la capacité de l'opérateur qui offre ce service à tirer de ces données personnelles une source de revenus. Le modèle économique de la majorité des réseaux sociaux et des moteurs de recherche repose, pour une part importante, sur la mise à disposition d'un service gratuit en contrepartie d'une collecte de données de l'utilisateur pour une utilisation commerciale.

Il ne s'agit pas d'une utilisation gratuite des services mais d'un échange de valeur : requête et réseau contre données personnelles. La sensibilité économique des données personnelles devrait marquer une évolution du régime juridique de celles-ci. Le modèle de 1978 établi par la loi Informatique et Liberté, adapté à une circulation rapide et transnationale en 1995, visait en premier lieu au contrôle des données privées par les Etats, en défiance aux dérives des pouvoirs publics.

[17] On peut ajouter d'autres informations sensibles que celles visant directement une personne, notamment les informations collectées lors d'essais cliniques pour l'obtention d'une AMM, qui suivent le mécanisme de l'article R 5121-26 CSP.

L'évolution 2.0, voire 3.0, de ce cadre légal doit permettre de contrôler l'exploitation commerciale de données personnelles par des opérateurs non-Etatiques en réseau. La nature de la donnée reste la même qu'il y a plus de trente ans, mais le risque de l'utilisation est d'une toute autre nature.

D'une défiance face à l'hégémonie et l'arbitraire de l'Etat, on doit faire face à une défiance face aux opérateurs économiques en ligne. Cette sensibilité économique des données personnelles n'écarte pas le risque liberticide. Un arrêt de la CEDH de juillet 2012[18] est venu rappeler ce cadre. La Cour a précisé la nature des données personnelles pouvant être collectées lors d'opérations de visite et saisie ordonnées par un juge d'instruction dans un cabinet d'avocats. Sur le fondement de l'article 8[19] de la Convention européenne des droits de l'Homme, la saisie de données ne doit pas violer le principe de proportionnalité de la saisie à l'objet des investigations.

CEDH, *Robathin c. Autriche*, 2012

La Cour retient que l'ordonnance allait au-delà de ce qui était nécessaire pour l'enquête (le mandat n'était pas raisonnablement limité) et que les règles procédurales ne compensaient pas cette carence de l'enquête. L'intégralité des données électroniques de l'avocat avait été saisie, ce sur quoi la Cour relève que la juridiction qui a ordonné les opérations de visite et saisie **n'a pas motivé à suffisance de droit la nécessité de saisir l'intégralité des données électroniques,** non plus que le point de savoir s'il aurait été suffisant de rechercher les seuls supports concernant les données relatives à deux clients de l'avocat. Les faits de la cause étaient uniquement liés à la relation entre le demandeur et ses deux clients, de telle sorte qu'il aurait dû y avoir des raisons particulières pour permettre la recherche de toutes les autres données, eu égard aux circonstances spécifiques prévalant dans un cabinet d'avocats.

18 CEDH, le 3 juil. 2012, *Robathin c. Autriche*, requête n°30457/06.

19 Article 8 : 1. Toute personne a droit au respect de sa vie privée et familiale, de son domicile et de sa correspondance. 2. 11 ne peut y avoir ingérence d'une autorité publique dans l'exercice de ce droit que pour autant que cette ingérence est prévue par la loi et qu'elle constitue une mesure qui, dans une société démocratique, est nécessaire à la sécurité nationale, à la sûreté publique, au bien-être économique du pays, à la défense de l'ordre et à la prévention des infractions pénales, à la protection de la santé ou de la morale, ou à la protection des droits et libertés d'autrui.

1.2 APPROPRIATION DE DONNÉES PERSONNELLES PAR LE DROIT COMMUN DES BIENS

La valeur de la donnée personnelle suppose non seulement de disposer de moyens de réservation, d'appropriation, mais aussi de moyen de défense. Ainsi, au-delà de la protection des données personnelles par le RGPD, qui suit une logique essentiellement statique, il faut passer en revue les solutions du droit des biens, et éventuellement des matières connexes, pour les appliquer à la donnée personnelle dans une approche dynamique, patrimoniale.

On envisagera la donnée personnelle comme objet de propriété (A), puis les conséquences sur son exploitation et sa défense (B).

A / La donnée personnelle, objet de propriété

L'appropriation de la donnée par le droit commun des biens appelle d'envisager la notion de « donnée personnelle » avant de solliciter les mécanismes du droit commun des biens qui permettent cette appropriation. C'est l'emprise possessoire conférée par le secret ou le maintien dans l'intimité de la personne - le contrôle de la donnée par son émetteur initial - qui permet de déduire qu'un droit de propriété s'exerce. Pour l'appropriation des biens, l'articulation entre le droit commun et le droit spécial impose une méthodologie adaptée issue de l'adage traditionnel, *speciala generalibus derogant*[20].

Dans un premier temps, il est nécessaire de vérifier l'appropriabilité de la chose par le droit spécial. Si cette solution est possible, elle écarte le droit commun des biens sauf si le droit spécial offre une option au possesseur de la chose. En effet, l'appropriation possible par le droit spécial s'impose mais le droit spécial peut laisser un champ d'opportunité au possesseur du bien intellectuel et permettre, par le jeu d'une option, à ce dernier de revenir vers le droit commun des biens.

Il faut distinguer le possesseur qui renonce à revendiquer une propriété, de celui qui cherche à obtenir une appropriation par le droit commun. Dans le premier cas, la chose ne sera pas appropriée,

[20] Voir notamment CARBONNIER, *Droit civil – Introduction*, 27ème éd. Puf 2002, n° 107, p. 208 ; CORNU, *Droit civil – Introduction au droit*, 13ème éd. Montchestien 2007, n° 329.

elle pourra être librement exploitée par tout intéressé : elle constitue une *res communes*. Par exemple, en présence d'une invention, le régime spécial du droit des brevets permet l'appropriation du bien intellectuel si des critères de fonds sont réunis. Toutefois, même si ces critères sont réunis, l'inventeur n'est pas obligé de recourir au droit des brevets pour approprier son bien. Suivant l'option dont il dispose, il peut aussi le faire par une autre voie : celle du secret. L'invention sera alors appropriée par le droit commun des biens, sous réserve d'en conserver une emprise possessoire offerte par le secret. Il peut enfin simplement divulguer son invention avant toute demande de brevet et celle-ci devient une *res communes.* À l'inverse, le droit d'auteur n'offre pas une telle option : si le bien intellectuel répond au critère de fond imposé par ce régime de propriété, il est nécessairement approprié par ce biais. L'exclusion en est très incertaine.

Lorsque le droit spécial ne permet pas une appropriation ou si une option s'ouvre, et qu'il y a un désir d'appropriation, un recours au droit commun des biens est possible pour l'appropriation des biens intellectuels et, plus généralement, pour les données, dont les données personnelles.

Il faut alors exercer une emprise possessoire sur la donnée personnelle, chose incorporelle, permettant d'exercer efficacement une propriété.

L'emprise possessoire en droit spécial et en droit commun.

La possession des choses incorporelles est consacrée par le droit spécial des biens, notamment au travers de **la possession personnelle antérieure**[21], ou encore pour les valeurs mobilières[22]. Cette possibilité doit s'étendre à toutes les choses incorporelles dès lors qu'une emprise possessoire est possible. Seul le secret, ou le contrôle de l'accès, permet une emprise possessoire efficace. À défaut de maîtrise possessoire, la divulgation supprime tout exercice exclusif de la propriété de droit commun et fait perdre à celle-ci son caractère essentiel. Il n'y a alors plus de propriété mais une jouissance non exclusive d'une chose commune.

L'emprise possessoire - le contrôle - doit être comprise comme le pouvoir de fait s'exerçant sur la donnée personnelle. La possession[23] est le constat juridique d'une situation de fait qui assure au possesseur une emprise sur l'objet de sa possession indépendamment de toute appropriation. En droit commun, cet aspect matériel de l'appropriation se prolonge dans les fonctions probatoire et acquisitive de la possession. Ainsi que le relève M. Bergel, si la maîtrise d'un bien se conçoit aussi bien pour une chose corporelle que pour un bien incorporel et si l'article 2128 du Code civil définit la possession comme « *la détention ou la jouissance d'une chose ou d'un droit* », alors elle peut concerner des données personnelles mais son exercice pratique est moins évident et plus équivoque. La possession est fondée sur le cumul de deux éléments, le *corpus*[24] et l'*animus*[25].

---

21 Voir BINCTIN N., Propriété intellectuelle, *op. cit*, n° 608.

22 Voir BINCTIN N., « La possession des choses corporelles et incorporelles », *in Le patrimoine au 21ème Siècle : regards croisés franco-japonais*, Société de législation comparée 2012, coll. Droits Etrangers, vol. 12, p. 429.

23 COLIN et CAPITANT, *Cours élémentaire de droit civil français*, 11ème éd. Dalloz 1947 par Julliot de La Morandière, t. 1, n° 1162 et sq. ; TERRÉ F. et SIMLER Ph., *Droit des biens*, 9e éd. Dalloz 2014, n° 138 et sq. Vocabulaire juridique, op. cit., V° Possession : « Pouvoir de fait exercé sur une chose avec l'intention de s'en affirmer le maître ». Cette définition implique donc l'existence d'une chose. Sur la reconnaissance d'un droit réel au possesseur : CHAUVEAU, « Classification nouvelle des droits réels et personnels », *Rev. Crit.* 1931.539, affirme que « le possesseur a pouvoir sur la chose comme le propriétaire… ».

24 *Corpus :* ensemble des actes matériels d'utilisation exercés sur la chose. En ce sens : TERRÉ F. et SIMLER Ph., *Les Droit des biens*, 9e éd. Dalloz 2014, n° 142.

25 *Animus :* volonté de se comporter en propriétaire.

Une donnée personnelle ne peut être objet d'une possession que si celui qui s'en prétend possesseur détient ces deux éléments : le *corpus* et l'*animus*. Le *corpus* de la possession consiste dans des actes matériels accomplis par le possesseur sur la chose[26]. Ces actes matériels sont notamment, au sens de l'article 2228 du Code civil, des actes de détention — la chose doit être soumise à la puissance, au contrôle, du possesseur — et des actes de possession — c'est-à-dire l'utilisation économique de la chose. Le secret comme le contrôle permettent la détention comme la possession. Ces deux types d'actes peuvent être accomplis sur une donnée, ce qui prouve l'existence de son *corpus*.

Le pouvoir factuel réside dans la capacité à la maintenir sous le sceau du secret ou d'en contrôler l'accès. L'*animus* est indépendant de l'objet de la possession, il faut tenir cet élément pour acquis car il ne relève que du comportement du possesseur, il ne peut y avoir de rapport possessoire sans volonté[27]. Pour la donnée personnelle, la volonté de la conserver n'appelle pas de discussion, l'*animus* est donc par hypothèse établie. *« En matière mobilière, c'est par la possession que le propriétaire affirme pratiquement sa souveraineté sur la chose en la soustrayant de l'appréhension d'autrui[28] ».*

Ainsi, si la donnée personnelle ne trouve pas une appropriation par le droit spécial des biens et qu'elle est soumise à une emprise possessoire, il est alors libre pour le possesseur de retenir une appropriation du bien par le droit commun, à condition de maintenir en vigueur les éléments d'emprise possessoire. Elle est alors traitée comme un bien approprié par le droit commun.

Dès lors, et c'est la dernière étape permettant d'envisager l'application du droit commun des biens en présence d'une chose incorporelle, la possession vaut titre de propriété. Cela permet d'affirmer que la possession d'une donnée grâce au secret, ou au contrôle, offre au possesseur la propriété de cette donnée, en application du droit commun des biens[29].

---

26 En ce sens, CARBONNIER, *Droit civil – Les biens*, 19ème éd. PUF 2000, n° 119 p. 203.

27 En ce sens, JHERING, *Etudes complémentaires de l'esprit du droit romain – Du rôle de la volonté dans la possession*, tome III, 2ème éd. A. Marescq aîné, Paris 1891, p. 17 et sq.

28 ZENATI-CASTAING F.et REVET Th., Les biens, *op. cit.*, n° 194.

29 L'application du mécanisme de l'article 2276 du Code civil aux incorporels nourrit un débat dans la doctrine, mais doit être admise. On suit sur ce terrain les propositions de William Dross. La possession n'étant que l'exercice factuel d'un droit, elle peut porter sur des biens incorporels. Voir DROSS W., *Droit des biens*, 2e éd. LGDJ 2014, n°473.

Le secret, le contrôle, ou l'intimité permet une emprise possessoire sur le bien incorporel, dont des données, volontairement maîtrisé par une personne. Cette possession permet de constater l'exercice du mécanisme acquisitif de l'article 2276 du Code civil et en conclure que la donnée personnelle est appropriée par le droit commun des biens.

B / L'exercice de la propriété de droit commun sur la donnée personnelle

La propriété est caractérisée par son mécanisme exclusif. La donnée personnelle répond à ce phénomène central de la propriété de droit commun, la possession permise par le secret ou le contrôle est « une forme d'expression de la volonté du propriétaire d'exclure »[30].

L'exclusivité de la propriété assure que son propriétaire puisse jouir de son bien et écarter les autres de celui-ci[31]. L'exclusivité ne signifie pas que le propriétaire est seul à avoir le bien en cause. Par son comportement, le propriétaire entend se réserver les utilités de la chose. Cette exclusivité est protégée par les moyens de défense dont dispose le propriétaire contre les atteintes à son bien. Le pouvoir exclusif permet de retirer tous les avantages du bien. *« C'est parce qu'il jouit du pouvoir souverain d'interdire sa chose à autrui que le bénéficiaire bénéficie de toutes les utilités qu'elle est susceptible de procurer, et non en vertu de prérogatives spéciales qui seraient inhérentes au droit de propriété[32] »* .

La défense de l'exclusivité s'effectue en droit des biens notamment par la capacité à revendiquer un bien auprès d'un tiers. Pour la donnée personnelle, la jurisprudence a admis à plusieurs reprises l'application d'une action en revendication mobilière en présence d'un tiers s'étant illégitimement approprié le bien d'autrui[33]. On doit admettre, par analogie, les mêmes solutions pour les données personnelles.
Le constat d'un droit de propriété de droit commun permet de tirer des conséquences pour l'exploitation et la défense du bien.

---

30 ZENATI-CASTAING F. et REVET Th., *op. cit.*, n° 194.

31 DANOS V. F., *Propriété, possession et opposabilité*, Economica 2007.

32 ZENATI-CASTAING et Th. REVET, *op. cit.*, n° 208.

33 Voir CA Versailles, 19 mai 2006, n° 04/08720, *« la soustraction d'une invention constitutive de l'un des cas d'ouverture de l'action [en revendication] est susceptible de se trouver réalisée, s'agissant de la revendication d'un bien immatériel, même sans dépossession du bien, du fait d'une atteinte à la jouissance de la chose sur laquelle s'exerce l'emprise de la propriété ».*

• **L'exploitation**

Dès lors que l'on retient que le droit commun des biens s'applique à une donnée personnelle, il faut tirer toutes les conséquences de la situation de droit. L'ensemble des dispositions du Code civil s'applique *mutatis mutandis* au regard des spécificités de ce bien, notamment le régime de la copropriété[34]. En effet, visant l'article 16 du Code de procédure civile, la Cour de cassation a retenu qu'en présence d'un bien intellectuel secret, un savoir-faire, la cour d'appel qui appliquait à ce dernier les dispositions du Code de la propriété intellectuelle violait ce texte. Seul le droit commun pouvait régir la situation de ce bien sauf à ce que les copropriétaires aient volontairement opté pour une application du droit spécial à leur bien. L'indivision du savoir-faire, des données secrètes, n'est donc pas régie par le système de l'indivision du brevet mais par celui du droit commun.

La propriété de droit commun, à la différence des propriétés spéciales de la propriété intellectuelle, dure tant que l'emprise possessoire perdure. Elle est potentiellement perpétuelle, du moins, elle est efficiente aussi longtemps que le secret ou le contrôle est maintenu.

Suivant les logiques du droit commun des biens, la donnée personnelle peut être vendue, ou louée comme tout bien pour lequel le propriétaire dispose d'un *usus*, d'un *fructus* et d'un *abusus.* Le juge a aussi appliqué à la donnée appropriée par le secret d'autres mécanismes du droit commun des biens, telle la revendication mobilière[35]. La cour d'appel de Paris impose aussi, avec raison, l'application des garanties légales d'éviction en présence d'une cession de données personnelles[36], confirmant que le contrat emporte bien transfert d'un droit de propriété, une propriété de droit commun et non de droit spécial.

La loi République Numérique, en instaurant un droit à la portabilité des données au profit des internautes qui souhaitent changer de prestataire de services numériques, consacre aussi nécessairement le contrôle des données par l'émetteur initial et donc l'appropriation de ces données et l'organisation de leur circulation. Elle confirme l'appropriation des données malgré elle. Ils peuvent récupérer leurs données (courriels, photos, préférences musicales, etc.) dans un format ouvert et facilement réutilisable, afin de les transférer vers un nouveau prestataire. La loi devance certains aspects du règlement européen du

34 Cass. com., 7 déc. 2010, n°10-30034.

35 CA Versailles, 19 mai 2006, Rep. 04/08720, *op. cit.*

36 CA Paris, 11 avril 2013, n° 12/21643, voir *contra*, J. Passa, RDC déc. 2014, p. 739.

27 avril 2016 sur la protection des données personnelles, applicable en mai 2018. Elle instaure un droit à la libre disposition de ses données numériques personnelles :

• **Le droit à l'oubli** numérique pour les mineurs : un mineur pourra obtenir plus facilement et plus rapidement l'effacement d'un contenu en ligne le concernant ;

• **Le droit à la mort numérique :** chacun pourra de son vivant exprimer ses volontés sur la conservation et la communication de ses données après son décès ou demander leur effacement ;

• **L'application stricte de la règle du secret des correspondances privées,** quel que soit le vecteur ou la technologie de communi cation utilisé (mails, réseaux sociaux, etc.).

Les compétences de la CNIL pour la protection des données personnelles sont élargies et son pouvoir de sanction renforcé ; le plafond maximal de ses sanctions passe de 150 000 à 3 millions d'euros. Ces actes sont des actes de disposition de biens propres à une personne qui peut exercer des prérogatives et imposer des choix sur l'utilisation ou la circulation de ses données.

Etat de la réflexion sur la patrimonialisation des données en droit interne et en droit de l'Union européenne.

Le statut du *text and data mining*, tant en droit interne qu'en droit de l'Union européenne nourrit les mêmes réflexions sur la patrimonialisation des données. Cette forme d'exploitation de bases de données est sujette à d'importantes discussions pour déterminer si elle relève de l'emprise du producteur de la base de données ou si elle en est exclue.

En droit interne, l'article L.122-5 10° permet désormais les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale. Il s'agit de l'insertion en droit français d'une exception de *text and data mining*, de portée limitée. Un décret fixe les conditions dans lesquelles l'exploration des textes et des données est mise en œuvre, ainsi que les modalités de conservation et de communication des fichiers produits au terme des activités de recherche pour lesquelles elles ont été produites. Ces fichiers constituent des données de la recherche.

En droit de l'Union européenne, la proposition de directive sur certaines utilisations autorisées d'œuvres en faveur des aveugles, et modifiant la directive Info Soc de septembre 2016, intègre une exception pour les reproductions et extractions effectuées par des organismes de recherche, en vue de procéder à une fouille de textes et de données sur des œuvres ou autres objets protégés auxquels ils ont légitimement accès à des fins de recherche scientifique. Cette exception serait limitée par des mesures destinées à assurer la sécurité et l'intégrité des réseaux et bases de données où les œuvres ou autres objets protégés sont hébergés. Elle serait aussi limitée par un encouragement à une organisation d'un commun accord des bonnes pratiques concernant l'application des mesures.

2.1 • La défense

La défense est un enjeu central pour la donnée appropriée. La directive Secret d'affaires va dans ce sens : l'essentiel de ses dispositions vise le procès et les sanctions, afin de permettre de donner une efficience aux sanctions de la violation du secret et donc à la maîtrise des données appropriées. On pourrait aussi ajouter les solutions du règlement 2016/679.

Le droit spécial compte plusieurs dispositions permettant d'envisager une sanction de la violation d'une maîtrise d'une donnée (voir encadré ci-contre).

Toujours au chapitre du droit spécial, la donnée trouve une protection en fonction de l'endroit où elle est stockée. Ainsi, si la donnée est intégrée dans une base de données, le régime du droit de producteur de base de données s'applique pour préserver cette dernière. L'extraction qualitativement ou quantitativement substantielle ouvre la possibilité de mobiliser l'arsenal légal attaché à ce régime de propriété et obtenir tant des mesures avant dire droit que des sanctions civiles ou pénales. Une procédure analogue à l'action en contrefaçon est ouverte. Dans le même esprit, si l'accès à une donnée est le fruit de la violation d'un système d'information, il est possible d'agir contre l'auteur de ce méfait en droit pénal[37]. Ce n'est plus la donnée telle quelle qui est en cause mais la façon dont on se l'est procurée. Le fait d'accéder ou de se maintenir, frauduleusement, dans tout ou partie d'un système de traitement automatisé de données est puni de deux ans d'emprisonnement et de 30 000 euros d'amende. Ainsi, le droit pénal tourne autour de la donnée et renforce l'image d'un régime d'appropriation de celle-ci par les outils de défense qu'il procure.

Plus encore que ces mesures spéciales, il est retenu une application du droit commun pénal pour l'atteinte à un bien soumis au droit commun. En effet, si l'atteinte à une donnée n'emporte pas le dessaisissement de cette dernière pour le propriétaire, ce qui écarte la possibilité d'agir pour vol, le propriétaire d'une donnée peut engager une action en abus de confiance, incrimination visant les atteintes aux biens.

---

37 Art 323-1 CP.

La sanction de la violation d'une maîtrise d'une donnée en droit.

Le Code de la propriété intellectuelle aborde les données sous l'angle des sanctions en cas de violation du secret de fabrique. L'article L. 621-1 renvoie au Code du travail, article L. 1227-1, pour définir les peines frappant la violation des secrets de fabrique. Cet article prévoit que le fait pour un directeur ou un salarié de révéler ou de tenter de révéler un secret de fabrication est puni d'un emprisonnement de deux ans et d'une amende de 30 000 euros.

Les sanctions de l'atteinte aux données secrètes ou contrôlées sont renforcées en présence d'un enjeu de défense. L'article 413-7 du Code pénal dispose qu'est puni de six mois d'emprisonnement et de 7 500 euro d'amende le fait, dans les services, établissements ou entreprises, publics ou privés, intéressant la défense nationale, de s'introduire, sans autorisation, à l'intérieur des locaux et terrains clos dans lesquels la libre circulation est interdite et qui sont délimités pour assurer la protection des installations, du matériel ou du secret des recherches, études ou fabrications.

Sur le fondement de cette disposition, un décret[38] institue une catégorie particulière de zones protégées, les « zones à régime restrictif » (C. pén., art. R. 413-5-1). Un arrêté non publié du Premier ministre détermine, au sein des secteurs scientifiques et techniques concernés, la liste des spécialités dont les données émises sont susceptibles d'être détournés à des fins de terrorisme ou de prolifération d'armes de destruction massive et de leurs vecteurs.

38 Décret n° 2011-1425, 2 nov. 2011 relatif à la protection du potentiel scientifique et technique de la nation, *JORF* 4 nov. 2011, p. 18562.

L'abus de confiance en droit.

Suivant l'article 314-1 CP, l'abus de confiance est « *le fait par une personne de détourner, au préjudice d'autrui, des fonds, des valeurs ou un bien quelconque qui lui ont été remis et qu'elle a acceptés à charge de les rendre, de les représenter ou d'en faire un usage déterminé* ».
En application de cette disposition, la chambre criminelle a retenu que **constitue un abus de confiance le fait pour un salarié en charge de mettre au point un projet de borne informatique, d'en disposer comme d'un bien propre au profit d'un tiers,** alors que dès sa réalisation, le projet était la propriété de son employeur, et qu'il n'en était que le détenteur[39]. Cette solution a été confirmée par un arrêt de la chambre criminelle du 22 octobre 2014[40].

Dans l'affaire ayant donné lieu à cet arrêt, un salarié avait informé son employeur, un cabinet de courtage d'assurances, de son intention de démissionner de son emploi de chargé de clientèle en vue de rejoindre un autre cabinet de courtage. Durant la période de préavis, un contrôle interne établit que le salarié démissionnaire avait capté un grand nombre de données issues d'une base informatisée à usage interne de la société, protégée par une charte de confidentialité signée par tous les salariés. Les données de cette base répondent à la définition des données personnelles. **Le salarié est alors poursuivi pour avoir détourné, au préjudice de son employeur, plus de trois cents fichiers informatiques qui ne lui avaient été remis qu'à charge d'en faire un usage déterminé,** conforme à la charte informatique interne proscrivant l'extraction de ces documents de l'entreprise.

La cour d'appel de Bordeaux avait retenu sa culpabilité dans un arrêt du 5 février 2013. La Cour de cassation a suivi cette analyse. Elle rejette le pourvoi de ce dernier et retient « *que le prévenu a, en connaissance de cause, détourné en les dupliquant, pour son usage personnel, au préjudice de son employeur, des fichiers informatiques contenant des informations confidentielles et mis à sa disposition pour un usage professionnel* », ce qui caractérise tous les éléments, tant matériel qu'intentionnel, du délit d'abus de confiance.

La Cour de cassation poursuit cette construction en retenant qu'une personne n'ayant pas entendu donner à une autre la disposition des documents personnels dont elle était propriétaire, le **libre accès à des informations personnelles sur un réseau informatique d'une entreprise n'est pas exclusif de leur appropriation frauduleuse** par tout moyen de reproduction[41].

39 Cass. crim., 22 sept. 2004, *Bull. crim*. n° 218, *Dr. penal 2004*, comm. 179, obs. M. Véron ; D. 2005, j. 411, note B. de Lamy ; *RTD. civ*. 2005.164, obs. T. Revet ; *Rev. sc. crim.* 2005.852, obs. R. Ottenhof.

40 Cass. crim, 22 oct. 2014, n° 13-82.630.

41 Cass. crim., 28 juin 2017, n° 16-81113.

2.3 LE POUVOIR DE COLLECTE DE L'INFORMATION

La troisième série d'éléments qui influence le régime des données relève de la nature et du pouvoir que détient celui qui collecte la donnée. Il ressort notamment de la jurisprudence une prise en considération importante de ces éléments pour l'appréciation du pouvoir collecte/diffusion de l'information, en particulier sur la capacité à jouir d'une donnée collectée par un tiers. On pense, en premier lieu, aux jurisprudences relatives aux infrastructures essentielles. Dans l'affaire *IMS Health*[42], si la base de données relevait du droit d'auteur (et aujourd'hui elle pourrait aussi relever du droit de producteur de base de données), les données collectées et exploitées ne semblaient pas, en tant que telles, éligibles à cette appropriation. La solution retenue par la Cour de Justice n'est pas totalement étrangère à cette qualité, ce qui a conduit plusieurs auteurs à identifier une catégorie spécifique de bien ou œuvre, le bien informationnel, ou une donnée.

Le pouvoir de collecte est un élément clé du régime juridique des données. En principe, la donnée non secrète ou non contrôlée est collectée librement par tout un chacun, dans la mesure où il est possible d'accéder librement à celle-ci. La collecte qui viole la vie privée d'une personne, le contrôle qu'elle souhaite avoir de ses données, ou un secret est fautive et doit être sanctionnée. Toutefois, certaines personnes disposent d'un pouvoir de collecte permettant d'imposer la communication de données, en général grâce à l'exercice de prérogatives de puissance publique ou grâce à des monopoles de droit.

Le statut de la donnée collectée peut être influencé par le pouvoir ou la qualité juridique du collecteur (A). Ce mouvement connaît un système correctif volontaire ; sans remettre en cause le pouvoir de collecte et son privilège, la donnée sera ensuite diffusée largement, voire librement, ce qui permettra sa réutilisation par des tiers, ce qui aura aussi une influence sur l'analyse patrimoniale des données en cause (B).

42 CJCE, 29 avr. 2004, C-418/01, Rec. I-5039.

A / L'influence de la prérogative juridique justifiant la collecte

Le statut de la donnée est influencé par le pouvoir du collecteur.
Il s'agit de déterminer si le statut de la donnée diffère selon qu'elle relève d'une libre collecte ou fruit de l'utilisation de prérogatives de puissance publique. Dans ce dernier cas, quel serait l'impact sur le statut de la donnée ? On ne recherche pas le cadre légal autorisant ou contrôlant la collecte, selon les procédures instaurées depuis 1978, mais le statut de la donnée ainsi collectée.

Contribution de la CADA.

La Commission d'accès aux documents administratifs apporte un contrôle de l'accès à ces données. Toutes les données publiques peuvent être réutilisées suivant les dispositions de l'article 10 de la loi de 1978. Aucun texte de portée générale n'interdit la réutilisation de données non publiques même s'il peut exister des textes particuliers limitant l'usage susceptible d'en être fait.

Pour qu'une donnée soit regardée comme publique, il faut tout d'abord qu'elle figure dans un document administratif. L'article 10 prévoit toutefois une série d'exceptions, notamment les documents librement communicables sur le fondement de l'article 2 de la loi du 17 juillet 1978, des articles L. 124-1 et suivants du Code de l'environnement, de l'article L. 2121-26 du Code général des collectivités territoriales sont en principe des données publiques, tout comme les données contenues dans des documents produits ou reçus par les autorités administratives dans l'exercice d'une mission de service public à caractère industriel ou commercial[43].

Ce modèle a largement été renversé par la loi République Numérique qui impose le principe d'une ouverture des données publiques, sauf exception. Cette ouverture générale de données collectées, et donc leur diffusion large, peut venir contredire, dans certains cas, l'attractivité d'une patrimonialisation des données. À tout le moins, cela traduit la tension existant actuellement entre une approche par l'appropriation et une approche par la communautarisation des données. Le partage de la chaîne de valeur ne fut malheureusement guère pris en compte par l'Etat pour retenir cette solution très politique.

43 CADA, décision n° 20090221, 15 janvier 2009.

Contribution de la CJUE.

La Cour de Justice apporte une contribution importante à cette analyse du statut de la donnée collectée par le biais de pouvoirs exorbitants. Dans un arrêt du 12 juillet 2012[44], la Cour a considéré que le refus de l'Etat autrichien d'ouvrir aux opérateurs privés l'accès aux données figurant dans le registre du commerce et des sociétés pour la constitution de banques de données enrichies, ne relève pas d'une activité économique opérée par une entreprise au sens du droit de la concurrence, parce que légitime au regard de la mise en œuvre de prérogatives de puissance publique. Cette affaire est d'autant plus intéressante qu'elle a conduit la Cour à proposer une interprEtation de la directive 2003/98/CE concernant la réutilisation des données du secteur public[45].

Cette dernière énonce, à son considérant 5, que l'un des principaux objectifs du marché intérieur est de créer les conditions qui permettront de développer des services à l'échelle de l'Union. **Les données émanant du secteur public constituent une matière première** importante pour les produits et les services de contenu numérique et forment une ressource importante au fur et à mesure que les services de contenu sans fil se développent. L'amélioration des possibilités de réutilisation des données émanant du secteur public devrait notamment permettre, suivant le considérant, aux entreprises européennes d'exploiter le potentiel de ces données et de contribuer à la croissance économique et à la création d'emplois.

La directive ne contient aucune obligation d'autoriser la réutilisation de documents, la décision d'autoriser ou non la réutilisation est laissée à l'appréciation des Etats membres ou de l'organisme du secteur public concernés. La directive s'applique aux documents qui sont mis à disposition aux fins d'une réutilisation lorsque les organismes du secteur public délivrent des licences, vendent, diffusent, échangent ou donnent des informations. Ainsi, l'article 1er dispose que « *la présente directive fixe un ensemble minimal de règles concernant la réutilisation*

44 CJUE, 12 juill. 2012, aff. C-138/11, *Compass-Datenbank GmbH Vs Republik Österreich.*

45 Directive 2003/98/CE du 17 novembre 2003, concernant la réutilisation des informations du secteur public, *JOUE* L 345, p. 90, dite directive ISP.

et les moyens pratiques destinés à faciliter la réutilisation de documents existants détenus par des organismes du secteur public des Etats membres ». L'article 2, point 4 définit la réutilisation des documents du secteur public comme étant *« l'utilisation par des personnes physiques ou morales de documents détenus par des organismes du secteur public, à des fins commerciales ou non commerciales autres que l'objectif initial de la mission de service public pour lequel les documents ont été produits ».* Le conflit portait sur les données collectées dans le cadre du registre du commerce et des sociétés autrichien.

La Cour relève qu'une activité de collecte de données relatives à des entreprises, sur le fondement d'une obligation légale de déclaration imposée à ces dernières et des pouvoirs coercitifs y afférents, relève de l'exercice de prérogatives de puissance publique. Par conséquent, une telle activité ne constitue pas une activité économique[46].

Pour ce qui est de la diffusion de ces données, la Cour précise qu'une activité consistant à tenir et à rendre accessibles au public des données ainsi collectées, soit par une simple consultation, soit par la fourniture de copies sur support papier, conformément à la législation nationale applicable, ne constitue pas davantage une activité économique, dès lors que la tenue d'une base contenant de telles données et sa mise à la disposition du public sont des activités indissociables de l'activité de collecte de ces données. En effet, la collecte desdites données serait largement privée de son utilité en l'absence de tenue d'une base de données les répertoriant afin que le public puisse les consulter[47].

À la lumière de ces éléments, la Cour retient qu'un Etat qui constitue et maintient le registre du commerce et des sociétés n'agit pas en tant qu'entreprise au sens de l'article 102 TFUE. Dès lors, une telle activité relève de l'exercice de prérogatives de puissance publique et ne constitue pas une activité économique au sens du droit de la concurrence. Par suite, l'Etat n'est pas tenu d'autoriser une utilisation libre des données qu'il collecte et met à la disposition du public.

46 Pt. 40.
47 Pt. 41.

Ainsi, l'activité d'une autorité publique consistant : à sauvegarder, dans une base de données, des données que les entreprises sont tenues de communiquer en vertu d'obligations légales ; à permettre aux personnes intéressées de consulter ces données ; et/ou à leur fournir des copies sur support papier de celles-ci ; **ne constitue pas une activité économique.** Cette autorité publique ne doit pas, par conséquent, être considérée, dans le cadre de cette activité, comme une entreprise, au sens de l'article 102 TFUE.

Le fait que cette consultation et/ou cette fourniture de copies sont effectuées en contrepartie d'une rémunération prévue par la loi et non pas déterminée, directement ou indirectement, par l'entité concernée n'est pas de nature à faire modifier la qualification juridique de ladite activité. Dans la mesure où une telle autorité publique interdit tout autre usage des données ainsi collectées et mises à la disposition du public, en se prévalant de la protection *sui generis* qui lui est accordée en tant que fabricant de la base de données en question au titre de l'article 7 de la directive 96/9, ou de tout autre droit de propriété intellectuelle, elle n'exerce pas non plus une activité économique et ne doit donc pas être considérée, dans le cadre de cette activité, comme une entreprise, au sens de l'article 102 TFUE.

Le pouvoir, né de la collecte et de la capacité corrélative à devenir producteur de base de données et exercer un droit sur l'extraction et la réutilisation de ces données, n'emporte pas une modification de l'analyse de la CJUE. La Cour met fin par cette décision à l'ambition de la société *Compass-Datenbank* d'obtenir de l'Etat autrichien un transfert massif des données récentes recueillies via le registre du commerce et des sociétés et ce, en contrepartie d'une rémunération raisonnable, accompagné du droit de réutiliser ces données brutes afin de lui permettre d'offrir un service élaboré sur le fondement des données déjà accessibles à tous par le biais des agences intermédiaires. La CJUE ayant conclu que l'Etat autrichien n'agissait pas en tant qu'entreprise sur le marché, elle n'a pas eu à répondre à la dernière question préjudicielle qui lui était soumise, à savoir l'application de la théorie des facilités essentielles à l'espèce. La question était pourtant pertinente, les faits en cause semblent très proches de ceux constatés dans l'affaire *IMS*.

Ainsi, l'interdiction de la réutilisation des données contenues dans le registre du commerce et des sociétés relève de l'exercice de prérogatives de puissances publiques, et n'est pas détachable des autres activités de puissances publiques de l'Etat en cause. Une autorité publique peut légitimement considérer qu'il est nécessaire, voire obligatoire au regard des dispositions de son droit national, d'interdire la réutilisation des données figurant dans une base telle que celle en cause, afin que soit respecté l'intérêt que les sociétés et les autres sujets de droit, qui souscrivent des déclarations imposées par la loi ont, à ce que des informations les concernant ne soient pas réutilisées en dehors de cette base. La Cour semble faire référence à la nécessaire protection, sinon de secrets des affaires, du moins des données propres aux entreprises, contre une divulgation systématique et organisée.

Dans ces différentes hypothèses, la qualification juridique des données n'évolue pas, elles ne sont pas appropriées en tant que telle, mais le pouvoir du collecteur, ou la cause de la collecte[48], influence leur régime de réutilisation.

48 Voir *supra* le statut des informations incluses dans les AMM.

B / Des mécanismes correcteurs

Les mécanismes correcteurs trouvent siège dans le mouvement *Open Data*. Celui-ci réunit de nombreux pays ou organisations, notamment l'Arabie Saoudite, l'Australie, l'Autriche, la Belgique, le Canada, la Corée du Sud, les Etats-Unis, l'Union Européenne, Hong-Kong, le Kenya, le Maroc, le Mexique, les Pays-Bas, le Royaume-Uni, Singapour ou encore la Tunisie.

La France s'inscrit aussi dans ce mouvement. Une ordonnance[49] a modifié la loi de 1978 et instauré le droit pour toute personne physique ou morale de réutiliser les données publiques des administrations. Ce régime est inséré dans un chapitre II du titre Ier de loi de 1978. En accédant librement aux données publiques dont disposent les administrations, la communauté des développeurs et des entrepreneurs devrait être en mesure de créer de nouveaux usages et des services applicatifs innovants. Ce mouvement a été renforcé par la loi République Numérique de 2016.

Pour soutenir cette action, un décret de février 2011 a créé la mission *Etalab*[50] qui a pour mission de créer le portail unique interministériel data.gouv.fr destiné à rassembler et à mettre à disposition librement les données publiques de l'Etat, de ses établissements publics administratifs, des collectivités territoriales et des personnes de droit public ou de droit privé chargées d'une mission de service public. En mettant à disposition les données publiques, l'Etat s'inscrit dans une stratégie d'ouverture des données, suivant le mouvement *Open Data*, au travers de la mise en ligne de ces données sur data.gouv.fr.

Est-ce une approche en lien avec la patrimonialisation des données ? On peut y voir deux cas : soit l'ouverture est la marque du pouvoir de contrôle, et donc d'une forme d'appropriation ; soit l'ouverture est la conséquence d'un impossible contrôle juridique des données et une exclusion de toute forme d'appropriation. L'Etat ne prend pas explicitement partie sur cette question, mais il exploitait antérieurement économiquement une partie de ces données. On peut au moins soutenir que la première branche de l'alternative ne peut être exclue automatiquement.

---

49 Ordonnance n° 2005-650 du 6 juin 2005 relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques.

50 Décret n° 2011-194 du 21 février 2011 portant création d'une mission « Etalab » chargée de la création d'un portail unique interministériel des données publiques.

L'objectif des pouvoirs publics est que **la créativité des développeurs et des entrepreneurs ne se heurte pas à des cloisons juridiques perçues comme des freins au développement de l'innovation.** L'intérêt collectif prime sur le contrôle par l'Etat des données qu'il collecte. L'Etat retient le principe d'une utilisation d'une licence gratuite applicable aux données publiques mises en ligne sur le portail public, sans exclure totalement l'hypothèse de licences payantes. Le vocabulaire est ici essentiel. Parlant de licence, l'Etat retient que l'on est en présence d'un droit de jouissance qui ne peut, juridiquement, être accordé que sur des biens. La patrimonialisation de la donnée semble alors s'imposer .

La réutilisation de données diffusées sous cette licence impose au producteur, i.e. l'entité qui produit la donnée et l'ouvre à la réutilisation dans les libertés et les conditions prévues par la licence, garantit au réutilisateur un droit personnel, non exclusif et gratuit, de réutilisation dans le monde entier et pour une durée illimitée, dans les libertés et les conditions du contrat. Il est ainsi possible de reproduire, copier, publier et transmettre la donnée, la diffuser et la redistribuer, l'adapter, la modifier, l'extraire et la transformer, notamment pour créer des données dérivées. Il est aussi possible d'exploiter la donnée à titre commercial, par exemple en la combinant avec d'autres données, ou en l'incluant dans un produit ou une application.

Ces vastes possibilités sont soumises à la condition de mentionner la paternité de la donnée, à savoir sa source (*a minima* le nom du « Producteur ») et la date de sa dernière mise à jour. Le régime est ici construit en analogie avec le droit d'auteur et ne peut que renforcer l'idée d'une approche patrimonialisée des données.

La loi République Numérique renforce l'ouverture des données publiques avec la création d'**un service public de la donnée.** L'ouverture des données publiques devient la règle et non plus l'exception. Les administrations au sens large doivent publier en ligne, dans un standard ouvert, leurs principaux documents, y compris leurs codes sources, ainsi que leurs bases de données et les données qui présentent un intérêt économique, social, sanitaire ou environnemental. Cette obligation va concerner les administrations d'Etat, les collectivités locales de plus de 3 500 habitants, les établissements publics et les organismes privés chargés d'un service public, à l'exception des petites entités. L'ouverture

des données concerne aussi les algorithmes publics, de plus en plus fréquents dans les décisions administratives : par exemple pour le calcul de l'impôt ou l'affectation des élèves dans les établissements scolaires ou l'enseignement supérieur. Sauf exception, toute personne destinataire d'une décision fondée sur un traitement algorithmique pourra demander à l'administration les règles définissant ce traitement et ses principales caractéristiques. De plus, les administrations doivent publier en ligne les règles de leurs principaux traitements algorithmiques fondant des décisions individuelles.

La loi introduit la notion de données d'intérêt général. Ces données, qui recouvrent un vaste champ, sont ouvertes à tous. Sont notamment concernées les données des délégations de service public (dans les transports, l'eau, la gestion des déchets, etc.), les données relatives aux subventions publiques au-delà d'un certain seuil, les données de jurisprudence sous conditions ou encore les données de consommation d'énergie. La loi confie à l'administration publique une nouvelle mission : celui du service public de la donnée. Celui-ci est chargé de faciliter la réutilisation des principales bases de données de l'Etat par les acteurs privés ou publics en leur garantissant un niveau élevé de qualité de service. Il s'agit de construire une infrastructure nationale autour de quelques grandes bases de « données de référence » comme le répertoire SIRENE des entreprises, qui est en accès ouvert et gratuit depuis le 1er janvier 2017, ou le cadastre. Cette considération publique de la donnée impose d'y rechercher des indices pour ce qui est envisageable pour les données des personnes physiques.

2. Le RGPD, un pas dans la bonne direction ?

PAR ISABELLE LANDREAU avec la contribution de RUBIN SFADJ[51]

Adopté par l'Union européenne le 27 avril 2016, le **Règlement général** 2016/679 sur la protection des données personnelles (RGPD) entrera en vigueur le 25 mai 2018 et abrogera la Directive 95/46/CE sur la protection des données personnelles, en vigueur depuis vingt ans. Il supplantera également la loi Informatique et libertés françaises[51].

C'est dans le double objectif de réaffirmer le principe fondamental de libre circulation des données — il était devenu difficile pour les entreprises de déterminer quelle loi nationale appliquer à leurs traitements de données personnelles — et d'**offrir un niveau de protection uniforme au sein de l'Union** (tout en évitant le *dumping* éthique, le niveau de protection variant selon les Etats membres) que le Parlement et la Commission européenne ont adopté le RGPD.

D'un régime déclaratif, basé en France sur les fameuses « déclarations CNIL » que les entreprises doivent soumettre préalablement à tout traitement de fichier, **le RGPD fait passer le droit des données personnelles dans l'ère de la *compliance*** en exigeant des organisations, privées comme publiques, qu'elles soient à tout moment en mesure de démontrer aussi bien au régulateur qu'aux personnes dont elles traitent les données que leurs pratiques, leurs traitements et leurs systèmes sont conformes à un certain nombre de principes directeurs.

Ainsi, à chaque droit conféré aux personnes physiques correspond, en regard, une obligation pour l'organisation. Ce changement de paradigme s'accompagne d'une nouvelle obligation de très large portée, puisque le RGPD impose la mise en place d'un niveau de sécurité « renforcé » et « adapté au risque » relevant de la protection de la vie privée en prenant en compte non seulement la nature, la portée,

[51] Avocat au barreau de Marseille, Rubin SFADJ est co-fondateur de Proposition47 et co-auteur de *Réussir votre mise en conformité GDPR : Guide pratique*, Broché, 2017.

le contexte et la finalité du traitement de données personnelles, mais aussi les risques pour les droits et libertés des personnes, les coûts de mise en œuvre et l'Etat des connaissances (article 32). C'est donc pour les organisations **une obligation de sécuriser spécifiquement les données personnelles** en fonction des risques encourus non plus seulement par l'organisation, mais par la personne elle-même.

Notre projet de propriété des données personnelles s'inscrit dans la droite ligne des dispositions de la Loi pour une République Numérique et celles du nouveau Règlement Général sur la Protection de la donnée (RGPD) en ce qu'il consacre implicitement à tout le moins *l'usus* et désormais *l'abusus* de la donnée personnelle, notamment par des articles 26 de la loi et des articles 1, 17 et 20 du règlement.

Notre projet de propriété des données personnelles va plus loin car il organise un véritable mécanisme de *fructus* de la donnée personnelle. Il prévoit que le citoyen qui a consenti librement et explicitement (article 7 du RGPD) à l'exploitation de ses données personnelles, doit être en mesure d'en retirer les fruits (revenus) que d'autres font sur son dos…

Les articles 17 et 20 du RGPD consacrent bien que la donnée personnelle est un bien meuble incorporel puisque le cybercitoyen peut demander l'effacement de ses données (article 17) et la récupération et le transfert de ses données (article 20 : portabilité des données).

Dès 2015, la Loi pour une République Numérique entérine déjà une appropriation de ses données par le citoyen dans ses articles 1er, 12, 26, 40-1.

L'article premier dispose que « l'individu doit conserver la **libre disposition** de ses données ».

L'article 1 de la loi du 6 janvier 1978 est devenu : « L'informatique doit être au service de chaque citoyen. Son développement doit s'opérer dans le cadre de la coopération internationale. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques.

Toute personne dispose du **droit de décider et de contrôler les usages** qui sont faits des données à caractère personnel la concernant, dans les conditions fixées par la présente loi ».

L'article 12 permet la **récupération de ses données auprès d'un fournisseur** et accepte donc implicitement l'appropriation de la donnée par le citoyen, qui en aurait délégué l'*usus* que pour une période donnée au fournisseur.

L'article 26 de ladite loi inscrit donc le droit à la libre disposition de ses données personnelles.

Enfin, l'article 40-1 aménage le droit de disposer par avance de ses données après sa mort, à une personne de son choix : « Toute personne peut définir des directives relatives à la conservation, à l'effacement et à la communication de ses données à caractère personnel après son décès. » Cette loi donne également la possibilité de désigner un **« tiers de confiance numérique certifié par la CNIL ».**

Ces articles reconnaissent donc en amont du RGPD l'*usus* et l'*abusus* du droit de propriété que le cybercitoyen a de ses données personnelles.

Premier principe : la libre circulation des données au sein de l'Union. La circulation de la donnée n'est ni limitée, ni interdite. Conforme au principe fondateur lors de la création du Marché commun, la donnée devient un bien par sa libre circulation (Article 1er). Ce règlement s'applique à tout établissement situé dans le périmètre de l'Union. Ainsi donc des filiales françaises de grands groupes américains sont soumises au RGPD (article 3). Il s'applique à la personne physique. Il est regrettable que ce règlement n'envisage pas l'objet connecté de la personne physique car l'objet connecté est relié à l'usage d'un citoyen et va générer des données personnelles. Elles doivent être la propriété de l'usager citoyen, *« le cybercitoyen*[52] *»,* le nouvel homme dans la cité numérique (article 4.1).

Le consentement doit être la cheville ouvrière de l'exploitation catégorielle de la donnée du citoyen numérique. Or dans son article 7, il est prévu que la charge du consentement appartient au responsable de traitement. À l'heure actuelle, **ce consentement est donné par défaut.** Les données personnelles sont siphonnées par le jeu de l'acceptation des conditions générales de vente que le citoyen numérique ne lit pas. Il faut donc un acte positif (déclaration écrite article 7.2). Cependant, toute velléité de peser sur l'exploitation de votre donnée est anéantie car si vous n'êtes pas d'accord, le traitement reste licite même si vous retirez postérieurement votre consentement (article 7.3). Il semble aussi possible d'attaquer lorsqu'il y aura un déséquilibre significatif (article 7.4). À la lumière de cette rédaction, il apparaît que nous pourrions utiliser le nouvel article sur les clauses abusives 1171 du Code civil *(dans un contrat d'adhésion, toute clause qui crée un déséquilibre significatif entre les droits et obligations des parties au contrat est réputée non écrite)* qui vise les contrats d'adhésion. Or les CGV peuvent être considérées comme des contrats d'adhésion.

L'article 17 est tout à fait intéressant car il met en place un droit à l'effacement aussi appelé « droit à l'oubli » et dont la Cour de justice de l'Union européenne avait entamé la reconnaissance. Le citoyen numérique peut demander la cessation de la diffusion de ses données lorsque ces données ne sont plus nécessaires à la finalité pour laquelle elles ont été collectées ou traitées, ou parce que le citoyen numérique retire son consentement car la finalité a changé (cf. article 6-1 a). Cette cause de retrait vise les finalités spécifiques. Le système du consente-

52 Ibid. 11

ment par exploitation catégorielle présenté ici rentre dans le schéma de l'article 17 1 a) combiné avec l'article 6-1 a). Il peut aussi tout simplement s'opposer à cet usage. Le responsable de traitement dans ce cas doit procéder sans délai à l'effacement. Il existe aussi des cas de limitation de l'utilisation des données.

L'article 20 prévoit la portabilité des données : la récupération et la transmission des données personnelles à un autre système (opérateur en pratique) et le responsable de traitement ne peut s'y opposer. Il reste à définir dans le concret quelles sont les modalités de la transmission et les normes techniques. La portabilité des données n'est pas en pratique très effective aujourd'hui. Des questions d'interopérabilité des données et des traitements voient le jour. *Quid* des traitements effectués par un fournisseur d'accès à Internet qui me sont utiles pour des déclarations administratives et fiscales et qui ne me seraient plus accessibles par mon changement de fournisseur ? Il y a un verrouillage technique par l'opérateur.

La portabilité des données est inscrite dans le livre II du Code de la consommation « Art. L. 224-42-1.-Le consommateur dispose en toutes circonstances d'un droit de récupération de l'ensemble de ses données ». Les données doivent être récupérables par le cyber-citoyen dans un standard ouvert qui doit être réutilisable et exploitable par un autre système de traitement automatisé (article L. 224-42-3 du Code de la consommation suite à l'article 48 de la loi pour une république numérique).

La portabilité consacre donc bien l'*abusus* du cyber-citoyen sur ses données personnelles. Il ne faut pas confondre portabilité et transfert. De fait, la portabilité n'implique pas l'effacement des données. Les données pourront toujours être conservées par l'opérateur en ligne avec sa finalité de traitement.

Dans le système du RGPD, il y a l'obligation de faire une étude d'impact sur la protection des données (articles 33 et 34). Il y a donc matière pour se mettre en conformité avec l'utilisation des données personnelles pour pouvoir transformer le modèle asymétrique en un modèle symétrique de gains réciproques dans le respect des données personnelles. De plus, il suffit de charger le futur Délégué à la Protection des Données (DPD) de contrôler que le consentement du cyber-citoyen

a bien été expressément donné dans ses missions de l'article 39 du RGPD. Le DPD devient naturellement le garant de l'exploitation catégorielle consentie des données du cyber-citoyen.

En outre, le RGPD opère un resserrement des obligations entre « responsable de traitement » (l'organisation pour le compte de laquelle les données sont traitées, et qui définit les finalités des traitements) et « sous-traitants » (les prestataires chargés d'opérer les traitements selon les instructions de leurs donneurs d'ordre). Alors que, jusqu'à présent, les responsabilités pouvaient assez largement se répartir par contrat entre ces deux acteurs-clés du traitement de données, le RGPD impose aux premiers de s'assurer que leurs prestataires offrent un niveau de sécurité conforme aux exigences vues plus haut, et aux seconds d'assister leurs donneurs d'ordre dans l'accomplissement de leurs obligations.

Par ailleurs, les articles 73 à 77 prévoient des actions contre les autorités de contrôle ou les responsables de traitement par la voie traditionnelle du recours en justice devant les juridictions nationales. Il n'y a pas eu à l'heure actuelle de recours par un citoyen sur l'utilisation non conforme aux dispositions sur la protection des données en France devant une juridiction nationale. Il est possible d'envisager des recours en actions de groupe ou en action conjointe par l'intermédiaire d'associations.

Nous avons vu le cas de la cyberattaque d'Equifax qui a engendré des troubles graves pour le citoyen américain dont la vie repose sur sa capacité de crédit. Des actions de groupe sont en cours. En France, la loi N° 2016-1547 du 18 novembre 2016 et son décret N° 888-2017 du 6 mai 2017 ouvrent la possibilité de faire des **actions de groupe** pour des problèmes d'Internet et de protection de vie privée.

Enfin, l'entrée en vigueur du RGPD s'accompagne d'une augmentation considérable des amendes encourues : en cas de perte, fuite ou compromission des données personnelles, ou encore de faillite à garantir les droits prévus par le règlement, les sanctions se monteront en effet à **20 millions d'euros ou à 4 % du chiffre d'affaires annuel du groupe** (le plus élevé des deux montants).

Pris ensemble, ces nouveaux éléments ne créent certes pas un véritable régime de patrimonialité des données personnelles au bénéfice des personnes physiques, mais incitent néanmoins les entreprises et institutions à mettre en place une véritable « gouvernance des données ».

Sans permettre au citoyen de recevoir une rétribution pour le traitement de ses données personnelles, **le RGPD replace ainsi les entreprises dans un rôle de gardiennes et non de propriétaires de ces données.** L'individu revient au centre du jeu. **Mais si la valeur de ses informations est reconnue, celle de ses données personnelles lui échappe encore.**

Il reste néanmoins la question **des données agrégées.** Le citoyen devrait recevoir une rétribution sur la masse de données brutes de la part de ceux qui vont produire la donnée agrégée. Des entreprises en France et à l'étranger commencent à créer des systèmes ad hoc pour permettre ce flux sur des transactions de données. Plusieurs expérimentations sont menées en Europe sur la monétisation des données personnelles, comme *MiData* au Royaume-Uni ou *MesInfos* par la Fondation Internet Nouvelle Génération (Fing) en France.

Aux USA, les *databrokers* existent déjà. Les courtiers de données, tels ACXIOM[53] ou BLUEKAI génèrent des revenus sur les données personnelles qu'ils vendent aux entreprises. Acxiom aurait déjà recollecté 600 données par foyer sur 6 millions de foyers français. A priori Acxiom ne fait pas le commerce des données sensibles (données de santé par exemple). Toutefois, la société collecte les données des réseaux sociaux et peut dessiner votre profil de consommateur. La *Federal Trade Commission* (FTC) enquête sur la collecte effectuée par cette société.

[53] www.zdnet.fr/actualites/data-brokers-aux-etats-unis.

Pour ce qui est de rééquilibrer la chaîne des revenus issus des données, il faut avouer que le système américain des *databrokers* favorise les GAFA, puisque par exemple Facebook a signé des partenariats avec 4 des plus grands courtiers de données. Le système actuel manque cruellement de transparence sur les clients auxquels les données sont revendues et de réversion de revenus aux cyber-citoyens pompés de ses données. En outre, il existe peu de droits d'accès pour les cyber-citoyens américains.

©visualhunt

3. Vers une patrimonialité des données personnelles : les solutions juridiques

Après avoir établi que la donnée relève du régime commun des biens, nous venons aux propositions concrètes pour garantir un droit de propriété.

3.1 DISSOCIER LA DONNÉE DE L'INFORMATION

• La donnée est un droit de propriété de la chose et des droits

Considérer la donnée comme un bien comme un autre bien, revient à appliquer le droit des biens à la donnée. Pour être un bien en droit, il faut être :

- un objet de **désir,** un objet qui a une valeur qui découle de son utilité ou de sa rareté ;
- un objet **appropriable :** la personne établit un rapport d'exclusivité avec le bien, une situation privative vis à vis des tiers ;
- un objet qui **circule :** l'objet circule de façon licite entre les individus, l'objet est transmissible et aliénable.

À l'heure actuelle, la donnée est monétisée mais elle a fait l'objet d'une appropriation unilatérale, sans véritable circulation au profit du premier fournisseur de donnée : le citoyen.

Les personnes physiques et morales ont un patrimoine. Ici on rattachera le patrimoine des données des objets connectés et de l'humanoïde à la personne physique qui **utilise** l'objet connecté et l'humanoïde et ceci constituera le seul patrimoine de l'*homo numericus*[54].

[54] Expression de Mme le professeur Solange Ghernaouti.

Par conséquent, le citoyen a un patrimoine de données. Nous privilégierons ici la conception large de la propriété, chère au professeur Ginossar[55], qui permet d'englober dans le patrimoine personnel les droits et obligations d'une personne juridique. Cela comprend donc les choses animées ou inanimées, mobilières ou immobilières, corporelles ou **incorporelles,** actuelles ou **futures** relevant d'une personne physique ou morale.

Ainsi figurent dans le patrimoine personnel : des choses corporelles et incorporelles, et l'objet de la réalisation des droits, droits réels et personnels.

La donnée personnelle est donc un bien meuble incorporel, tel que l'ont bien décrit Fabrice Mattatia et Morgane Yaïche[56]. Cette conception permet de créer une certaine homogénéité à des choses et des droits qui peuvent avoir actuellement des régimes juridiques différents (les données composant la Méga Data). Faut-il rappeler les articles 544, 545 et 546 du Code civil :

> Article 544 : La propriété est le droit de jouir et de disposer des choses de la manière la plus absolue, pourvu qu'on n'en fasse pas un usage prohibé par les lois et règlements.

> Article 545 : Nul ne peut être contraint de céder sa propriété, si ce n'est pour cause d'utilité publique, et moyennant une juste et préalable indemnité.

> Article 546 : La propriété d'une chose soit mobilière, soit immobilière, donne droit sur tout ce qu'elle produit, et sur ce qui s'y unit accessoirement soit naturellement, soit artificiellement. Ce droit s'appelle « droit d'accession ».

Le citoyen est donc propriétaire de la donnée première et des données générées et agrégées en tant que fournisseur de la matière première, des choses de son patrimoine personnel. Il devient créateur/générateur de données, la matière première sur sa personne et sur ses activités.

55 GINOSSAR S., *Pour une meilleure définition du droit réel et du droit personnel*, RTD civ. 1960, p37

56 MATTATIA Fabrice et YAÏCHE Morgane, « Etre propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ? » *Revue Lamy de droit immatériel*, 2015/114, pp60-63.

Il doit donc pouvoir la maîtriser : la vendre, la louer, la céder, voire la gager. C'est une opportunité énorme qui s'ouvre au citoyen, ne passons pas à côté. Il serait envisageable de construire des patrimoines distincts et de loger certaines des données du citoyen dans une fiducie.

Les personnes morales auront aussi un patrimoine de données. À l'heure actuelle, 84 % d'entre-elles collectent les données de leurs clients mais seulement 30 % s'occupent de la gestion de ces données.

• Un nouveau paradigme : le citoyen au centre du business model de l'exploitation de la donnée

Le business model des GAFA qui fait du citoyen un bon fournisseur « endormi » de données à potentiel énorme peut être bousculé par une prise en compte française et européenne des droits de l'*homo numericus*[57] et la mise en place d'un nouveau modèle économique basé sur une réversion des revenus générés au citoyen, prix de son consentement sur une exploitation catégorielle ou selon la finalité poursuivie de ses données.

Nous passerons donc d'un modèle gratuit à un modèle rétribué qui sera non seulement facteur de croissance mais aussi générateur de sécurité.

Si on s'en tient à la définition de l'agrégat telle donnée dans le dictionnaire Larousse 2015, l'agrégat est un *« ensemble d'éléments constituant un tout mais n'ayant pas de forme définie, d'organisation, d'unité véritable ou de finalité ».*

La donnée appartient à celui qui la fournit (conception classique) et le *business model* doit se fonder sur le premier fournisseur de donnée : le citoyen, qui sera rétribué sur la plus-value produite par la donnée, qu'elle soit première, générée ou agrégée.

57 Ibid.

A / Les acteurs de la chaîne de l'exploitation de la donnée

Nous envisageons les différents acteurs de la chaîne de l'exploitation de la donnée avec des métiers possibles :

- **Le citoyen ou la personne morale doté(e) d'un patrimoine de données :** premier fournisseur de données : soit ses données personnelles, soit les données de son entité .

- **Le collecteur de données** (*data centers*, FAI).

- **L'agrégateur de données :** sera l'entité privée (commerciale ou associative) ou publique (pourquoi pas une API d'Etat qui collecterait les données catégorielles des EPIC) qui aura la capacité technique et financière de gérer et analyser ces données. Le citoyen doit avoir un revenu de cette collecte sur la base du volume de la donnée et de la pertinence de la donnée. L'agrégateur peut vendre ces données aux plateformes.

- **La plateforme :** il faut aussi réformer le statut des plateformes. Le Conseil d'Etat dans son rapport annuel[58] suggère de réserver un statut particulier aux plateformes qui *proposent ses services de classement ou de référencement de contenus, biens ou services mis en ligne par des tiers.*

- **Le détaillant**[59] (ou analyseur) de données : le détaillant ou courtier est celui qui va vendre un service lié à l'exploitation des données.

- **Le Délégué à la Protection de la donnée** (DPD ou en anglais DPO pour *Data Privacy Officer*) : le chargé de la donnée personnelle dans l'entreprise selon le nouveau RGPD.
Le DPD deviendra celui qui sera chargé de définir le contenu de la donnée exploitable et pourrait accompagner la vente de données catégorielles selon une stratégie définie avec l'entreprise, en ligne avec le responsable de traitement.

---

58 Conseil d'Etat, rapport annuel 2014.

59 Expression de Gérard Peliks.

B / Un business model d'exploitation des données encadré par le droit

La réversion des acteurs de la chaîne d'exploitation de la donnée peut se faire sur la base de divers mécanismes juridiques existants : le schéma du contrat de licence de marque, le droit de suite cher au droit d'auteur, ou encore une redevance par exploitation consentie (*pay per loyal use*, PPLU[60]) intégré dans le système de collecte de données.

• Premier modèle : la marque et le contrat de licence

Dans ce mécanisme, le citoyen dépose son patrimoine de données à titre de marque. C'est un signe distinctif, un nom, un pseudonyme, un chiffre (adresse IP, carte d'identité, carte vitale), qui est enfermé dans un monopole de propriété.

Il est facile d'utiliser la classe 45 de la classification de Nice qui comprend « les services de sécurité pour la protection des biens et des individus » ou aussi les « services de réseautage social en ligne ». On pourrait ajouter aussi une catégorie dans cette classe telle : « services numériques d'exploitation de données personnelles ou professionnelles ».

L'exploitation de cette marque ne peut se faire que par **un contrat de licence de marque** qu'il concède à un tiers (ex : son assureur) et le citoyen est payé par **redevance,** sur le volume de la donnée et l'usage.

L'avantage ici est que l'on crée **un monopole de propriété en amont** et que les exploitants sont obligés de passer par le citoyen. L'inconvénient est que l'accès à ce titre de propriété n'est pas gratuit et qu'il crée une discrimination entre les citoyens. De plus, la donnée qui est exploitée n'est pas la donnée statique. Or la marque fige une partie de ses données à un moment précis.

• Second modèle : le droit de suite du droit d'auteur

La donnée peut être intégrée dans **l'article L.111-1** du Code de propriété intellectuelle en considérant que **c'est une œuvre de l'esprit** avec des attributs intellectuel et moraux et patrimoniaux.

60 Terminologie de l'auteur.

L'article L 111-2 permet que l'œuvre soit réputée créée indépendamment de toute divulgation, du seul fait de la réalisation, même inachevée de la conception de l'auteur.

Ainsi donc on permettrait au citoyen internaute dont la donnée est absorbée par des mécanismes informatiques tels que les *cookies* ou autres, de considérer *de facto* que **sa donnée issue de lui par ses activités sur Internet,** est une œuvre de l'esprit, dont le contenu est exploitable sur d'autres supports.

Il reviendra à la société collectrice de données d'informer l'internaute sur les données collectées (ce qui se fait un peu actuellement souvent *system d'opt out* ; attention que le consentement explicite *system opt in* ne revienne pas à un blanc seing de l'internaute à une exploitation dont on ne connaît ni le contenu exact, ni le périmètre), le type d'exploitation de ses données (reste à faire) et verser un pourcentage ou une somme forfaitaire sur la valorisation de la base de données ou des résultats de la donnée exploitée à l'internaute (reste à faire).

L'article L 111-3 permet cette exploitation puisque la propriété incorporelle est distincte de la propriété de l'objet matériel. Cela permet donc l'exploitation de la donnée sur de nombreux supports (TV, portables, tablettes, montres, etc.)Il suffit de rajouter dans l'article L 122-1 du Code de propriété intellectuelle la mention : droit de collecte (de données) comme suit : *« Le droit d'exploitation appartenant à l'auteur comprend le droit de représentation et le droit de reproduction et* ***le droit de collecte numérique*** *»*

On pourrait aussi en conséquence modifier **l'article L. 122-3** sur le droit de reproduction en ajoutant un troisième alinéa : *« Pour les données personnelles, la reproduction consiste en la collecte implicite ou explicite des données de l'internaute et de ses objets connectés dans un but lucratif »*

Dès qu'une société exploite la donnée de l'internaute, elle devra verser un pourcentage ou une somme forfaitaire à l'internaute fabricant de la matière première.

La controverse[61] est tout à fait possible car il est en effet difficile de trouver le caractère original à des données de connexion ou aux données numériques telles le numéro de téléphone, la carte vitale, la carte bancaire. Le système n'est ici attrayant que pour l'organisation de l'exploitation de la donnée, avec par exemple une société de gestion collective des données personnelles.

• **Troisième modèle : le citoyen créateur de base de données et le système déclaratif - Déclaration de destination limitée et *pay per loyal use* encadré par le droit des contrats**

Le citoyen devient créateur de base de données car il consent à l'investissement humain et matériel dont il est le sujet-objet. **Il est propriétaire de la donnée première et de la donnée générée.**

Il faudrait introduire un amendement à la définition du producteur de bases de données pour intégrer le créateur (reconnu par la CJUE dans son arrêt de 2015) de base de données, le citoyen qui consent à ce que ses activités génèrent et de la donnée exploitable.

Deux conditions pour être reconnu comme producteur de bases de données : personne physique ou société dont le siège social/administration centrale/établissement principal est en Europe et avoir réalisé l'investissement financier, humain, matériel substantiel (article L. 341-1 et L. 341-2 du Code de propriété intellectuelle).L'investissement ici est humain. On passerait donc du producteur de bases de données au créateur (citoyen) de base de données.

Le système déclaratif fonctionne assez bien et a fait ses preuves. Le seul inconvénient est que la masse de données que nous allons avoir avec les objets connectés va devenir ingérable. Le mécanisme serait le suivant.

[61] MATTATIA Fabrice et YAÏCHE Morgane, « Etre propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ? », *Revue Lamy de droit immatériel*, 2015/114, pp60-63.

1. Le citoyen créateur de base de données déclare auprès de la CNIL sa volonté de faire exploiter ses données par un gestionnaire de plateforme de la donnée et d'en tirer un revenu soit via des objets connectés, soit via des plateformes ou des FAI, pour une catégorie de données et une finalité précise .

2. Le citoyen créateur de base de données consent par écrit à un gestionnaire de plateforme de la donnée, dans un contrat de licence à **l'exploitation catégorielle** de sa donnée selon une **finalité déterminée.**

On soumet cette exploitation au consentement du citoyen dont les données sont issues, pour une exploitation catégorielle et limitée dans le temps et quantitativement. On utilise le système déclaratif de la CNIL et on fait du citoyen le décideur de telle exploitation catégorielle de la base de données. Ceci doit être valable pour tout citoyen mais aussi ouvert à toute entreprise qui travaille sur la Méga Data[62]. On a un système de Déclaration Selon l'Usage (DSU ou DWYU pour *Declare what you use*[63]). Ce consentement requis doit être explicite. On peut imaginer que le citoyen aille sur une plateforme de la CNIL ou sur une API de gestion de la donnée ou sur une plateforme d'une entreprise gérant de la donnée, où il s'enregistre avec ses données personnelles. Cette plateforme serait un centre de gestion de l'exploitation de la donnée personnelle. Le citoyen recevrait une demande d'exploitation catégorielle de ses données sur cette plateforme, et il répondrait à l'alerte et cocherait activement pour accepter un usage loyal et monétisé de ses données.

3. Le gestionnaire de la plateforme revend l'usage de cette base de données à des plateformes ou à des FAI ou à des GAFA par un « smart contact » dans la chaîne de blocs. Dès lors par la chaîne de blocs, la finalité et la catégorie de données sont enfermées et chiffrées .

62 Etude INPI, PI et Économie numérique, 2014.

63 Terminologie de l'auteur.

4. Les plateformes, FAI et GAFA reversent un pourcentage au détaillant pour l'exploitation de la catégorie de données selon la finalité déclarée .

5. Le détaillant de la donnée reverse un revenu par des micro-paiements en fonction de la valeur d'usage de la catégorie de données .

6. On soumet cette exploitation à un paiement par le gestionnaire de la plateforme de données vers le citoyen fournisseur de matière première, paiement effectué sur la plateforme, tout en respectant les exceptions de l'article L 342-2 du Code de propriété intellectuelle. Les données sont stockées sur la plateforme qui en devient le courtier de données.

7. On introduit un paiement par usage loyal (PUL ou PPLU pour *pay per loyal use for what you declare you use*). Le contrôle de l'exploitation de la donnée revient à la CNIL. Tout usage illicite et déloyal, dont le citoyen serait averti par un système d'alertes intelligentes, entraînerait la suspension ou le retrait par le citoyen de l'exploitation sélective de ses données sur la plateforme ou l'API.

Il faudrait une correspondance égale entre la liste DWYU et celle PPLU, contrôle exercé par la CNIL, pour respecter les droits du citoyen créateur de données. La CNIL resterait l'autorité de sanctions dans le cas de dépassement de l'exploitation catégorielle non conforme à la finalité déclarée.

Ce dernier système est aussi facile à mettre en place car il s'assoit sur une autorité administrative existante, la CNIL et des mécanismes actuels. Il s'insère dans la chaîne de production de la donnée en respectant l'ensemble des acteurs.

Il s'inscrit dans le nouveau Règlement Général sur la Protection de la Donnée par le recours au consentement exprès et entérine l'*usus* et l'*abusus* de la donnée personnelle faits par les articles 17, 20 notamment et va au-delà car il prévoit des revenus (*fructus*) pour le cyber-citoyen.

On nous rétorquera que les données personnelles ne doivent pas être dans le patrimoine, elles doivent rester des biens extrapatrimoniaux, droits incessibles, pour éviter les abus d'exploitation de la donnée et voir aussi se multiplier les cas d'usurpation de la donnée personnelle[64].

Mais au regard des faits actuels, est-ce que la soi-disant non patrimonialisation des données empêche une quelconque exploitation abusive de ses données ? Voyons-nous une baisse du nombre d'usurpations d'identité ? Force est de constater que non. On voit au contraire que le cyber-citoyen est désapproprié de son droit d'*abusus*.

L'idée ici avancée est de créer un équilibre économique dans un jeu déficitaire pour le cyber-citoyen. Or réduire cette dissymétrie est bien un des enjeux du RGPD de 2016. Nous mettons dans le système proposé, la CNIL comme acteur majeur du contrôle qu'elle est, de l'exploitation loyale et licite des données.

64 MATTATIA Fabrice et YAÏCHE Morgane, « Etre propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ? », *Revue Lamy de droit immatériel*, 2015/114, p62.

3.2 LE PARTAGE DE LA CHAÎNE DE VALEUR D'EXPLOITATION DES DONNÉES

PAR NICOLAS BINCTIN

Le statut des données a une influence importante sur le partage de la chaîne de valeur. Les données publiques, librement exploitables, ne semblent guère permettre une inclusion du collecteur initial des données ou de l'émetteur des données dans la chaîne de valeur. Ainsi, sans exclure une approche patrimoniale des données, l'Etat a, en revanche, écarté de participer à la chaîne de valeur attachée à l'exploitation de ces données.

Il en va différemment de la donnée secrète ou de la donnée personnelle collectée par des solutions logiciels au gré des usages des terminaux. Dans les deux cas, la donnée n'est pas publique, elle est même au cœur de **la transaction :** soit elle est l'objet direct du contrat (c'est le cas d'un transfert de données secrètes), soit elle constitue l'une des contreparties du contrat (c'est le cas de l'utilisation des données personnelles par les moteurs de recherche ou des réseaux sociaux). On retrouve la fausse gratuité de ces services : ils sont basés sur un échange de valeur, contenu contre données personnelles, au lieu de rechercher une contrepartie monétaire. Dans ces circonstances, il est nécessaire de rechercher les conditions de participation de l'émetteur de la donnée à la chaîne de valeur de celle-ci.

La question est complexe et peut, sous certaines formes, laisser penser qu'il est possible d'opérer **une monétisation directe de la donnée.** Par exemple, à chaque donnée collectée, on affecte un micro-paiement au bénéfice de l'émetteur de la donnée et on rétablit ainsi un équilibre. Cette idée reviendrait à dire que la donnée à une valeur supérieure au service rendu par l'opérateur en ligne car ce dernier, non seulement, fournirait le service mais devrait aussi payer l'émetteur de données client du service. Si cette situation n'est pas impossible, elle semble complexe à identifier et à mettre en œuvre. On l'écarte en tant que telle de réflexions développées ci-après.

Au stade actuel de développement du troc en ligne, services contre données, il nous semble que le partage de la chaîne de valeur ne doit pas être appréhendé dans une approche individuelle mais dans une approche collective, plus en phase avec l'enjeu de la collecte massive de données. Dans les deux cas, la question de la patrimonialisation des données demeure un enjeu car cela constitue un support juridique de l'analyse.

Dans l'économie de la donnée, ce n'est pas une donnée qui a de la valeur mais une masse de données. Dans ces conditions, la valeur pertinente est collective et non individuelle et il faut qu'une partie de la valeur attachée à ces opérations soit collectivisée pour permettre le financement de l'action publique. Dès lors, en présence d'une opération de troc, services contre données, entre un professionnel, le moteur de recherche ou le réseau social, et un consommateur, l'internaute, il est nécessaire de rechercher les conditions dans lesquelles il serait possible de **soumettre l'opération à la TVA.** La TVA collectée viendrait nourrir les budgets des Etats de l'Union européenne et assurerait un partage de la chaîne de valeur en lien avec l'intérêt général.

A / Une transaction à valeur ajoutée

Le troc est une opération ancestrale qui trouve une nouvelle vie dans l'économie en ligne et la collecte de données. Le troc entre entreprises représenterait, selon l'*International reciprocal trade association* (IRTA), l'équivalent de 12 à 14 milliards de dollars par an. Il serait aujourd'hui nécessaire de revoir cette étude pour y intégrer la valeur des échanges dans le cadre de l'économie de la donnée en ligne. En présence de tels échanges, le droit fiscal prévoit que, si l'on est assujetti, l'échange est considéré comme une double vente dont chaque transmission est soumise à la TVA. Si l'échange se fait entre deux assujettis imposés, la TVA n'est due qu'en principe que sur la marge bénéficiaire.

Le troc est aussi connu pour caractériser l'économie collaborative et les projets de monnaies locales et solidaires, tel le SEL[65]. L'appréhension fiscale de ces initiatives est une source récurrente de tension, notamment car elles n'ont pas vocation à s'exclure du paiement des charges sociales et de la TVA.

[65] MAGNEN J.-Ph. et FOUREL Ch., Mission d'étude sur les monnaies locales complémentaires et les systèmes d'échange locaux, *D'AUTRES MONNAIES POUR UNE NOUVELLE PROSPÉRITÉ*, Rapport remis au Secrétaire d'Etat chargé du commerce, de l'artisanat, de la consommation et de l'économie sociale et solidaire le 8 avril 2015.

Dans le monde numérique, l'échange doit être pensé non pas dans les rapports entre professionnels, ou entre particuliers, mais dans les rapports entre professionnels et consommateurs-internautes. Il appartient alors au professionnel d'établir la valeur de la transaction et de déclarer la TVA due pour celle-ci. En effet, selon le CGI, la TVA est due pour toutes « livraisons de biens et les prestations de service relevant d'une activité économique effectuée à titre onéreux ». Ces prestations de service visent toutes les opérations autres que les livraisons de biens meubles corporels, c'est-à-dire de toutes les prestations non corporelles, dont l'utilisation d'un moteur de recherche, d'un réseau social, d'une plateforme de films en ligne, etc.

La TVA est due en présence d'une activité économique effectuée à titre onéreux. Il faut entendre cela comme couvrant les activités de marché, opérations économiques moyennant une contrepartie, un échange. Cette contrepartie est le plus souvent représentée par un élément pécuniaire mais elle peut aussi bien intervenir sous la forme d'un paiement en nature, telle la collecte de données. La TVA s'impose ainsi en cas de troc, pour dissuader les entreprises de préférer le troc au monnayage. L'intervention sur le marché doit être à titre onéreux, c'est-à-dire avec une juste contrepartie, ce qui ne se réduit pas au seul cas de la recherche d'un but lucratif qui suppose en plus la recherche d'un bénéfice. Ainsi, même les opérateurs du Net qui n'invoqueraient pas un but lucratif pourraient tout de même avoir une action sur le marché à titre onéreux, ce pourrait être le cas de *Firefox* par exemple.

À la lumière de ces éléments, il ne fait pas de doute que la collecte de données à l'occasion de l'utilisation des outils en ligne, en contrepartie de l'utilisation de ces outils, est une forme de troc, qui doit être qualifié, d'un point de vue fiscal, d'activité économique effectuée à titre onéreux.

Il faut donc que les opérateurs qui collectent de la donnée déclarent une valeur pour la transaction et supporte de la TVA sur celle-ci. Si pour l'internaute, l'échange est commutatif et synallagmatique car il lui permet d'accéder à un service rechercher, à condition que l'internaute soit conscient que l'opération n'est pas gratuite mais simplement fournie grâce à une contrepartie en nature, ses données, il est impératif que les Etats puissent collecter l'impôt sur la valeur ajoutée attachée à chacune de ces transactions. À l'heure où l'OCDE cherche à mettre en œuvre des politiques fiscales coordonnées afin de lutter contre l'érosion de la base d'imposition, il faut intégrer dans l'analyse de la chaîne de valeur ce point essentiel de l'économie en ligne.

Pour mettre en œuvre une telle solution fiscale, permettant une intégration dans l'économie locale des opérations réalisées en ligne et permettant aux Etats de mettre en lien leurs recettes fiscales avec l'évolution des pratiques des opérateurs économiques, il serait nécessaire de définir une assiette, un taux et l'opérateur redevable.

Tous ces points doivent trouver une réponse à l'échelle de l'Union européenne et non localement pour qu'une telle TVA puisse s'appliquer efficacement car, non seulement la TVA fait l'objet d'une harmonisation européenne mais, en plus, l'économie numérique est par nature transnationale. Il n'est pas nécessaire ici d'avancer sur le détail de ces éléments. On pense qu'il faudrait envisager une TVA au regard de la masse de données personnelles collectées par les opérateurs en ligne plus que à l'unité de donnée collectée. On pourrait prendre en considération des catégories de données, la présence de cookies, la présence d'outils d'analyse comportemental, etc. Par exemple, les données de géolocalisation auraient une valeur supérieure à celles relatives à l'âge de l'internaute.

Par ailleurs, la mise en place d'une TVA sur les transactions faussement gratuites, prenant la forme d'un troc services contre données, permettrait de neutraliser partiellement les effets des modèles économiques fondées sur cette fausse gratuité. Le coût de la TVA viendrait réduire un peu les marges étonnantes de certains opérateurs. On rétablirait une concurrence plus frontale et moins déloyale entre différentes solutions de prestations de service en ligne. Il serait ainsi possible d'ouvrir des espaces pour d'autres modèles économiques pour le commerce

et les services en ligne et un choix pour le consommateur entre le troc et l'achat des services dont il a besoin.

B / Une localisation dans le pays de collecte

L'essor de l'économie numérique soulève des défis qui se rapportent à la fiscalité internationale. Un rapport de l'OCDE analyse en détail ces défis[66]. Il observe que l'économie numérique s'impose comme l'économie au sens propre, de sorte qu'il serait difficile, voire impossible, de l'isoler du reste de l'économie à des fins fiscales. Il ajoute toutefois que certains modèles économiques et attributs essentiels de l'économie numérique peuvent exacerber les risques de réduction de la base d'imposition des opérateurs économiques, et décrit les effets attendus des mesures issues de l'ensemble des actions qui constituent le projet BEPS. Il présente en particulier les règles et mécanismes d'application qui ont été définis pour faciliter la collecte de la TVA à partir du pays où se trouve le consommateur lors de transactions transfrontalières entre entreprises et consommateurs, et qui permettraient d'établir des règles du jeu équitables entre fournisseurs nationaux et étrangers.

Le rapport étudie et analyse des solutions possibles aux défis fiscaux de plus large portée posés par l'économie numérique, et souligne la nécessité de suivre les évolutions de l'économie numérique au fil du temps. L'idée d'une TVA sur la collecte et le traitement massif de données, notamment de données personnelles, n'apparaît pas contraire aux solutions préconisées par ce rapport.

Les initiatives de l'UE en 2016[67] pour engager des réformes du cadre fiscal de l'UE visent à la fois la fiscalité directe et la fiscalité indirecte au sein de l'UE. Le paquet de mesures contient plusieurs initiatives législatives et non législatives visant à aider les Etats membres à protéger leur base d'imposition, à créer un environnement équitable et stable pour les entreprises et à préserver la compétitivité de l'Union à l'égard des pays tiers.

Pour le commerce en ligne, la Commission propose d'améliorer l'environnement de la TVA dans l'UE, afin de permettre aux consommateurs et entreprises, notamment les start-up et les PME, d'acheter

66 OCDE, *Relever les défis fiscaux posés par l'économie numérique, Action 1 - Rapport final 2015*, Éditions OCDE, Paris.

67 Justice fiscale : la Commission présente de nouvelles mesures contre l'évasion fiscale des entreprises, janvier 2016, http://ec.europa.eu/taxation_customs/business/company-tax/anti-tax-avoidance-package_fr.

et vendre plus aisément des biens et services en ligne. Libérer le potentiel du commerce électronique en Europe et créer un marché unique numérique font partie des grandes priorités de la Commission Juncker. La mise en place d'un portail paneuropéen pour les paiements de TVA en ligne (le « guichet unique ») devrait sensiblement diminuer les coûts liés au respect des règles attachées à la TVA au sein de l'UE, estimé à 9000 € par entreprise et par pays de déclaration, permettant ainsi aux entreprises dans toute l'UE d'économiser environ 2,3 milliards € par an.

Les nouvelles règles garantiront que la TVA est payée dans l'Etat membre du consommateur final, ce qui se traduira par une répartition plus équitable des recettes fiscales entre les pays de l'UE. Ces propositions devraient permettre aux Etats membres de récupérer la TVA non perçue sur les ventes en ligne chaque année, qui est actuellement estimée à 5 milliards €. Selon les estimations, les pertes de recettes devraient atteindre 7 milliards € d'ici à 2020 si aucune mesure n'est prise. La proposition d'une TVA sur la collecte et le traitement massif de données attachés à une offre de service en ligne s'inscrit pleinement dans ces perspectives.

La Commission a proposé d'instaurer de nouvelles règles permettant aux entreprises qui vendent des biens en ligne d'accomplir facilement toutes leurs obligations en matière de TVA dans l'UE en un seul lieu. Cela devrait conduire à simplifier les règles de TVA pour les start-up et les micro-entreprises qui effectuent des ventes en ligne. La TVA due sur des ventes transfrontières d'un montant inférieur à 10 000 € sera gérée au niveau national et les PME bénéficieront de procédures plus simples pour les ventes transfrontières d'une valeur maximale de 100 000 € afin de leur faciliter la vie. Elle souhaite avec ces mécanismes accroître la lutte contre la fraude à la TVA provenant de l'extérieur de l'UE, qui peut fausser le marché et entraîner une concurrence déloyale.

Ces initiatives pourraient être utilisées pour la collecte de la TVA liée à l'échange de valeur en ligne et pourraient être aussi enrichies d'un mécanisme de TVA dédié à cette catégorie d'opérations en ligne.

05

PARTIE 3

La technologie au secours de votre vie privée ?

PAR GÉRARD PELIKS - LUCAS LÉGER

Comme nous venons de le voir, un utilisateur peut vouloir tirer des avantages financiers de ses données numériques personnelles qui sont collectées, pour être corrélées et exploitées avec les technologies du *Big Data*. Des outils lui sont proposés gratuitement, comme les moteurs de recherche, mais sans que ce soit explicite, ses data recueillies sont revendues à diverses organisations telles que les régies publicitaires ou les sociétés qui veulent mieux cibler leurs clients. L'utilisateur devrait pouvoir refuser ce modèle économique, quitte à payer pour l'utilisation des outils jusqu'alors gratuits, mais en revanche être payé pour la fourniture de ses données personnelles.

Mais encore faut-il que le créateur de données numériques puisse prouver son authenticité et établir que ces données et métadonnées lui appartiennent quand sa bonne foi ne suffit pas. Il faut aussi que ces données soient facilement accessibles aux acheteurs potentiels et restent intègres. Nous analysons ici plusieurs méthodes possibles pour s'authentifier et pour mettre ses données à disposition, et nous proposons un modèle basé sur une chaîne de blocs qui gère des « contrats intelligents ».

Nous montrerons d'abord qu'**une adresse IP** ne peut établir de manière sûre l'authenticité d'un utilisateur, bien qu'elle soit reconnue par la loi comme étant une donnée à caractère personnel. Nous expliquerons comment **une signature électronique** authentifie un utilisateur et prouve l'intégrité d'un document. La confiance repose sur l'autorité qui a signé le certificat numérique détenu par cet utilisateur. Nous évoquerons ensuite comment **une chaîne de blocs** peut être utilisée pour recueillir les données d'un utilisateur qui veut les rentabiliser. La confiance est donc fondée sur les multiples duplications des données et le nombre de ceux qui valident les transactions. Nous évoquerons enfin comment **les contrats intelligents** de cette chaîne de blocs

peuvent établir les conditions de transfert de propriété entre ceux qui fournissent leurs données et ceux qui les exploitent, par exemple pour les corréler entre elles, et avec d'autres données extérieures, par les algorithmes du *Big Data.*

Le modèle que nous proposons ici suppose de grands changements dans les outils et les méthodes et ne peut se réaliser que dans un rapport de force par lequel les utilisateurs l'imposent aux sociétés qui développent les outils et aux sociétés qui exploitent leurs données.

1. Comment prouver son identité en ligne ?

1.1 LES LIMITES DE L'ADRESSE IP

Une adresse IP (*Internet Protocol*) est-elle une donnée à caractère personnel ? Oui, répond aujourd'hui le législateur. Mais si on analyse cette qualité sur le plan technique, on voit que dans la réalité, il est compliqué voire impossible d'établir une correspondance crédible entre l'adresse IP d'un terminal et l'utilisateur qui est derrière.

Chaque élément physique connecté à l'Internet possède une adresse IP. Cette adresse est une suite de quatre octets, séparés par un point, dans le cas de l'IPv4 (exemple : 192.23.34.1) et de 16 octets dans celui de l'IPv6. Pour rendre la manipulation de ces adresses plus simple à mémoriser, on associe, à chaque adresse IP, un nom d'ordinateur suivi d'un nom de domaine, par exemple : noeud.domaine.fr, ou encore pour la messagerie : nom@domaine.fr. Les serveurs de noms DNS (*Domain Name Servers*) se chargent d'établir la relation entre la représentation en adresses IP et la représentation en noms de domaine. Il est plus facile pour l'utilisateur d'utiliser des adresses avec noms de domaine, faciles à retenir que de se rappeler un nombre. Les machines sur l'Internet ne manipulent

que des adresses IP. Dans le texte qui suit, nous caractériserons un élément connecté par son adresse IP uniquement plutôt que par les noms de domaines car ces derniers n'ajoutent rien à la relation entre un utilisateur et le terminal d'accès à l'Internet qu'il utilise.

Une adresse IP permet-elle d'authentifier un utilisateur ? Pas vraiment. Elle ne caractérise qu'un terminal (PC, tablette, smartphone, etc.) avec lequel l'utilisateur se connecte à l'Internet. Une adresse IP n'authentifie pas l'utilisateur. On pourrait penser que si un smartphone appartient à tel utilisateur, l'adresse IP de ce smartphone caractérise cet utilisateur, mais cette affirmation est hasardeuse car le smartphone a pu être emprunté ou volé par une autre personne.

Une adresse IP seule permet-elle d'être certain de l'identité d'un utilisateur qui se trouve derrière le terminal d'accès à l'Internet ?

Dans l'IPv4, la pénurie d'adresses IP disponibles aujourd'hui fait qu'un particulier qui se connecte se voit attribuer par son prestataire de services Internet une adresse IP dynamique le temps de sa session (en utilisant le protocole DHCP). Cette adresse sera ensuite attribuée par ce prestataire à un autre de ses clients. Ayant connaissance du moment de la connexion d'un utilisateur, le prestataire peut savoir à tout moment, en faisant une recherche, à quel client il a attribué une adresse donnée. En effet, le prestataire est tenu de conserver, au moins un certain temps, les journaux de connexion de ses clients. **Mais ce niveau additionnel augmente la difficulté de lier telle personne à telle adresse IP.** Pour savoir qui est derrière une adresse IP identifiée, outre le moment précis de la connexion dont il faut avoir connaissance, il faut aussi savoir qui est le prestataire de service et si ce prestataire accepte de faire la liaison entre son client et l'adresse IP qu'il lui a attribuée pour la durée de sa session.

De plus, dans les serveurs ou PC fixes situés sur l'Intranet d'une entreprise et protégés de l'Internet par un firewall, un employé sort en général sur l'Internet avec l'adresse IP du contrôleur réseau extérieur (celui qui voit l'Internet) du firewall. Ici, c'est son entreprise qui attribue, en interne, à chaque serveur ou PC connectés sur son Intranet, une adresse IP fixe, non routable, dite du « RFC1918 » et qui commence par 10 (exemple : 10.20.3.41), par 172 ou par 192. L'entreprise sait à quel ordinateur elle a attribué telle adresse IP, mais là encore, hors de l'entreprise, sur l'Internet, il existe donc une couche qui augmente la difficulté d'attribuer telle transaction à tel utilisateur puisque toutes les transactions extérieures semblent venir de l'adresse IP extérieure du firewall qui protège l'organisation.

Nous voyons donc qu'une adresse IP ne peut caractériser une personne de manière sûre. Elle peut seulement caractériser le moyen d'accès à l'Internet d'une personne, mais sans l'avoir authentifiée. Au-dessus se trouvent des couches logicielles qui ne rendent pas évident l'établissement de la relation entre telle personne et telle adresse IP. **En conclusion, si une personne souhaite tirer parti de ses données, il faut qu'elle soit authentifiée par autre chose que par son adresse IP.**

1.2 PROUVER SON AUTHENTICITÉ PAR UNE SIGNATURE ÉLECTRONIQUE

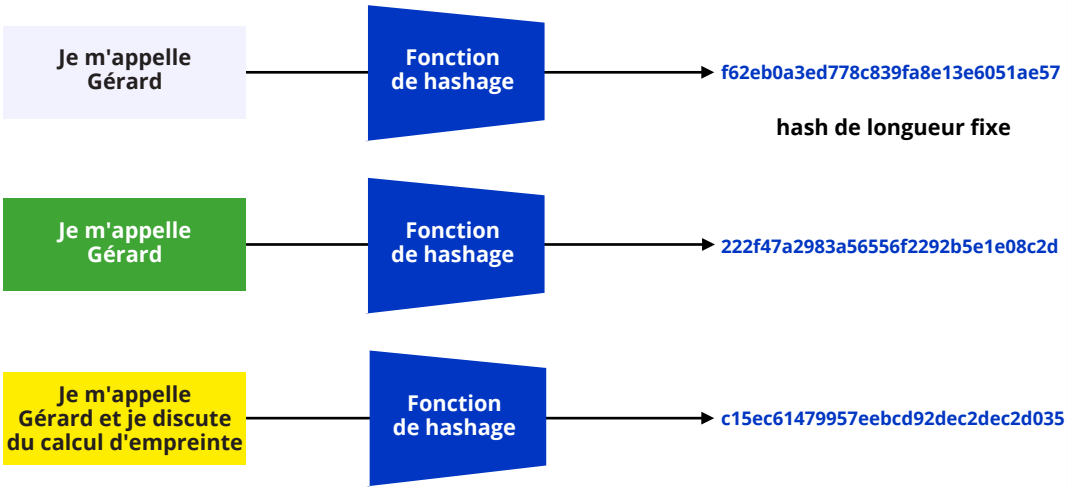
La signature électronique, basée sur des mécanismes cryptographiques, permet d'établir l'authentification forte d'une personne, quel que soit son moyen d'accès à une donnée numérique, et même s'il n'est pas connecté.

Nous allons donner ici des explications simples sur les mécanismes cryptographiques utilisés par la signature électronique. Nous expliquerons comment une signature électronique établit une relation entre une personne et un document numérique de manière au moins aussi valable, sous certaines conditions, que peut le faire une signature manuscrite entre une personne et un document papier.

Il convient de comprendre d'abord ce qu'est un calcul d'empreinte, puis ce qu'est **le chiffrement à clé publique,** dit encore chiffrement asymétrique. Nous ne rentrerons pas dans des détails techniques, notamment ceux des algorithmes utilisés, qui font la complexité et la beauté de la cryptographie. Nous donnerons juste un minimum d'explications nécessaires pour appréhender ce domaine vaste et complexe. Dans les prochains paragraphes, l'idée principale est de comprendre comment on signe électroniquement un document, et comment cette signature électronique établit à qui appartient le document.

- **Le calcul d'empreinte (ou hash) par les fonctions de hachage**

Une fonction de **hachage,** dite aussi fonction mathématique à sens unique, fait correspondre à un document de longueur variable une empreinte qui est une chaîne de caractères de longueur fixe.
Par exemple, si ce texte passe par la fonction de hachage SHA256, son empreinte sera une chaîne de caractères de 256 bits, quelle que soit la longueur du texte. Prenons pour illustrer un autre exemple.
Si l'ensemble des œuvres d'Emile Zola est traité avec la fonction SHA256, son empreinte sera aussi une chaîne de caractères de 256 bits, bien évidemment différente de l'empreinte du texte que vous êtes en train de lire, mais qui, passant par cette fonction de hachage, serait aussi une chaîne de caractères de 256 bits. Si on ajoute ne serait-ce qu'une virgule ou n'importe quel autre caractère à un des textes de Zola, la nouvelle empreinte calculée de l'ensemble de son œuvre sera différente de l'empreinte calculée de son œuvre initiale.



L'empreinte d'un document original, attachée à ce document, établit-elle **l'intégrité du document ?** Oui, mais tant qu'on ne le modifie pas en recalculant ensuite sa nouvelle empreinte, qu'on attache à ce document modifié. **Donc, pour établir l'intégrité d'un document, il faut lier à ce document son empreinte. On dit que le document est scellé. Mais pour autant, si l'intégrité du document commence à pouvoir être établie, ceci ne dit pas qui en est l'auteur.** Pour cela, nous allons expliquer maintenant les mécanismes du chiffrement à clé publique.

- **Le chiffrement à clé publique pour chiffrer l'empreinte**

Le chiffrement à clé publique fait intervenir deux clés mathématiquement liées : une clé privée et une clé publique. Si on chiffre avec une des deux clés, on ne peut déchiffrer qu'avec l'autre. La clé de chiffrement et celle de déchiffrement sont donc différentes, c'est pourquoi ce chiffrement est dit « asymétrique », mais les deux clés sont mathématiquement liées pour que chiffrant par l'une, on déchiffre par l'autre. La clé privée doit être gardée secrète par son propriétaire, il ne doit la diffuser à personne. En revanche, la clé publique mathématiquement liée à la clé privée, comme son nom l'indique, est publique et peut être distribuée à tout le monde. Bien entendu, possédant une clé publique, il n'est pas possible, ou en tout cas très coûteux en calculs, et donc très long en durée, de reconstituer la clé privée correspondante.

Mais qu'est-ce qui établit qu'une clé publique, qui peut être largement distribuée à tous ceux qui la demandent, est bien celle de telle personne qui détient la clé privée correspondante ? La réponse est qu'une clé publique n'est pas fournie seule et se trouve incluse dans **un certificat numérique** qui contient, outre la clé publique, plusieurs autres éléments qui caractérisent son propriétaire comme son nom, son prénom, éventuellement son organisation. Il contient également les dates de début et de fin de validité de ce certificat. Et surtout il est signé électroniquement par une autorité de confiance qui prouve que le contenu du certificat n'a pas été falsifié. Nous verrons comment cela est possible lorsque nous aurons fini d'expliquer ce qu'est la signature électronique. Retenez ici que le certificat a été signé électroniquement par une autorité à laquelle tous ceux qui vérifieront la signature électronique font confiance.

• **La signature électronique pour garantir à qui appartient une donnée numérique**

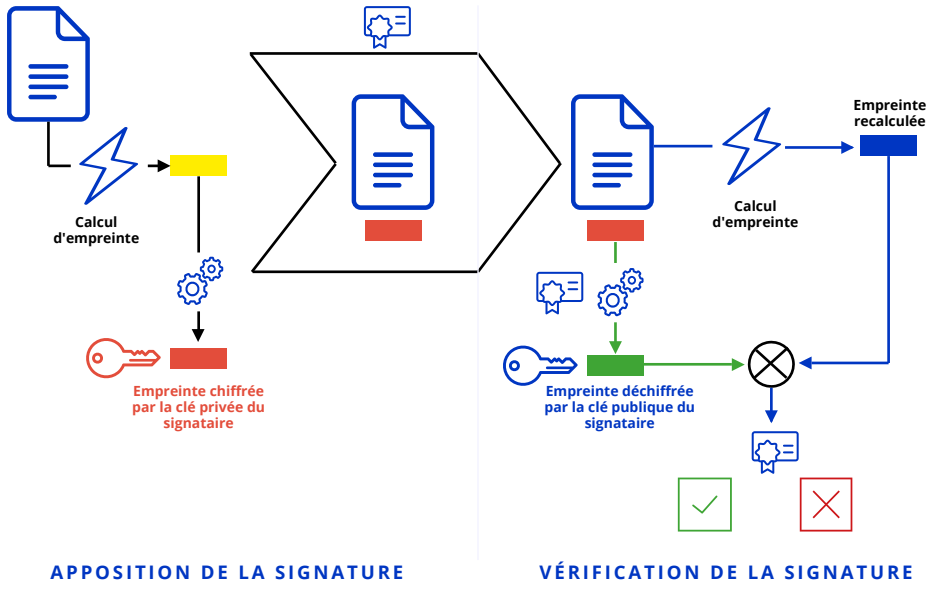
NB : Ce paragraphe est plus technique mais le lecteur peut se reporter directement au schéma ci-dessous pour une illustration de ce qui suit.

L'auteur d'un document calcule l'empreinte de son document par **une fonction de hachage.** Il chiffre cette empreinte avec sa clé privée, que lui seul possède, par un algorithme de chiffrement asymétrique tel que le RSA ou les courbes elliptiques. Il joint cette empreinte chiffrée à son document. Le document est alors scellé.

Celui qui veut établir l'authenticité du propriétaire et en même temps l'intégrité du document s'arrange pour obtenir **le certificat contenant la clé publique du propriétaire** en le lui demandant ou en allant le chercher dans un annuaire de certificats. Il regarde qui est l'autorité qui a signé le certificat. S'il fait confiance en cette autorité, et si le certificat se trouve bien dans les dates de validité, il sait que la clé publique qu'il extrait du certificat est bien celle mathématiquement liée à la clé privée du propriétaire et dont les coordonnées se trouvent dans le certificat. Il déchiffre l'empreinte du document à l'aide de la clé publique de celui qui a signé le document, par le même algorithme de chiffrement asymétrique. Il obtient une empreinte déchiffrée. Bien entendu ce n'est pas un utilisateur qui fait tous ces calculs, mais l'application de signature électronique qu'il utilise comme la messagerie ou le navigateur web dans le cas d'une connexion sécurisée https.

Le destinataire du document signé calcule alors **l'empreinte du document** avec la même fonction de hachage et obtient l'empreinte recalculée. Si l'empreinte recalculée est la même que l'empreinte déchiffrée, seul l'auteur, dont les coordonnées ont été trouvées dans le certificat d'où la clé publique a été extraite et attesté par l'autorité de confiance qui l'a signé, a pu avoir chiffré l'empreinte du document. Ceci parce qu'il est le seul à posséder la clé privée qui correspond mathématiquement à la clé publique trouvée dans le certificat de celui qui a signé le document. Le destinataire est aussi assuré de l'intégrité du document car il

peut être sûr que le document n'a pas été modifié depuis sa signature, cas de figure dans lequel l'empreinte recalculée n'aurait pas correspondu à l'empreinte déchiffrée.



Le fonctionnement de la signature électronique.

Ainsi sont établies l'authenticité de l'auteur et l'intégrité de son document. L'autorité de confiance a signé le certificat numérique de celui qui a signé le document par le même mécanisme que nous avons décrit. Quand un utilisateur veut signer un document, il doit donc posséder une clé privée et un certificat contenant la clé publique qui correspond à la clé privée, le certificat étant signé par une autorité de confiance.

Mais tous les certificats numériques n'ont pas la même valeur en fonction de la façon dont ils ont été obtenus. De même, les autorités de confiance n'ont pas le même poids, surtout dans un pays qui n'est pas celui où réside cette autorité de confiance.

Une signature électronique est donc un bon moyen pour un utilisateur de prouver qu'il est le propriétaire d'une donnée numérique et que cette donnée n'a pas subi de modification depuis qu'elle a été signée, ce qui prouve son intégrité. Le document a bien été signé

par son propriétaire quel que soit le terminal qu'il a utilisé. La confiance repose sur l'autorité qui a signé le certificat numérique, car elle possède aussi une clé privée qu'elle garde secrète, et un certificat contenant sa clé publique qu'elle distribue largement, et qui, pour un navigateur web, est souvent inclue dans le magasin de certificat que le navigateur possède.

Nul ne peut donc modifier un certificat signé par une autorité de confiance sans qu'on s'en aperçoive, car le résultat du calcul de l'empreinte de ce certificat ne correspondrait pas à l'empreinte déchiffrée du certificat avec la clé publique de l'autorité de confiance.

2. La chaîne de blocs pour garantir l'authenticité des données

Une chaîne de blocs est une « base de données distribuée, décentralisée de façon autonome »[1]. La plus connue des chaînes de blocs est celle de Bitcoin [2], un protocole informatique qui permet notamment l'émergence d'applications distribuées se passant de tiers de confiance, telles que la cryptomonnaie.

Loin d'être un feu de paille, certains y voient une technologie de rupture tout aussi révolutionnaire que l'avènement de l'Internet puis celui du Web. Dans le cas qui nous concerne, nous allons voir que **les chaînes de blocs peuvent apporter des solutions originales** au modèle économique que nous évoquons ici.

Où placer, dans le cyberespace, les données que l'on veut commercialiser ? Le plus simple est de les mettre dans un lieu facilement accessible, que tous peuvent consulter, mais que personne ne peut modifier. Une sécurité supplémentaire sur la disponibilité de ces données peut être garantie si ces données sont automatiquement dupliquées sur de nombreux serveurs largement répartis dans le cyberespace. La technologie d'une chaîne de blocs fournit cette possibilité, mais les tailles limitées de blocs des chaînes de blocs actuelles ne permettent pas de stocker de larges volumes de données. Aussi, nous conseillons de stocker uniquement, dans les blocs, les empreintes des données transférées, qui, elles, seront conservées par leur propriétaire sur un portefeuille spécialement dédié. En utilisant une application comme Zcash, ces transferts peuvent même être anonymisés.

Une chaîne de blocs peut être vue comme un grand registre ouvert à tous, dupliqué automatiquement et constamment sur beaucoup de serveurs, et dont les données numériques se situent dans des blocs infalsifiables. En effet, ces blocs sont chaînés ; chaque bloc est horodaté et dépend des précédents par des mécanismes cryptographiques. Une fois inclus dans la chaîne de blocs, un bloc ne peut être

[1] Source : https://www.youtube.com/watch?v=wKu_nZcgy3w.
[2] Source : https://bitcoin.org/bitcoin.pdf.

modifié car d'autres blocs s'ajoutent à la suite et si on modifie l'un d'eux, il faudrait aussi modifier tous les blocs qui précèdent. Des vérifications de **l'intégrité des blocs de la chaîne de blocs** s'opèrent, à chaque inclusion de nouveaux blocs, par de nombreuses personnes appelées les mineurs. Techniquement, toutes les duplications devraient être modifiées, car elles sont très sécurisées contre ce genre d'attaque. **Le mécanisme des chaînes de blocs peut être considéré comme raisonnablement sûr.**

Les données numériques des blocs sont signées électriquement par leur propriétaire. Elles pourraient être chiffrées mais sont généralement en clair car chacun doit pouvoir les consulter. Ajoutons que le propriétaire peut s'identifier par un pseudonyme ; s'il veut rester anonyme, seule sa clé privée utilisée pour inclure ses données dans la chaîne de blocs lui permettra de s'authentifier et de prouver que les données entrées sont bien les siennes.

En conséquence, si les données numériques d'une personne qui souhaite les commercialiser se trouvent dans une chaîne de blocs, le propriétaire de ces données peut attester qu'elles lui appartiennent jusqu'à ce qu'une transaction atteste que le propriétaire les a vendues, auquel cas les données numériques vendues ne lui appartiennent plus mais appartiennent à celui qui les a acquises. La chaîne de blocs permet à tous de s'en assurer.

Avec les standards actuels des chaînes de blocs, la taille de chaque bloc étant limitée, on met juste dans les blocs l'empreinte des données et leur horodatage.

3. Commercialiser ses données grâce à la technologie ?

• Le contrat auto-exécutant pour établir les conditions de vente des données

Ethereum, une des chaînes de blocs qui gère des contrats intelligents dits *smart contracts* semble être dès aujourd'hui une solution pour gérer les identités et les données numériques que l'on accepte de commercialiser.

Dans les contrats entrés dans ce type de chaîne de blocs, infalsifiables et ineffaçables à partir du moment où les mineurs les ont validés, il est possible d'établir les conditions sous lesquelles les données peuvent être acquises et comment doivent être effectués les paiements. Dans le cas d'Ethereum, **les paiements se font en *ethers.*** L'*ether* est l'unité d'une des nombreuses cryptomonnaies que l'on trouve dans les chaînes de blocs. Les possesseurs d'*ethers* peuvent ensuite convertir cette cryptomonnaie en euros ou dans autre monnaie fiduciaire via une plateforme d'échange.

Dans la chaîne de blocs Ethereum, les contrats intelligents sont écrits dans un langage informatique dit *Solidity*, qu'il faut évidemment maîtriser. Il est aussi indispensable de prévoir tous les cas d'application des contrats car ici, *code is law*. Mais il faut dire qu'après l'effondrement de *TheDao* suite à une vulnérabilité dans le code du contrat intelligent utilisant la récursivité, le code *is law* est remis en question. D'autres chaînes de blocs publiques ou privées seront sans doute créées[3] sur le modèle d'Ethereum, avec des langages de description des contrats intelligents plus simples d'emploi, de meilleures performances et utilisant des procédés de vérification et de validation, par les mineurs, plus adaptés à notre modèle économique.

[3] On peut déjà citer *Cardano, EOS, Dfinity* ou *Tezos* comme quelques concurrents potentiels.

• Utilisation possible de ces outils pour gérer les données personnelles

L'utilisation de ce qui précède implique de grands changements par rapport à ce qui se fait aujourd'hui. **Avec ce modèle économique, les données utilisées par les GAFA ou par des régies de publicité ne seront pas directement prélevées à la source, comme ce qui se fait actuellement, avec ou sans le consentement de leur propriétaire. Elles seront accessibles sur des serveurs dédiés où le propriétaire les mettra à disposition moyennant certaines conditions de paiement qu'il définira dans un contrat intelligent.**

L'utilisateur ou les programmes qu'il utilise doivent approvisionner cette chaîne de blocs. Ceci n'est envisageable que si le processus est automatisé. Les outils qui nous sont familiers : navigateurs, outils bureautiques, etc. devront être modifiés pour permettre aux utilisateurs de placer leurs données directement dans une chaîne de blocs que chaque utilisateur pourra choisir. **Si, dans le futur, cette méthode devenait universelle, les mécanismes actuels des chaînes de blocs devront être améliorés pour offrir les performances requises par l'avalanche des transactions qui seront à prévoir.**

Il n'est bien sûr question ici que de faire subir ces traitements aux données personnelles qu'on souhaite commercialiser. Mieux vaut garder chez soi, ne pas divulguer et chiffrer les données que l'on souhaite garder confidentielles. **Et mieux vaudra alors utiliser des outils qui s'adressent à une chaîne de blocs pour fournir des données, plutôt que d'utiliser des outils qui fournissent directement aux concepteurs les données que vous souhaitez commercialiser.** Si vos données apparaissent dans l'écosystème de l'Internet, alors que la chaîne de blocs qui est censée les gérer ne fait aucune référence à cette transaction, par exemple la transmission à un tiers de vos données sans votre accord préalable, cela prouve que vos données numériques n'ont pas été prises à cet endroit, et ont donc été acquises illégalement. On pourra alors parler de vol ou de plagiat.

• De nouveaux services pour un nouveau métier

Il est certain qu'on ne peut demander à chaque utilisateur de mettre en œuvre ce modèle. On ne peut lui demander de trouver l'acheteur à qui proposer les données dont il souhaite tirer des avantages, et qui inté-

ressent les régies publicitaires, et toutes les organisations qui vont traiter ces données pour mieux connaître leurs clients. **Il manque, dans la chaîne de la mise à disposition des données, un intermédiaire entre ceux qui produisent les données et ceux qui peuvent en tirer parti en les corrélant avec d'autres données et en les faisant traiter par des algorithmes pour obtenir des résultats qu'eux-mêmes vendront. Cet intermédiaire est le « détaillant en données numériques ».**

De même que les hommes et femmes de spectacle passent par des chargés de communication, les utilisateurs pourront passer par des détaillants en données numériques qui leur proposeront plusieurs services. Celui de leur fournir des outils pour que les données qui sortent de leurs terminaux, parfois à leur insu, n'aillent plus directement vers les GAFA, par exemple, mais transitent par les détaillants en données numériques. Ces nouveaux prestataires récolteront les données de tous types dans un lac de données dit *data lake*, et en assureront la duplication et la sauvegarde si les données en valent la peine, qualité de Véracité et de Valeur dans le *Big Data*[4].

Il est évident qu'une donnée isolée a peu de valeur et que le producteur ne peut donc pas en tirer grand-chose, mais une donnée corrélée avec beaucoup d'autres peut produire une information de grande valeur. La corrélation des données qui exige des algorithmes et des logiciels de traitement sera faite par le détaillant en données numériques.

Autre service qui sera très prisé, l'aide à l'écriture, ou la conception clé en main des contrats intelligents, en langage *Solidity* dans le cas de la chaîne de blocs *Ethereum*, ou en d'autres langages sur d'autres chaînes de blocs gérant des contrats intelligents. Ces contrats fixeront les conditions de la commercialisation des données.

Un service complémentaire sera d'anonymiser ou de pseudonymiser les données, et un autre service du détaillant en données numériques pourra être encore de proposer à ses fournisseurs de données ce qu'il faut pour qu'ils les signent, en particulier la clé privée et le certificat contenant la clé publique dont l'utilisateur se servira pour signer ses données. Le détaillant en données numériques pourra être aussi une autorité de confiance qui signe les certificats qu'il émet.

Le détaillant sait également où trouver les clients qui accepteront de lui payer un pourcentage de la somme due aux producteurs de données.

[4] Cela fait référence aux fameux 5V du Big Data.

Le reste se fait par les termes du contrat intelligent qui va lier les obligations du producteur et de l'utilisateur des données. Si l'idée se concrétise, poussée par les utilisateurs qui souhaitent tirer profit de leurs données numériques aujourd'hui ponctionnées et utilisées souvent à leur insu sous prétexte que l'outil qu'ils utilisent (navigateur, antivirus, compteur *Linky*, etc.) est gratuit, c'est un vaste chantier qui s'annonce.

Quand un logiciel est gratuit, on a coutume de dire que le produit est l'utilisateur. Avec ce modèle, l'utilisateur reprend la main sur la valeur de ses données et peut alors accepter de payer les logiciels qu'il avait coutume d'utiliser gratuitement. C'est un autre modèle d'affaires, et peut-être un modèle d'avenir.

• Les conditions de réussite et quelques exemples concrets

Notre proposition s'inscrit d'une part dans un débat beaucoup plus large sur le recours à différentes technologies pour assurer une meilleure protection de l'intimité des personnes lorsqu'elles naviguent sur Internet. On peut distinguer deux approches, qui ne sont d'ailleurs pas exclusives[5]. La première solution consiste à utiliser des moyens techniques pour limiter la diffusion des données personnelles. La deuxième tend à mettre en place des droits liés directement à l'utilisation et/ou la diffusion de ces mêmes données. L'exposé rapide de notre solution s'appuie sur ces deux aspects.

D'autre part, on note que des solutions existent déjà pour monétiser le contenu d'un site en ligne. Le souci de protection de la vie privée ou la rétribution juste des producteurs de contenu sont au cœur de ces solutions. Avec son *Smart Media Token*, la start-up *Steem* propose de rétribuer et encourager la création de contenu web par l'intermédiaire de sa chaîne de blocs[6].

[5] Nous renvoyons directement le lecteur aux travaux de : LE METAYER D., « Whom to Trust? Using Technology to Enforce Privacy », Chapter 17, in Wright D., De Hert P. 2016. *Enforcing Privacy Regulatory, Legal and Technological Approaches. Law, Governance and Technology Series*, Volume 25.

[6] https://smt.steem.io/smt-whitepaper.pdf. Sur un principe un peu similaire, voir aussi le projet *Akasha* sur Ethereum : https://akasha.world/. Ce dernier a le souci de protéger la vie

D'ailleurs, il n'est pas nécessaire d'avoir recours à ce genre de technologie pour tracer et monétiser des données[7].

Dans notre cas, le *smart contract* ici sert de pont entre des données personnelles, qu'elles se trouvent dans une base de données ou qu'elles soient directement détenues par son propriétaire.

La chaîne de blocs n'est ici que le registre des transactions et ne sert pas au stockage des données. En effet, une transaction effectuée sur une chaîne de blocs apporte une preuve de publication[8], qui permet d'éviter la double dépense[9], ou dans notre cas, la duplication d'une donnée sans le consentement de son détenteur. Dans ce contexte, le transfert de responsabilité est total. La donnée vous appartient[10] et le droit garantit la propriété. Si *Bitcoin* vous permet d'être votre propre banque, notre solution est relativement comparable, dans la mesure où vous êtes désormais responsable de la conservation et de la protection de vos données. Vous décidez ou non de les partager. Mais, à la différence de *Bitcoin*, vous détenez un droit de propriété garanti par l'Etat.

privée des utilisateurs via une plateforme de publication de contenu (sur le modèle de *Medium*), mais sans stockage des données sur un serveur dédié. Voir également BAT (*Basic attention token*) : https://basicattentiontoken.org/.

7 Sans nécessairement monétiser les données, des chercheurs au MIT ont mis au point le protocole HTTPa (pour Accountable Hypertext Transfer Protocol en anglais), qui traque l'utilisation des données d'un serveur à un notre. Certaines restrictions peuvent être intégrées en amont. Voir O. Seneviratne et L. Kagal, HTTPa : Accountable HTTP, novembre 2010, https://www.iab.org/wp-content/IAB-uploads/2011/03/oshani_seneviratne.pdf. Sweeney et al. proposent de leur côté d'associer à chaque fichier transféré un système de '*datatags*'. Cette notion introduit dans le partage d'un fichier un 'tag' de condition d'accès en fonction du degré de sensibilité des données transférées. Voir : https://techscience.org/a/2015101601/. Une solution plus radicale serait de tendre vers des plateformes décentralisées, où toutes les données de l'utilisateur sont stockées sur les serveurs de son choix. Voir : https://www.w3.org/2008/09/msnws/papers/decentralization.pdf.

8 Ce terme est bien entendu à distinguer de celui de 'preuve de travail', qui, dans les protocoles Bitcoin ou Ethereum, permet d'atteindre un consensus sur la légitimité de l'ensemble des transactions passées.

9 Source : https://petertodd.org/2014/setting-the-record-proof-of-publication#what-is-timestamping.

10 On voit déjà émerger ce genre de solutions avec des applications distribuées qui fonctionnent sur la chaîne de blocs Ethereum, comme par exemple IPFS (https://github.com/ipfs/ipfs).

S’il est désormais évident que la chaîne de blocs apporte la preuve de transfert et de cession d’une donnée personnelle d’un individu à une tierce partie, notre solution « technico-juridique » comporte néanmoins, à l’heure actuelle, certaines limites qu’il faut mentionner ici.

La première est liée à la territorialité du droit, qui peut rentrer en conflit avec le fait que la donnée est géographiquement agnostique. La seconde concerne le problème de l’identité sur une chaîne de blocs qui pourrait entraîner certaines complications juridiques. Nous verrons ensuite que les places de marché décentralisées peuvent difficilement monétiser les données (cf. annexe 2).

4. Les interrogations socio-économiques que pose une solution technologique

- **La non territorialité des données**

Au-delà des limites techniques[11] de mise à l'échelle, la complexité des interactions humaines introduit des difficultés relatives aux sciences humaines qu'il faut noter, même si elles ne sont pas insurmontables à long terme.

Par ailleurs, la nature de la donnée, qui peut être aisément copiée, rend parfois complexe le maintien de droits de propriété exclusifs et non rivaux. On pourrait même aller plus loin dans cette discussion théorique : la localisation d'un serveur et d'une base de données a nécessairement un effet sur l'efficacité d'un droit de propriété. Prenons un exemple bien réel pour illustrer notre propos. La publication des journaux académiques n'est pas faite directement par des universités, mais par des éditeurs reconnus tels que *Springer* ou *Elsevier.* Les chercheurs soumettent leurs articles à ces mêmes éditeurs, qui se chargent de ce que l'on appelle « l'évaluation par les pairs ». Les articles sont anonymisés
et envoyés à d'autres chercheurs pour commentaires. Ce processus est organisé par les éditeurs qui, une fois ce travail terminé, publient l'article. Celui-ci est ensuite vendu par l'éditeur soit à l'unité, soit sous forme d'abonnement. Très coûteux pour les universités, l'accès à la recherche académique se limite donc à quelques journaux ou plateformes.

Contestant ce monopole des éditeurs, une neuroscientifique kazakh, Alexandra Elbakyan, fonde Sci-hub en 2011. Le site abriterait aujourd'hui plus de 64 millions d'articles[12]. Il s'agit donc de la plus grande bibliothèque virtuelle d'articles académiques. Le site contourne les systèmes de paiements des revues en utilisant des identifiants existants et télécharge directement dans un serveur dédié l'article ainsi débloqué.

---

11 Mise à l'échelle, coûts de transaction élevés pour des nanopaiements sur des protocoles encore en construction ou à perfectionner. La mise à l'échelle des chaînes de blocs publiques (https://blockgeeks.com/guides/blockchain-scalability/) est, à l'heure actuelle, l'obstacle le plus important pour la mise en œuvre de notre solution, dans la mesure où l'on sait d'avance qu'un grand nombre de *smart contracts* seront déclenchés pour la rémunération des données personnelles sous forme de micropaiements. Sur Ethereum, des travaux sont en cours : https://plasma.io/plasma.pdf.

12 Source : https://en.wikipedia.org/wiki/Sci-Hub.

Le plus souvent, le contenu est récupéré illégalement. Cependant, le site est tellement simple d'utilisation que la plupart des chercheurs l'utilise même s'ils ont des identifiants fournis par leur université[13] !

La fondatrice du site est poursuivie par de nombreux éditeurs pour violation du droit de propriété intellectuelle[14]. Si plusieurs cours américaines ont déjà statué sur l'illégalité du site et condamné sa fondatrice par contumace à des amendes considérables, Sci-hub continue d'être utilisé et de copier du contenu initialement protégé par la propriété intellectuelle. Elbakyan utilise plusieurs noms de domaines et le site est également accessible par Tor, un réseau informatique qui permet d'anonymiser l'origine des connexions. Enfin, les serveurs ne se trouvent pas sur le sol américain et la base de données illégale ne peut donc être saisie ou détruite.

Quel intérêt dans le cadre de nos data ? **Cet exemple montre deux limites à notre solution. D'une part, que des données peuvent être facilement copiées et stockées sans le consentement de la contrepartie. D'autre part, que le droit a des limites territoriales. Si une personne mal intentionnée décide de piller des données personnelles et se trouve géographiquement à un endroit où la règle de droit est différente, il sera très difficile de la faire respecter.**

L'une des questions est donc de savoir si les Etats qui ne souhaiteraient pas appliquer cette solution seraient à même de rentrer dans ce genre de rapport de force. À l'heure actuelle, il n'existe pas de solution technique permettant de restreindre la duplication des données.

• Identité numérique et identité physique

Une deuxième limite provient des chaînes de blocs elles-mêmes. **Le système de transaction est par construction entièrement pseudonyme,** si ce n'est anonyme. Prenons à nouveau un exemple d'une transaction courante. Si je cède un bien immobilier, le notaire sert comme tiers de confiance. Ce dernier disparaît, en supposant que la transaction se fait désormais par l'intermédiaire d'un contrat autoexécutant. Cela ne pose aucun problème du point de vue de l'échange, à partir du moment où les deux parties se sont accordées sur les termes

13 Source : http://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone.

14 Source : http://www.sciencemag.org/news/2017/11/court-demands-search-engines-and-internet-service-providers-block-sci-hub?utm_source=sciencemagazine&utm_medium=facebook-text&utm_campaign=scihubblock-16248.

de celui-ci. Cependant, le notaire tient également le rôle de garant que les co-contractants sont en pleine possession de leurs moyens au moment de la transaction. La même transaction via un contrat auto-exécutant sur une chaîne de blocs publique a donc deux limites. La première, c'est qu'il est difficile de vérifier que la clé privée n'a pas été subtilisée et que la transaction est bien légitime. Deuxièmement, même s'il n'y a pas eu de vol de cette clé privée, rien ne garantit que je ne suis pas sous tutelle par exemple, et que je suis donc en pleine possession de mes moyens au moment de l'échange.

Ces difficultés liées à l'identité sont réelles et ont un impact potentiel en termes juridiques. Cela renvoie directement à un problème de répudiation du contrat, où le co-contractant mécontent aura un plus grand intérêt à ne pas honorer ses engagements. À partir du moment où l'identité n'est pas formellement vérifiable, on peut toujours arguer que l'on s'est fait dérober ou que l'on a perdu sa clé privée. Par ailleurs, nous avons déjà montré dans ces pages la propension des individus à négliger leur sécurité sur Internet. L'intermédiaire sera probablement incité à faciliter la tâche de son client en lui proposant de conserver la clé privée et de protéger son accès par un mot de passe. Ces deux aspects ont pour effet d'accroître l'instabilité et l'incertitude liées au « contrat ».

Certains garde-fous technologiques existent mais restent à l'heure actuelle encore imparfaits, et doivent encore faire l'objet de recherche et développement. **On peut envisager de contourner en partie ce problème grâce à la multi-signature. Une transaction, pour être valide, doit-être signée par plusieurs parties, ce qui réduit les risques évoqués plus haut.**

Plus globalement, les individus devraient pouvoir gérer eux-mêmes leur identité numérique[15]. Ce n'est pas le cas aujourd'hui, dans la mesure où elle est devenue un produit utilisé par des plateformes pour maximiser leurs revenus publicitaires. Ainsi, les défenseurs de la souveraineté individuelle dite *self-sovereingty* entendent redonner le contrôle à l'individu de son identité numérique[16]. Existent, ou sont en cours de développement, des applications qui permettent à leurs utilisateurs de mieux protéger leurs données liées à leur identité[17]. Ces applications apportent également une partie de la réponse aux limites liées à l'usurpation d'identité citée plus haut. Ces techniques permettent d'identifier formellement chaque utilisateur.

15 Source : https://arxiv.org/pdf/1712.01767.pdf.

16 Vitalik Buterin décrie brièvement ces points dans une récente interview : http://unchainedpodcast.co/.

17 On pense ici à uPort : https://whitepaper.uport.me/uPort_whitepaper_DRAFT20170221.pdf,

• **Octroi de droits vs contrôle**

L'un des aspects développés dans notre analyse s'appuie notamment sur la notion de contrôle des données. Dans l'environnement actuel, cette thématique est centrale, car il s'agit bien de replacer le consommateur-citoyen au cœur de notre solution, en lui redonnant, du moins en partie, un certain contrôle sur l'utilisation ou non de ses données par un tiers. La notion même de contrôle, sous un double aspect technique et juridique peut parfois être ambiguë, suivant que l'on se place du côté de la technique ou du juridique[18]. Si nous avons vu les contours juridiques de la notion de contrôle des données personnelles, il nous faut dire quelques mots sur ses aspects techniques.

En tant que solution de production décentralisée de preuves, la chaîne de blocs pourrait être un registre de référence enregistrant non pas les transactions des données elles-mêmes, mais les octrois de droit d'utilisation de ces données personnelles. **On pourrait imaginer un marché des droits relatifs à l'utilisation des données. Une sorte de marché de la diffusion hertzienne appliquée à l'utilisation des données personnelles.** La durée et le périmètre d'utilisation de ces données pourraient prendre la forme de métadonnées intégrées dans les transactions. Les transactions enregistrées dans la chaîne de blocs pourraient constituer la preuve d'octroi d'usage des données et de ses modalités d'utilisation. Cela pourrait constituer une preuve légale de l'utilisation de données personnelles. Le citoyen pourrait notamment s'y référer en cas de litige. C'est cette référence, cet élément de preuve d'octroi de droit, qui est susceptible de prendre de la valeur et donc de prendre la forme d'un actif échangé, plutôt que les données elles-mêmes.

ou à CIVIC : https://www.civic.com/products/how-it-works.

[18] La difficulté de contrôle sur les données personnelles, en tout cas d'un point de vue technique, est particulièrement vraie à l'ère du *Big Data*, voir par exemple : https://script-ed.org/article/control-over-personal-data-true-remedy-or-fairy-tale/. C'est notamment pour cette raison qu'il faut associer à une solution technologique une garantie juridique, voir WERBACH K., *« Trust but verify: Why the Blockchain Needs the Law ? »*, Berkeley Technology Law Journal, 2018

L’enjeu pourrait donc résider dans la construction d’un marché de droits d’utilisation des données préalable à toute tentative de création de marché de la donnée.

06

CONCLUSION

Nous croyons en un Internet décentralisé, où l'identité individuelle a encore une signification. L'individu ne se résumant pas à un flux de données. C'est dans cet esprit que nous avons conçu nos propositions.

L'avènement de l'ère numérique au début des années 2000 portait avec elle cet espoir d'une certaine décentralisation des pouvoirs, et de la libre circulation de l'information. Nous avons assisté au contraire à une concentration des pouvoirs dans les mains de ceux qui ont un accès sans limite à l'information que nous produisons chaque jour par l'intermédiaire de nos données.

Nous croyons en un Internet décentralisé, où l'identité individuelle a encore une signification. L'individu ne se résumant pas à un flux de données. C'est dans cet esprit que nous avons conçu nos propositions qui s'articulent autour deux axes : l'une technique et l'autre juridique, afin de permettre à chacun de se réapproprier son identité numérique.

Faire avancer le débat public sur ces enjeux est l'ambition de ces quelques lignes, qui ne peuvent malheureusement pas couvrir tous les aspects de questions aussi vastes que complexes. Si l'analyse ici exposée est essentiellement juridique, une seconde publication portera sur l'aspect économique et la valorisation de la donnée personnelle.

La technicité de certains de nos propos peut néanmoins se traduire en quelques pratiques, en guise de conclusion.

9 septembre 2019. Gérard ne conduit que le dimanche pour aller rendre visite à sa fille et ses petits-enfants. Quelques dizaines de kilomètres les séparent et Gérard connaît bien la route. Il utilise néanmoins une application de trafic en temps réel pour lui éviter les encombrements du périphérique parisien. De son côté, Camille, parcourt les routes de France dans son camion pour le compte d'une société de transport routier. Elle aussi utilise la même application.

Dans les deux cas, l'accès à l'application « coûte » la même chose (zéro) à un chauffeur routier qui circule tous les jours, et à un retraité qui ne prend sa voiture que le dimanche. En échange, ils sont tous deux obligés de partager leur géolocalisation, afin de renseigner l'application sur l'Etat de la circulation : Camille contribue donc bien davantage que Gérard à la valeur de l'application.

Dans l'hypothèse d'une patrimonialité des données, Camille pourrait exiger d'être rémunérée, la valeur qu'elle produit pour l'application étant largement supérieure à celle qu'elle en retire. À l'inverse, Gérard pourrait choisir de payer pour accéder à l'application sans lui céder ses données, la dédommageant ainsi de son comportement de passager clandestin. Ainsi chacun, en fonction de ses habitudes de vie et de ses usages d'Internet, pourrait bâtir le modèle économique qui lui convient le mieux[1].

21 juin 2021. Karim se réveille à peine. Une lumière estivale pénètre dans la pièce à mesure que les stores de son appartement se lèvent automatiquement. La radio s'allume sur son *smartphone.* Karim a faim. Il a néanmoins ignoré l'alerte de son réfrigérateur sur la pénurie d'aliments nécessaires à son petit déjeuner, et se contente d'un simple café déjà préparé par sa machine reliée à l'alarme de son téléphone. Chez lui, tout est programmé à distance et interconnecté pour lui assurer le meilleur réveil possible. Tout est calculé en fonction de ses phases de sommeil et de son temps de préparation avant de se rendre au travail.

L'interconnexion constante entre différents appareils génère une quantité importante de données sur ses va-et-vient quotidiens, son Etat de santé, ses habitudes de consommation[2]. Ces appareils sont aussi source de risques informatiques. Karim en a conscience et souhaite protéger cette intimité, tout en limitant le piratage éventuel de cette domotique.

Dans ce cas, deux options s'offrent à lui. La première consiste à sélectionner chacun des appareils qui intègrent dès la conception un mécanisme de protection des données, notamment par chiffrement[3]. Néanmoins, le transfert de données entre les différents appareils n'est pas toujours garanti. Karim opte pour une solution distribuée. Chacun des appareils est connecté à une même API. La transmission de toutes les données se fait via une chaîne de blocs dédiée[4] et protège ainsi la domotique de Karim de potentielles attaques tout en garantissant la confidentialité de sa vie privée.

---

1 L'application *Steamr* va dans ce sens : https://www.streamr.com/#howItWorks.

2 Sur les risques de l'Internet des objets et la fin de la confidentialité, voir par exemple : https://dl.acm.org/citationcfm?id=3105843&dl=ACM&coll=DL&CFID=850514586&CFTOKEN=59441772.

3 C'est ce que les anglo-saxons appellent *privacy by design.*

4 On pense ici à *IOTA* : https://iota.org/IOTA_Whitepaper.pdf. Bien que cette chaîne de blocs soit critiquée (https://hackernoon.com/why-i-find-iota-deeply-alarming-934f1908194b), l'idée d'une chaîne de blocs dédiée à l'Internet des objets reste intéressante.

22 janvier 2022. Alice a 19 ans et veut ouvrir un compte sur un réseau social pour la première fois. Au lieu de signer électroniquement un « accord utilisateur » qu'elle ne lira pas, une page d'accueil s'ouvre. On peut y lire les données pour lesquelles la plateforme offre une rémunération à ses utilisateurs en échange de leur utilisation par l'entreprise.

Pour chaque type de donnée (âge, ethnicité, ville, analyse des photos, préférences politiques, etc.), deux prix s'affichent[5]. Le premier correspond au prix d'achat par la plateforme des données d'Alice, le second au prix de la protection de ses données personnelles. Ces prix sont calibrés en fonction de quelques questions sur ses habitudes de navigation auxquelles elle a préalablement répondu.

Majeure, elle a déjà pris les devants et a déjà déposé les données qu'elle consent partager sur le réseau social à son courtier en données, qui les conserve sur des serveurs protégés et installés dans le pays d'origine d'Alice. Elle n'aura qu'à renseigner le nom de son courtier et son numéro de contrat[6] sur la page dédiée de son nouveau compte sur le réseau social en question. Ce contrat se déclenche automatiquement à chaque fois que toutes les conditions préétablies sont remplies. Par exemple, si Alice a consenti à partager son âge et son sexe avec la plateforme, elle sera rémunérée annuellement via un *smart contract*, pour leur utilisation. La validation par Alice via son numéro de contrat constitue l'élément déclencheur du *smart contract.*

Chaque année, son courtier lui enverra un bilan de ce qu'elle doit à la plateforme pour protéger ses données. Dans le cas contraire, elle la plateforme lui versera le montant en euros correspondant. Alternativement, Alice pourrait choisir de gérer seule ses données, et de les protéger sur un support informatique dédié. Dans ce cas, c'est elle qui devra informer le réseau social des données qu'elle consent à rendre publiques ou non. Un contrat sera généré avec le réseau social directement.

---

5 Cela reflète les derniers travaux en économie comportementale, où le prix que l'on est prêt à payer pour protéger ses données est souvent différent du prix que l'on est prêt à recevoir pour les céder. Voir ACQUISTI A., BRANDIMARTE, L., LOEXENSTEIN G., Op. cit. Et https://www.cmu.edu/dietrich/sds/docs/loewenstein/WhatPrivacyWorth.pdf. Cette stratégie de prix pourrait éventuellement être intégrée par les plateformes pour refléter ce biais cognitif.

6 Lorsque les conditions d'utilisation évoluent, alors le *smart contract* devient obsolète. Cela a un double effet. D'abord, il incite le fournisseur d'un service à ne pas les changer trop régulièrement, faute de voir l'accès à son service se réduire dans une phase de transition. Ensuite, on voit ici une certaine limite au *smart contract*, assez rigide, qui n'évolue pas à mesure des mises à jour d'un site non payant.

Alice paie désormais son service avec son information personnelle.
Dans les deux cas, la validation de la transaction via une chaîne de blocs permet d'authentifier une cession de droit sur ces données, et d'être rémunéré en conséquence.

07

ANNEXES

Précisions techniques.

Annexe 1. Analyse et évaluation de la donnée

Que ce soit du point de vue du gouvernement ou des grandes entreprises de l'Internet, les données personnelles jouent un rôle de plus en plus important dans l'analyse des comportements. Le plus souvent, ce ne sont pas vos données individuelles en soi qui sont intéressantes et monétisées, mais l'agrégation de toutes celles-ci[1].

Un changement de *business model* implique en conséquence une modification de la chaîne de valeur. Nous laissons cette question en suspens dans ce rapport, qui se concentre avant tout sur l'aspect juridique d'une donnée et tente de répondre à l'interrogation suivante : comment rendre le pouvoir au consommateur sur ses données. Si l'approche économique a un réel intérêt pour amorcer un débat, cette question trop technique et trop large pour la présente publication et sera abordée dans un second temps. La question est déjà épineuse, dans la mesure où elle implique des questions de politiques publiques qui sont seulement traitées actuellement par le législateur.

L'objectif est donc de brosser à grands traits les enjeux économiques principaux et de donner au lecteur un bref aperçu de la littérature[2].

Sur le plan macroéconomique, allouer un droit de propriété intellectuel permet en effet de gommer en partie le paradoxe de l'information énoncé par Arrow en 1962. Supposons que je veuille vendre à un tiers mes données liées à mon identité. Lors de cet échange, l'acheteur estime un ordre de grandeur du prix à payer pour cette information, c'est ce que l'on appelle en microéconomie sa propension à payer. Cependant, pour faire cette évaluation, l'acheteur peut demander

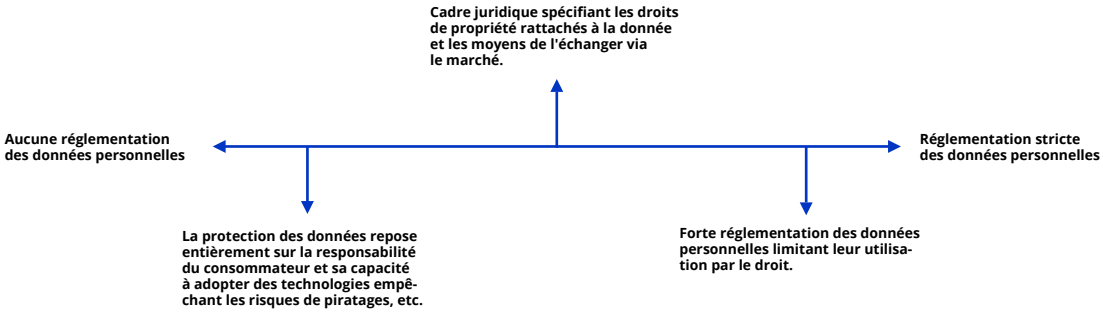
[1] On notera que les économies d'échelle ne sont pas toujours au rendez-vous. L'augmentation du nombre d'observations n'améliore pas nécessairement le retour sur investissement. C'est le cas par exemple dans la publicité. Lewis, R. and Rao J., *The unfavorable economics of measuring the returns to advertising*, Quarterly Journal of Economics, 130(4), 2015.

[2] Des revues de littérature sont disponibles, voir en particulier : ACQUISTI A., TAYLOR C., and WAGMAN L., Op. cit. ou : https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf.

au vendeur de lui révéler l'information afin d'en déterminer le prix. Mais à la différence d'un bien ou d'un service, les données ainsi divulguées n'ont plus aucune valeur. Dans ce contexte, la propriété intellectuelle apparaît comme un candidat naturel à la protection des données personnelles.

Si le paradoxe d'Arrow semble surmontable, on observe une autre difficulté liée à la commercialisation des données : elles sont non-rivales et non-excluables[3]. D'une part, cela signifie que l'on peut copier l'information sans que cela affecte leur utilisation par un autre consommateur. D'autre part, il est difficile d'exclure l'accès à utilisateur à cette information, même par l'intermédiaire d'un système de prix. En effet, un acheteur peut choisir de revendre les données qu'il vient d'acquérir, à partir du moment où le droit n'est pas spécifique sur ce point.

Une approche en termes de droits de propriété permet d'encadrer de façon plus ou moins stricte l'échange des données[4], comme l'illustre le graphique suivant :



En revanche, ce genre de modèle conduirait à réduire le surplus du consommateur[5]. Il y a donc un coût/bénéfice à l'excluabitlité qu'il faut pouvoir évaluer. Mais avant de pouvoir calculer ce coût, il est nécessaire d'établir un prix de marché à la donnée.

---

3 Source : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2858171.

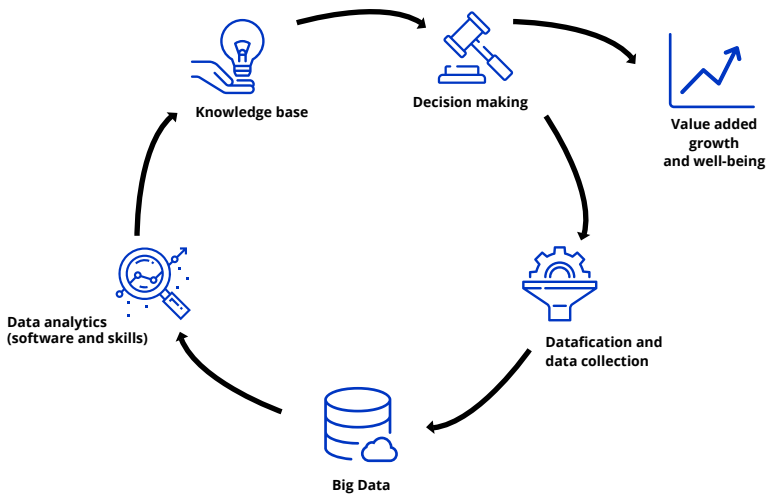
4 À ce titre, l'Europe et les Etats-Unis ont des approches différentes. Nous l'avons vu, le RGPD ouvre la porte à une patrimonialité des données. Tandis que les Etats-Unis défendent une position qui favorise l'appropriation des données personnelles par les databrokers. Ce marché est peu connu des consommateurs, à tel point que la *Federal Trade Commission* appelle à plus de transparence et un meilleur encadrement. Source : https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

5 ACQUISTI A., TAYLOR C., and WAGMAN L., *Op. cit*

Par ailleurs, la collecte le retraitement des données représentent un coût pour les organisations, qui doivent l'évaluer[6]. Cette évaluation est nécessaire, car elle est au cœur de nos économies de la connaissance, où l'on tire de la valeur à partir de l'information.

En premier lieu, les nouvelles technologies de l'information et des télécommunications ont des modèles économiques qui fonctionnent à partir de ces données[7]. Elles deviennent la matière première de la nouvelle économie numérique. La dématérialisation de nos relations s'est matérialisée par une nouvelle sorte de "surveillance" bienveillante. Si l'information a une valeur économique, on peut en déduire qu'elle implique de nouvelles relations de pouvoir.

En second lieu, l'analyse de la donnée s'inscrit elle-même dans une chaîne valeur, qui peut avoir un impact positif pour la croissance. Le graphique ci-dessous illustre ce cycle de la chaîne de valeur :



Source : OCDE

6 Nous renvoyons directement le lecteur aux travaux de l'OCDE sur ce sujet, qui permettent de mieux comprendre les difficultés que l'on rencontre dans cet exercice complexe. Source: https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/IE/REG(2011)2/FINAL&docLanguage=EN et http://predipubcn.sistemaip.net:8096/intranet-tmpl/prog/img/local_repository/koha_upload/DSTI-CDEP(2016)4-ENG.pdf.

7 VARIAN H., FARRELL J., and SHAPIRO C., *The economics of Information Technology : An Introduction*, Cambridge University Press, 2004.

En troisième lieu, les plateformes représentent un mécanisme de marché plutôt inédit. Depuis les travaux de Rochet et Tirole[8], on sait que ces marchés bi-faces doivent satisfaire des clients au profiles hétérogènes. D'une part, les utilisateurs, et de l'autre les publicitaires. Dans les premiers modèles, les mécanismes de prix reposaient sur un effet de réseau, et non la collecte et l'utilisation des données. Ces modèles aujourd'hui évoluent afin de mieux intégrer les effets sur le bien-être. Les plateformes permettraient effectivement de réduire les coûts de transaction et d'information[9].

Enfin, il existe une composante comportementales évidente qui influence les choix individuels et met en évidence la difficulté d'établir un prix de la donnée, qui dépend essentiellement du contexte[10].

8 Rochet J. C., and Tirole J., *Two-sided markets: a progress report*, Rand Journal of Economics, 37(3), 2006.

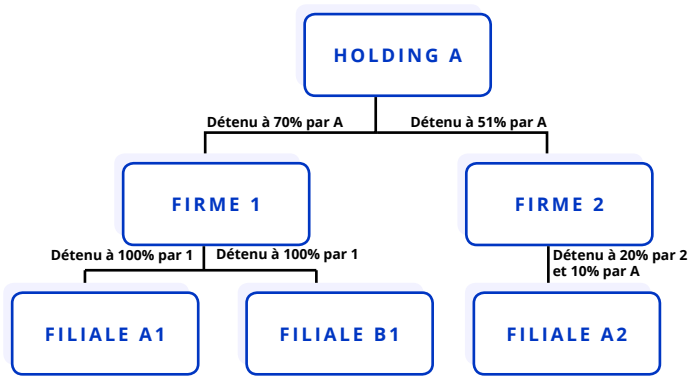
9 Source : https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf.

10 ACQUISTI A., BRANDIMARTE, L., LOEXENSTEIN G., *Op. cit.*

Annexe 2. Le cas des « Data Market places » décentralisées

Un marché de la donnée est **un marché de l'information.** Allouer **des droits de propriété** à cette dernière permettrait-il de corriger certaines **failles de marché ?** L'une des particularités de l'économie de l'information a été énoncée en 1962 par l'économiste Arrow[11]. Le partage de l'information introduit en effet ce paradoxe.

Considérons quelques instants cette structure très simplifiée de propriété ; les grands groupes ont en général plusieurs dizaines voire centaines de filiales qu'ils contrôlent plus ou moins directement. Une holding A détient des parts majoritaires dans plusieurs entreprises. Faisons l'hypothèse que toutes ces entreprises sont considérées comme des "établissements" par la nouvelle réglementation européenne et que ces derniers se trouvent tous au sein de l'UE. Cette hypothèse évacue certaines difficultés et dispositions spécifiques à la RGPD et liées aux transferts de données de l'UE vers des pays tiers. Dans ce cas, une telle structure en cascade permettrait à une holding de partager ses bases de données sur ces clients avec plusieurs des filiales, sans que ceux-ci n'aient nécessairement donné leur consentement, en l'occurrence avec les firmes 1 et 2 et leurs filiales a1, et a2. Cela est dû à la structure de propriété dans notre exemple. Dans ce contexte, la donnée échappe en partie à son détenteur. En résulte une difficulté de contrôle supplémentaire pour le régulateur. Elle n'est pas impossible, mais dans ce cas la réglementation devra prévoir la mise en place d'un système suffisamment transparent pour que l'utilisateur ait facilement connaissance où se trouvent ses données.



[11] Source : http://www.nber.org/chapters/c2144.pdf.

Si un contrat juridique peut spécifier en détails quelle entité détient quelles données, cela n'engendre pas moins **des coûts de surveillance** difficiles à dissocier dudit contrat. En effet, il serait bien trop coûteux de stocker chaque donnée liée à chaque transaction sur une chaîne de blocs. Cette nouvelle architecture n'est pas faite pour cela. Dans ce contexte, le suivi du respect de la loi devra se faire au cas par cas et là où se trouvent réellement les données échangées, au sein de chaque entreprise. L'information nécessaire au juge pour appliquer le droit devra donc être facilement vérifiable par ce tiers. Cependant, le juge est lui-même soumis à deux contraintes bien connues des économistes : l'aléa moral et la sélection adverse. La première suppose que le juge n'a pas fait l'effort nécessaire (peut-être par contrainte de temps ou de moyens) pour bien comprendre les tenants et aboutissants du cas. La seconde stipule que le juge n'a pas la connaissance nécessaire et/ou l'information pour traiter de ce cas en toute équité, ou qu'il a ses propres préférences en termes jurisprudence. Ces asymétries d'informations sont renforcées par le fait que les technologies utilisées ici sont nouvelles et ne sont pas encore maîtrisées par les professionnels du droit. Sur ce dernier point, certains juristes parlent d'"ossification du droit"[12]. Ces incertitudes inhérentes à la décision judiciaire font porter un risque supplémentaire pour les parties prenantes, à savoir que ni la lettre, ni l'esprit du contrat ne soient mis en œuvre[13].

Imaginons une plateforme d'échange numérique de données personnelles, sur laquelle on pourrait mettre en vente les données de notre choix et y trouver un acheteur intéressé. Les transactions des données seraient enregistrées sur le registre d'une chaîne de blocs publique en contrepartie d'une rémunération en crypto-monnaie. Seul hic, une fois que les données sont échangées, rien ne peut empêcher leur duplication. Un seul échange suffit, théoriquement, à détruire son propre marché. Une chaîne de blocs, si elle contient la preuve de la transaction, semble ici impuissante face au risque de « contrefaçon ».

12 MC GARITY, Thomas O. 1992. Some thoughts on deossifying the rulemaking process. Duke Law Journal 41: 1385, 1385–1462 ; Pierce, Richard J. Jr., 1995. Seven ways to deossify agency rulemaking. Administrative Law Review 47: 59, 60.

13 Source : https://www.jstor.org/stable/2999457?seq=1#page_scan_tab_contents.

08

SOURCES

Rérérences principales.

Ouvrages.

• BENYAYER L-D., CHIGNARD S., *Datanomics, les nouveaux business models des données*, FYP Editions, 2015.

• LANIER J., *Who owns the future ?*, Simon and Schuster, 2013.

• LUCAS André et LUCAS Henri-Jacques, *Traité de la propriété littéraire et artistique*, Litec, 2012.

• SIMLER C., *Droit d'auteur et droit commun des biens*, Litec, 2008.

• WRIGHT D., DE HERT P., « Enforcing Privacy Regulatory, Legal and Technological Approaches. Law », *Governance and Technology Series*, Volume 25, 2016.

• SFADJ Rubin et GRANGER Elodie, *Réussir votre mise en conformité GDPF : Guide pratique*, Broché, 2017

Rapports.

• ARROW K., *Economic Welfare and the Allocation of Resources for Invention*, The Rand Corporation, 1962.

• Organization for Economic Co-operation and Development (OECD), *Thirty years after the OECD Privacy Guidelines*, 2011.

• OECD, *Exploring the economics of personal data : a survey of methodologies for measuring monetary value*, 2013.

• SEN A., STIGLITZ J., FITOUSSI J-P, Rapport de la Commission sur la mesure des performances économiques et du progrès social, La Documentation française, sept. 2009.

Articles.

• ACQUISTI A., TAYLOR C., WAGMAN L., « The Economics of Privacy », *Journal of Economic Literature*, 54(2), 442–492, 2016.

• CANTERO Isabelle, « En attendant le règlement européen sur les données personnelles… 'l'essentiel fait vertu' », *L'Usine Digitale*, 19 février 2016.

• LECHENET Alexandre, « Données de santé : une base saine mais peut mieux faire », *Libération*, 5 mai 2016.

• HERAULT S., BELVAUX B., « Privacy paradox et adoption de technologies intrusives. Le cas de la géolocalisation mobile », *Décisions Marketing*, 74, 2014.

• MATTATIA F., YAÏCHE M., « Etre propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ? », *Revue Lamy de droit immatériel*, 114, 2015, p.62.

• NAKAMOTO S., Bitcoin : « A Peer-to-Peer Electronic Cash System », 2008 : https://bitcoin.org/bitcoin.pdf

• ROCHET J. C., TIROLE J., « Two-sided markets: a progress report », *Rand Journal of Economics*, 37(3), 2006.

• ROUX D., « La résistance du consommateur : proposition d'un cadre d'analyse », Recherche et Applications en Marketing, 22, 4, 2007, pp.59-80.

• Science, *The End of Privacy,* vol . 347 (6221), 2015.

• WRIGHT P., « Marketplace Metacognition and Social Intelligence », Journal of Consumer Research, 28, 4, 2002, pp. 677-83.

Sites internet.

• Une carte interactive de la CNIL pour savoir où vos données vont être exportées : https://www.cnil.fr/fr/la-protection-des-donnees-dans-le-monde

• Ethereum White Paper : https://github.com/ethereum/wiki/wiki/White-Paper

Conférence.

• #Dataday, Conférence débat sur la stratégie d'open Data pour le développement de l'économie de la donnée , 12 janvier 2016.

09

REMERCIEMENTS

Auteurs.

Nicolas Binctin.

Professeur agrégé des facultés de droit, Nicolas Binctin enseigne le droit des affaires et le droit de la propriété intellectuelle, dans ses différentes composantes, au sein des universités de Poitiers et Paris XII, et de l'École de Droit et de Gestion de l'Université Paris II.

Isabelle Landreau.

Avocat au Barreau de Paris et Docteur en droit, Isabelle Landreau exerce en droit de la propriété intellectuelle et droit des nouvelles technologies. Elle assiste ses clients dans la protection et la valorisation de leurs créations immatérielles et exerce également en tant que avocat-médiateur en propriété intellectuelle.

Gérard Peliks.

Ingénieur en cybersécurité, Gérard Peliks travaille dans le domaine de la sécurité de l'information depuis plus de 20 ans. Président de CyberEdu, il est aussi directeur adjoint du MBA Management de la Sécurité des Données Numériques de l'Institut Léonard de Vinci.

Virginie Pez-Pérard.

Enseignant-chercheur, maître de conférences à l'Université Paris II Panthéon-Assas, Virginie PEZ est spécialiste du comportement du consommateur, de la psychologie de la consommation et des problématiques de privacy et d'intrusivité dans les pratiques commerciales.

10

GENERATION LIBRE

La raison d'être du think tank.

Tocqueville déplorait déjà, dans *L'Ancien Régime et la Révolution*, « l'effrayant spectacle » des philosophes français, coupés du reste de leurs semblables, ignorants de la vie de la Cité, aveugles au reste du monde. « Même attrait pour les théories générales, les systèmes complets de législation et l'exacte symétrie dans les lois ; même mépris des faits existants ; même confiance dans la théorie. »

A l'inverse, les politiques restent bien souvent détachés de toute réflexion philosophique, en se reposant trop exclusivement sur l'administration pour imaginer les projets de réformes.

> **« C'est donc à mieux marier théorie et pratique, principes philosophiques et action politique, que doivent travailler les think tanks »**

Sur le fondement d'une doctrine claire, ils rassemblent les compétences d'experts pour décliner des idées parfois inhabituelles en politiques publiques précises et chiffrées. S'agissant du revenu universel par exemple, GenerationLibre s'est emparé d'un concept puissant mais très abstrait pour élaborer une proposition économiquement viable sous la forme d'un impôt négatif.

Il est heureux que les think tanks jouent un rôle croissant sur la scène publique française. Au-delà des convictions de chacun, c'est la garantie d'un débat riche et informé sur les grands sujets de notre temps.

ACTIONS

Notre combat quotidien.

Nos objectifs.

1. **Vivre et laisser vivre,** pour permettre à chacun de définir ses propres valeurs dans une société ouverte.

2. **Briser les rentes,** parce que la libre concurence des échanges comme des idées est le meilleur moyen de contester l'ordre établi.

3. **Penser le progrès,** pour que les innovations technologiques demeurent au service de l'individu.

Nos dernières publications.

- « Redéfinir le contrat de travail : de la subordination à la coopération », janvier 2017 ;
- « LIBER, une proposition réaliste, tome II », janvier 2017 ;
- « Le sexe et l'Etat : de l'indisponibilité à la libre détermination », juin 2017 ;
- « Retrouver l'Europe, pour un Etat minimal européen », chapitre I, avril 2017 ;
- « Schumpeter et les robots, le cas de la France », novembre 2017.

NOUS SOUTENIR

Soutenir de nouvelles idées.

GenerationLibre est un think tank fondé en 2013 par le philosophe Gaspard Koenig. Son financement repose exclusivement sur la générosité de ses membres, seule garantie de sa liberté de ton et de son indépendance. Il refuse toute subvention publique et n'effectue aucune activité de conseil.

Nous écrire, nous rencontrer.

GenerationLibre
24, rue Saint-Lazare
75009 Paris
contact@generationlibre.eu

www.generationlibre.eu