cigref

Réseau de Grandes Entreprises

Economie des données personnelles

Les enjeux d'un business éthique

SYNTHÈSE

La donnée est au cœur des priorités stratégiques des entreprises. La donnée est partout, ses sources sont multiples, provenant à la fois des entreprises, des individus, des puissances publiques, voire des machines elles-mêmes. La donnée circule, se reproduit, se stocke, s'agrège, se corrèle ; elle devient la matière première de nombreux métiers et la raison d'être de nouveaux marchés, témoignant d'une tendance généralisée qui est la *data driven economy* (l'économie guidée par les données). Les mutations économiques, stratégiques, politiques et sociales n'en sont qu'à leur début, le déluge des données s'accroissant de manière exponentielle avec le développement récent de nouvelles technologies (objets connectés, *machine learning*, intelligence artificielle…).

La donnée personnelle, au cœur de la « *data driven economy* »

Nous sommes aujourd'hui plongés dans l'ère du Big data, de l'interconnexion des données et de leur valorisation. **Dans ce contexte Big Data, les données personnelles occupent une place de plus en plus importante**. En effet, leurs modes de production, leurs moyens de collecte et d'analyse ont explosé ces dernières années et offrent **des potentiels de valorisation gigantesques, mettant notamment l'expérience client au cœur des stratégies**. La relation client devient de plus en plus proactive, ce qui permet d'anticiper les besoins, d'adapter les stratégies, les opérations, le *retail*, etc.

De nombreuses entreprises ont élaboré les premiers modèles fondés sur l'analyse et/ou la revente des données personnelles pour développer leur service. L'économie des données personnelles repose en particulier sur l'évolution de deux modèles, très justement identifiés par S. Chignard et L.D. Benyayer dans « *Datanomics* : *Les nouveaux business models des données* » (FYP éditions, 2015) :

- Les modèles bifaces qui se sont construits sur une logique de troc implicite : données personnelles contre service gratuit.
- Les modèles serviciels pour lesquels les données personnelles sont un véritable carburant : grâce à la personnalisation, les services sont plus efficaces et pertinents pour l'utilisateur et bénéficient tant à l'entreprise qu'au client (dans la limite des règles de consentement).

… / …

"*Innovation and privacy, you can have it all!*"[1]

Le sujet des données personnelles fait l'objet de nombreux débats animés opposant souvent deux conceptions idéologiques et culturelles: celle libérale, d'influence anglo-saxonne et celle protectionniste, d'influence européenne. Autrement dit, la donnée personnelle est un véritable or noir pour l'économie numérique mais constitue également un risque majeur pour les libertés individuelles, ce qui impacte la manière d'innover. Cela alimente un cercle vicieux. La conciliation est-elle possible ? La protection des données personnelles brime-t-elle nécessairement l'innovation ? Selon certains spécialistes, comme Ann Cavoukian, l'instigatrice de la *privacy by design*, il est parfaitement possible d'innover en faisant du *Big Data* tout en protégeant la vie privée : cela n'est pas antinomique. En prenant en compte les enjeux de protection des données personnelles le plus en amont possible, dès la conception des technologies, concilier innovation et vie privée devient une démarche non seulement **plus responsable** mais aussi **plus rentable**. En étant conforme *by design,* on parie pour une **innovation durable**, et cela constitue *in fine* une réelle valeur ajoutée.

Les conditions de réalisation d'un *business* éthique

Ce livrable est le résultat de réflexions des membres du CIGREF issus de secteurs et de fonctions variés. La vision commune qui émerge est que **le sujet de l'économie des données personnelles ne doit pas opposer *business* et éthique** : ces deux dimensions doivent être pensées conjointement**.** Pour tenter d'élucider ces enjeux majeurs, nous avons fait intervenir des acteurs de divers horizons (professionnels d'entreprises, *start-ups*, *think tank*, conseillers, juristes…) et analysé des *business models* spécifiques aux données personnelles.

Afin d'apporter des solutions à la construction d'un *business* éthique, nous avons parcouru l'ensemble de la chaîne de valeur des données personnelles, partant des modes de gouvernance et de régulation de l'innovation en interne, de la production à la distribution des données personnelles, jusqu'aux innovations d'usage comme le *Vendor Relation Management*, attestant d'une relation durable avec le client. Nos entreprises peuvent en effet être les pionnières d'une nouvelle économie symétrique, dans laquelle les clients/consommateurs deviendront délibérément co-créateurs de valeur.

[1] Ann Cavoukian & Deloitte, Livre Blanc « *Using Privacy By Design To Achieve Big Data Innovation Without Compromising Privacy* », 2014

Le CIGREF, réseau de Grandes Entreprises, a été créé en 1970. Il regroupe plus de cent très grandes entreprises et organismes français et européens de tous les secteurs d'activité (banque, assurance, énergie, distribution, industrie, services...). Le CIGREF a pour mission de promouvoir la culture numérique comme source d'innovation et de performance.

TITRE DU RAPPORT : L'ÉCONOMIE DES DONNÉES PERSONNELLES : LES ENJEUX D'UN BUSINESS ÉTHIQUE

EQUIPE DU CIGREF

Jean-François PÉPIN – Délégué général
Sophie BOUTEILLER – Directrice de mission
Anne-Sophie BOISARD – Directrice de mission
Josette WATRINEL – Secrétaire de direction
Flora FISCHER – Chargée de programme de recherche
Frédéric LAU – Directeur de mission
Matthieu BOUTIN – Chargé de mission
Marie-Pierre LACROIX – Chef de projet
Josette LEMAN – Assistante de direction

REMERCIEMENTS

Nos remerciements vont à **Christophe BOUTONNET, DSI du CONSEIL D'ETAT**, qui a piloté cette réflexion ainsi qu'à toutes les personnes qui ont participé et contribué à ce groupe de travail CIGREF :

Cyril BARTOLO – LAGARDERE
Antoine BELIGNE – PSA PEUGEOT CITROËN
Christophe BINOT -TOTAL
Frédéric DAMEZ – ESSILOR
Christophe DIGUE - EDF
Bernard DUPRE - RATP
Dominique ENTRAYGUES – MICHELIN
Georges EPINETTE – LES MOUSQUETAIRES
Pierre FOURNIER – HARMONIE MUTUELLE
Marie-Noëlle GIBON – LA POSTE
Bruno LALANDE – RENAULT
Yann-Maël LARHER – TOTAL
Emilie MAURIN - AXA
Jean-Luc RAFFAELLI – LA POSTE
Yao SOGADZI – FONDATION DE FRANCE
Fabienne VILLARS – RENAULT

Le CIGREF remercie également sincèrement les personnalités extérieures suivantes pour leurs interventions et leurs contributions aux réflexions :

- Sanaa AIT DAOUD, Responsable R&D – DIGITAL & ETHICS
- Thomas BEAUGRAND, Avocat – STAUB & ASSOCIÉS
- Alessandro FIORENTINO – Ambassadeur du Privacy by Design, Consultant au Cabinet INFHOTEP
- Daniel KAPLAN, Délégué Général – FING
- Ludovic LEVY, Vice-président, en charge de la Data Monétisation – ORANGE
- Stéphane PETITCOLAS, Ingénieur Expert en charge du laboratoire de recherches et d'innovation – CNIL
- Hervé RANNOU, CIO Data – CITYZEN DATA
- Marc VALLEE, Directeur Associé, R&D/Produit – YSANCE

Ce document a été rédigé par Flora FISCHER, CIGREF avec la participation de Marie-Noëlle GIBON, LA POSTE, Jean-Luc RAFFAELLI, LA POSTE et Christophe BOUTONNET, CONSEIL D'ETAT.

POUR TOUT RENSEIGNEMENT CONCERNANT CE RAPPORT, VOUS POUVEZ CONTACTER LE CIGREF AUX COORDONNÉES CI-DESSOUS :

CIGREF, Réseau de Grandes entreprises
21, avenue de Messine 75008 Paris
Tél. : + 33.1.56.59.70.00
Courriel : contact@cigref.fr

http://www.histoire-cigref.org/
http://www.questionner-le-numerique.org
http://www.entreprises-et-cultures-numeriques.org
http://www.entreprise2020.fr/

Droit de propriété intellectuelle

Toutes les publications du CIGREF sont mises gratuitement à la disposition du plus grand nombre, mais restent protégées par les lois en vigueur sur la propriété intellectuelle. Est autorisée la copie du titre et d'extraits de 500 caractères, suivis chacun de la mention « Source : » assortie de l'url de la publication CIGREF. Toute autre reprise doit faire l'objet d'une autorisation préalable auprès du CIGREF cigref@cigref.fr .

SOMMAIRE

Introduction 1

1. Qu'est-ce qu'une donnée personnelle ? 2
   1.1. Définition et caractéristiques de la donnée personnelle 2
      1.1.1. Protéifome 2
      1.1.2. Relative 3
      1.1.3. Ouverte 4
   1.2. Les modes de production des données personnelles 4
   1.3. Les droits rattachés à la donnée personnelle 6
      1.3.1. Le droit au respect de la vie privée 6
      1.3.2. Le droit à l'oubli 6
      1.3.3. Le droit à l'auto-détermination informationnelle 7

2. La valorisation des données personnelles 8
   2.1. Les potentiels de valorisation des données personnelles 8
      2.1.1. Stratégiques 9
      2.1.2. Opérationnels 9
      2.1.3. De recherche et développement 9
      2.1.4. Marketing 10
      2.1.5. Le cas de l'administration 11
   2.2. Les modèles d'affaires autour des données personnelles 12
   2.3. Anonymisation *versus* valorisation ? 13
   2.4. Les innovations technologiques : quel avenir pour les données personnelles ? 15

3. Les enjeux de la règlementation sur la protection des données personnelles face au marché 18
   3.1. Innovation et conformité 18
   3.2. Convertir les contraintes en avantages concurrentiels 19
   3.3. Les bonnes pratiques pour l'accompagnement juridique du marché 20
   3.4. Le projet de règlement européen sur la protection des données personnelles 21
      3.4.1. Cadre général 21
      3.4.2. Discussions et positionnement des entreprises utilisatrices 23
   3.5. Le cas particulier des données de santé et des données de bien-être 23
      3.5.1. Données de bien-être ou données médicales, quelles frontières ? 24
      3.5.2. La m-Santé, au cœur de l'économie numérique 25
   3.6. La localisation des données personnelles 26

4. La gouvernance de la donnée personnelle 29
4.1. Les référentiels existants 30
4.1.1. Assurer la conformité : Label de gouvernance des données - CNIL 30
4.1.2. Evaluer les projets : Charte Ethique et *Big Data* – Alliance *Big Data* 30
4.2. L'urbanisation des données personnelles 31
4.3. La bonne gouvernance des données personnelles 32
4.3.1. Penser conjointement *security by design* et *privacy by design* 32
4.3.2. La *security by design* 34
4.3.3. La *privacy by design* 34
5. Innovations d'usages et bonnes pratiques pour construire une relation durable avec le client 42
5.1. Le pilotage des données par le client : la théorie 42
5.2. Le VRM (*Vendor Relationship Management*) ou *Self Data* : la pratique 43
Conclusion 46

INTRODUCTION

Toutes les minutes, l'humanité produit 350 000 Tweets, 15 millions de SMS ; 200 millions de mails ; 250 gigaoctets d'informations sont archivés sur Facebook et 1 740 000 gigaoctects d'informations sont publiés dans le monde. Tous les jours, Google traite plus de 24 peta-octets de données, soit 24 millions de milliards d'octets. En 2013, 1,01 milliards d'objets connectés peuplent la planète. En 2020, nous estimons qu'il y en aura près de 100 milliards.

Voici la conséquence logique du poids que le numérique a pris dans notre vie : les données se multiplient à un rythme effréné. Produites par nos ordinateurs, nos téléphones mobiles, nos outils de paiement, mais aussi par les multiples capteurs qui équipent désormais nos voitures, nos routes, nos maisons et nous-mêmes, ces milliards de milliards de données s'accumulent sur les ordinateurs de la planète. Cette démultiplication croissante des sources et des flux de données a généré un accroissement de la production de données structurées, non-structurées, semi-structurées, ainsi qu'une baisse des coûts de stockage et d'analyse.

Dans cette data-sphère, les données personnelles constituent la matière première essentielle au service du développement de l'économie numérique. Tandis que la stratégie numérique des entreprises oriente de plus en plus la donnée et sa gouvernance au cœur de ses priorités, il convient de s'intéresser au cas particulier des données personnelles qui alimentent à elles seules tout un pan de l'économie numérique. De nombreuses sociétés se créent entièrement sur le modèle de la valorisation des données personnelles. Mais lorsqu'il s'agit de créer de la valeur autour des données personnelles, les questions d'éthique et de conformité apparaissent. Comment dès lors préserver la confiance des clients, tout en garantissant la création de valeur ?

Ce rapport identifie les enjeux de l'économie des données personnelles et les conditions de réalisation d'un *business* éthique. Il analyse des *business models* spécifiques aux données personnelles et expose les solutions méthodologiques et techniques à mettre en œuvre pour être conforme tout en étant innovant. Les entreprises partagent en effet cette conviction selon laquelle **la pérennité du succès de l'entreprise passe par l'élaboration d'un système de confiance, véritable accélérateur de *business*.**

1. Qu'est-ce qu'une donnée personnelle ?

Nous ne pouvons mener une réflexion sur l'économie des données personnelles sans exposer la nature de plus en plus complexe des données personnelles, et la difficulté qu'il y a à saisir dans son ensemble l'immense spectre qu'elles recouvrent.

1.1. Définition et caractéristiques de la donnée personnelle

La définition de la loi Informatique et Libertés couvre un périmètre très large car comme le précise **son** article 2: « *Constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne* ».

La notion de donnée personnelle est vaste dans la mesure où elle s'étend à tout ce qui « indirectement » renvoie à une personne. Est-ce que dans ce contexte, toutes les **traces numériques sont des données personnelles ?** L'adresse IP a par exemple été définie par la CNIL comme étant une donnée personnelle, car elle renvoie indirectement à une personne physique ou à un foyer identifié. La donnée personnelle serait-elle finalement presque partout ? Cette complexification de la nature des données personnelles s'explique par le fait que la donnée personnelle est de plus en plus :

1.1.1. Protéifome

L'ère numérique marque le passage d'un monde où régnait un seul type de donnée vers un monde où règnent trois types de données : aux données structurées s'ajoutent les données semi-structurées et non-structurées (issues des réseaux sociaux, mobiles, web…). Les systèmes et processus de gestion doivent être agiles et adaptables pour prendre en compte ces caractéristiques de la donnée, en général, et d'autant plus lorsque des données personnelles se cachent derrière ces catégories.

Il est à noter que les données personnelles sont classées en deux catégories distinctes :

- **Les données personnelles identifiantes** : ce sont les données rattachées directement à l'identité d'une personne. Ce sont généralement le nom, l'adresse, l'*e-mail*, la situation familiale, ou encore toutes les données ou fichiers permettant d'identifier indirectement une personne, via par exemple un numéro d'identification, une adresse IP etc.
- **Les données comportementales** : ces données sont rattachées à l'ensemble des comportements d'un individu, collectées via le suivi de ses navigations, mettant en exergue ses comportements d'achats ce qui permet, par exemple, pour un *retailer* (détaillant) d'affiner le profil d'une personne, surtout si ces données comportementales

sont croisées avec des données personnelles identifiantes. En cernant mieux les centres d'intérêt d'un individu et ses préférences, il est possible d'optimiser son parcours d'achat et de lui proposer des offres personnalisées.

Ainsi grâce aux données comportementales, les analyses prédictives permettent d'anticiper les types de produits ou de services qui correspondraient le mieux à tel ou tel profil ou à telle ou telle zone géographique.

Pour expliquer concrètement la différence entre données personnelles et données comportementales, voici un scénario proposé par Médiapost Communication, groupe de médias relationnels intelligents, filiale du Groupe La Poste :

« Albert, père de deux enfants domicilié à Neuilly-sur-Seine, se renseigne sur Internet à propos de l'achat d'une nouvelle voiture. Grâce aux sites consultés, nous allons pouvoir lui proposer des essais. Qu'est-ce qui relève des données personnelles ? Il s'agit de son nom, du fait qu'il ait deux enfants rattachés à son foyer et qu'il habite Neuilly. En termes de données comportementales, nous savons qu'Albert est prêt à utiliser le web dans sa démarche d'achat, qu'il envisage l'achat d'une nouvelle voiture et qu'il dispose d'un revenu plutôt élevé en raison des marques qui l'intéressent. À partir de toutes ces données, il devient possible d'en déduire qu'Albert pourrait avoir une propension à acheter plutôt une grosse berline, car il habite dans une commune à fort pouvoir d'achat et qu'il a une famille de deux enfants. »[2].

1.1.2. Relative

La donnée personnelle est nécessairement liée au référentiel culturel des individus et à la réglementation en vigueur dans chaque pays. C'est un aspect essentiel à prendre en compte quand une société s'adresse à un marché international. Nous pouvons questionner trois perceptions relatives de la donnée personnelle qui impliquent trois approches différentes de la problématique :

- Du point de vue de **l'entreprise** : la donnée personnelle a une valeur stratégique et marchande. Elle permet de développer ou d'améliorer les modèles serviciels, d'enrichir l'expérience client, d'être analysée à des fins décisionnelles, d'opérationnalisation, de revente etc. Chaque entreprise fait donc un travail de définition en contexte, en rapport avec le *business* qu'elle souhaite développer.
- Du point de vue du **client** : ses données personnelles sont une émanation de sa personnalité et de sa vie privée, une « trace » de ses comportements. C'est évidemment un point sensible. La confiance qu'il accorde dans les services numériques est fragile si la transparence, les finalités de traitement et les conditions de consentement ne sont pas claires.

[2] Source : http://vrm.mediapostcommunication.net/

- Du point de vue **juridique** : le statut juridique de la donnée personnelle permet-il de faire face aux tensions qui existent entre la protection des personnes et les besoins du marché d'utiliser leurs données pour développer de nouveaux services ? Un Code de la Donnée est-il souhaitable pour prendre en compte les enjeux de la société du XXIème siècle, tout comme l'on a créé un Code de la Propriété Intellectuelle au XXième siècle ?

1.1.3. Ouverte

La CNIL, dans son cahier IP n°1 sur « Vie privée à l'horizon 2020 »[3], a recueilli les paroles d'experts sur le sujet des données personnelles. Selon eux :

- La donnée personnelle offre une vision fermée de la donnée. Or, aujourd'hui, ces données sont **ouvertes**, elles circulent et se partagent en permanence. **La notion même de donnée personnelle ne correspond plus à la réalité.** Parler de données personnelles n'a plus de sens, il faudrait plutôt parler de **données relationnelles, transactionnelles** puisque les données circulent en permanence et qu'**il est possible de déduire des informations personnelles à partir de données non personnelles**. Le véritable enjeu se trouve dans **l'interopérabilité** des données.
- La principale difficulté aujourd'hui est de pouvoir identifier **les traitements sensibles** et non pas uniquement de sécuriser les données personnelles en tant qu'entité propre.

1.2. Les modes de production des données personnelles

Le nouveau paysage de données qui nous entoure est avant tout caractérisé par une **grande diversité de la nature des données et de leurs sources** : « *données transactionnelles, données collaboratives, données issues des systèmes de gestion, registres officiels, médias sociaux, données issues de capteurs et objets connectés, traces numériques, etc.* »[4].

Les données personnelles proviennent d'une part des individus eux-mêmes sur internet : données déclaratives, questionnaires, inscriptions en ligne… Mais les données personnelles s'étendent bien au-delà de cette sphère déclarative. **Les hommes ne sont plus les seuls producteurs de données, car les machines en produisent également** via différents types de capteurs ou d'objets connectés et qui révèlent parfois plus d'informations personnelles grâce aux techniques de *data mining*, capables de corréler des données de manière informelle et d'en tirer de nouvelles informations. Par exemple, les voitures connectées dotées de systèmes de *machine learning* permettent d'analyser des styles de conduites particuliers (sportifs, familial…) : ces données de type comportemental sont utiles à la compréhension des usages et permettent via l'analyse prédictive d'anticiper les besoins et de mieux organiser le

[3] http://www.cnil.fr/fileadmin/documents/La_CNIL/publications/DEIP/CNIL-CAHIERS_IPn1.pdf

[4] S.Chignard, L.D.Benyayer, « Datanomics, Les nouveaux *business models* des données », FYP, p.26

merchandising[5]. On le voit bien, les modes de production des données personnelles se complexifient avec l'évolution technologique qui enrichit considérablement la nature même des données personnelles, qui ne sont plus seulement liées à l'identité officielle d'un individu mais qui s'étendent à ses habitudes (de navigation, d'achat…), à ses parcours d'achats en magasin (via la pratique du « *clienteling* [6]», qui consiste à récolter via la Wi-Fi les données des clients sur leur *Smartphone*) ou en ligne, etc.

D'autre part, **la façon dont les données sont produites n'est pas toujours, pour leurs auteurs, consciente ou volontaire**, et pourtant toutes les traces numériques issues de leurs requêtes sur les moteurs de recherches, tous leurs « clics », parcours de navigation ou l'activation de la géolocalisation génèrent autant de données, qui peuvent être ou non associées à une identité physique, via l'adresse IP notamment.

La part d'activité de l'individu est considérable dans l'économie numérique. Le ***digital labor*** *(« travail digital »)* par exemple **est en soi un mode de production**, une ressource dont les entreprises de l'internet jouissent naturellement en échange de l'accès gratuit de l'internaute à un service. Selon Antonio Casilli, sociologue, le *digital labor* est « *l'ensemble des activités des usagers des plates-formes sociales, des sites web et des applications mobiles. Cela concerne non seulement la publication de contenus générés par les utilisateurs (des photos, des vidéos, des textes), mais aussi toute forme de jeu, de navigation, de bavardage en ligne qui ne serait pas reconnaissable formellement en tant que travail, et qui pourtant produit de la valeur pour les entreprises* »[7].

Dans ce contexte, il faut souligner que **l'essor des données non structurées** (textes, images, requêtes sur les moteurs de recherches etc.) **est un phénomène récent** qui découle du perfectionnement des outils d'analyse à partir des réseaux sociaux et des navigateurs internet. **Ces données non structurées amènent à une compréhension plus précise des comportements des internautes** et se révèlent **très utiles pour le marketing** ou la **stratégie** d'une entreprise. « *Il [l'entreprise] lui est difficile d'accéder aux réclamations que traite le service clients de son principal concurrent ; à l'inverse, il est relativement aisé et tout à fait légal d'enregistrer puis d'analyser tous les tweets, les articles de blogs, les messages Facebook qui parlent des produits dudit concurrent !* » [8]. Ces données non structurées sont une véritable mine d'or pour les entreprises. Sans oublier un autre atout majeur : elles existent sous forme de flux et sont donc nécessairement toujours « à jour ».

[5] Le *merchandising* est l' « ensemble des techniques assurant la meilleure diffusion commerciale des produits, grâce à une adaptation de ceux-ci, prenant en compte les désirs des acheteurs et différents éléments de stratégie commerciale (emballage, répartition des produits dans les surfaces de vente) » (Larousse)

[6] Le *clienteling* désigne la stratégie de personnalisation de la relation avec le client.

[7] Entretien avec Antonio Casilli, propos recueillis par Xavier Bonnefond et Judith Chouraqui, http://jefklak.org/?p=1467

[8] S. Chignard, L.D. Benyayer, « Datanomics, Les nouveaux *business models* des données », FYP, p.29

1.3. Les droits rattachés à la donnée personnelle

1.3.1. Le droit au respect de la vie privée

Les données personnelles sont bien évidemment rattachées au principe de protection de la vie privée, qui se décline sous trois aspects :

- **Le droit au secret** : ne pas voir ses informations personnelles divulguées à son insu,
- **Le droit à l'anonymat** : rester protégé d'une attention ou observation par des tiers, non désirée,
- **Le droit à la solitude** : pouvoir choisir sa proximité physique par rapport aux autres.

Dans le monde numérique, le respect de ces droits est fragilisé.

Appliquée aux données personnelles se trouvant dans le monde numérique, la protection de la vie privée supposerait donc :

- de savoir et pouvoir contrôler quelles données sont stockées par quel(s) tiers, en vue de quel(s) traitement(s) ;
- d'être informé de la collecte de ces informations ;
- d'être en mesure d'en obtenir à la demande le retrait et l'effacement.

1.3.2. Le droit à l'oubli

Une décision sur le « droit à l'oubli » a été rendue par la Cour de justice de l'Union européenne (CJUE), rendue le 13 mai 2014. Pour la première fois, la Cour se prononce en faveur du droit à l'oubli numérique en demandant à Google d'adopter les mesures nécessaires pour retirer de son index des données à caractère personnel concernant un internaute et d'empêcher l'accès à celles-ci, via ledit moteur de recherche, dans l'avenir.

Cependant, ce « droit » soulève encore quelques questions en suspens :

- Techniquement le droit à l'oubli ne signifie pas la suppression définitive des données concernées.
- En théorie, beaucoup de solutions sont envisagées pour garantir le droit à l'oubli ; en pratique, nous observons une difficulté à régler et réglementer la problématique du droit à l'oubli numérique au sein d'un cyberespace où les lois sont multiples car spécifiques à chaque état.
- Octroyer systématiquement une durée maximale de détention de données personnelles, c'est indirectement accorder une forme d'oubli à l'individu concerné. Mais une question, qui reste ouverte, est de savoir comment gérer avec efficacité et efficience les durées de conservation des données.

Le droit à l'oubli numérique reste pour l'instant une spécificité européenne et n'est qu'une goutte d'eau dans cet océan numérique où les lois sont multiples.

1.3.3. Le droit à l'auto-détermination informationnelle

Ce droit fondamental a été consacré successivement par le droit international, le droit de l'Union européenne et la jurisprudence constitutionnelle. Le Conseil d'Etat l'affirme à son tour dans son étude annuelle 2014 portant sur "Le numérique et les droits fondamentaux" :

> *« Proposition n°1 : concevoir le droit à la protection des données personnelles comme un droit à « l'autodétermination informationnelle » (et non pas un droit de propriété), c'est-à-dire le droit de l'individu de décider de la communication et de l'utilisation de ses données à caractère personnel. »*

Ce statut particulier octroyé à la donnée personnelle, qui suppose une maîtrise par l'individu de ses propres données (et non une propriété), implique de revoir tous les *process* qui visent à informer et recueillir le consentement des utilisateurs avant l'exploitation des données et de leur fournir les moyens concrets pour décider de ce qu'il adviendra de ses données : droit d'opposition, choix des cookies etc.

Note du Conseil d'Etat sur « Le numérique et les droits fondamentaux »

- Le Conseil d'Etat a trois missions : « conseiller » en donnant un avis au Gouvernement sur des projets de loi, d'ordonnance ou de décrets ; « juger » comme juridiction suprême des contentieux administratifs ; « gérer » en assurant l'administration générale des juridictions administratives. Cette année le Conseil d'Etat s'interroge en particulier sur la société numérique et sur **l'équilibre entre le développement du numérique et les droits des citoyens.**
- Le numérique **modifie le regard sur le statut de la personne**. Ceci nécessite une modification du régime juridique.
- Le rapport du Conseil d'Etat du 9 septembre 2014, **« Le numérique et les droits fondamentaux »**, n'accorde pas le droit de propriété aux données personnelles. Il préfère le recours au **principe d'autodétermination informationnelle**. Ceci signifie que tout individu doit avoir la liberté et le droit de «*décider de la communication et de l'utilisation de ses données à caractère personnel.* » (Proposition n°1 du Conseil d'Etat).
- Il n y' a pas que la personne humaine qui discute sur Internet, mais aussi les outils. Il faut donc s'interroger non seulement sur le statut des données personnelles mais aussi sur celui des machines qui traitent ces données personnelles.
- Par ailleurs la proposition n°12 du rapport propose la **promotion d'une économie des données personnelles responsable**, se développant avec les garanties appropriées : *« Proposition n°12 : afin de sécuriser le Big Data en Europe, maintenir sans ambiguïté dans la proposition de règlement européen la liberté de réutilisation statistique des données personnelles, quelle que soit la finalité initiale de leur traitement, en prévoyant pour seule condition que cette réutilisation soit entourée de garanties d'anonymat appropriées ».*

2. LA VALORISATION DES DONNÉES PERSONNELLES

Dans le monde numérique, la valorisation des données personnelles ne se limite plus au stockage et à la gestion de fichiers clients, elle n'a aujourd'hui de sens que dans la circulation et l'interaction des données avec tout un écosystème qui peut être interne et/ou externe à l'entreprise. Cela ne va pas sans poser de nouvelles questions pour les régulateurs. La CNIL a en effet soulevé ce nouvel enjeu caractérisant le *Big Data*, qui est l'interopérabilité des données : *« Si avec la naissance de la CNIL, les enjeux étaient le recueil de données pour de mauvais usages, aujourd'hui c'est l'interopérabilité qui est la nouvelle question centrale* »[9].

Dans ce contexte *Big Data*, les données personnelles et leur valorisation prennent une importance croissante. D'après une étude du Boston Consulting Group **d'ici 2020 le marché européen des données personnelles représentera une valeur économique de mille milliards d'euros**. D'après les prévisions d'IBM, **d'ici 5 à 10 ans la plus grosse partie des données viendra non pas des usages d'internet mais directement de l'usage des objets** (Internet des objets), qui encadrent notre quotidien. Ces tendances économiques et technologiques témoignent d'un avenir prometteur pour l'économie des données personnelles, mais celle-ci ne se fera pas sans le concours des consommateurs/clients qui prennent une part de plus en plus importante dans les processus de valorisation.

2.1. Les potentiels de valorisation des données personnelles

Les grandes entreprises ont progressivement pris conscience du potentiel de création de valeur que leur apportait l'usage des données ciblées, notamment via les technologies qui sous-tendent le *Big Data*, appliquées en particulier à la masse croissante de données qu'elles peuvent mobiliser sur leurs clients. Mieux connaître ses clients, comprendre leurs comportements et leurs attentes, anticiper leurs réactions permet de les fidéliser et de leur proposer les offres personnalisées les mieux adaptées. C'est un enjeu majeur pour toutes les entreprises, enjeu souvent inscrit au premier rang dans leur stratégie.

Parallèlement à la croissance régulière des données internes des entreprises, corrélées plus ou moins à la croissance et à la diversité des activités des grands groupes, la montée en puissance d'acteurs du Web a révélé le potentiel immense des données externes à l'entreprise. Cette « masse » jugée inqualifiable et inexploitable jusqu'aux années 2000 constitue la base de recherches des opportunités de demain.

Grâce aux moyens et outils de mesures disponibles pour traiter ces *Big Data*, les données personnelles se révèlent être des actifs inédits et précieux pour les entreprises dans des domaines variés.

[9] CNIL, Vie privée à l'horizon 2020, Cahiers IP, n°01, p.20
http://www.cnil.fr/fileadmin/documents/La_CNIL/publications/DEIP/CNIL-CAHIERS_IPn1.pdf

2.1.1. Stratégiques

Les modèles de valorisation des données personnelles ouvrent de plus en plus la voie aux alliances stratégiques, à la co-création de services et/ou de produits, au partage de données avec des entreprises tierces et en *BtoB*. Les secteurs de l'automobile et de l'assurance sont par exemple amenés à travailler ensemble sur la voiture connectée par l'intermédiation de plateformes numériques. C'est aussi le pari qu'à fait Nike en industrialisant un produit –des chaussures- avec des services associés, c'est-à-dire la création d'une plateforme qui collecte, centralise et met à disposition des utilisateurs les données. C'est cet atout stratégique qui a fait la différence et lui a accordé un avantage concurrentiel sans précédent, avec au cœur, la donnée personnelle ou « de bien-être » comme actif stratégique.

2.1.2. Opérationnels

Via l'anonymisation de données, l'exploitation des données personnelles dans un contexte *Big Data* permet d'optimiser des processus opérationnels. C'est ainsi que procèdent de nombreux services de transport qui mesurent les flux de passagers après avoir anonymisé les données pour adapter leurs services en fonction; la connaissance des passages les plus fréquentés, ou les plus désertés à certaines tranches horaires permet d'adapter, parfois en temps réel, les services de surveillance, de régulation des flux, les ouvertures des guichets d'informations, des points de vente etc.

2.1.3. De recherche et développement

Il est courant que l'entreprise crée de la valeur sur les données via l'expérimentation, et donc selon un usage non défini à l'avance. Or, un des principes fondamentaux de la protection des données personnelles est d'informer la personne concernée de la finalité du traitement au moment de la collecte de ses données afin d'en obtenir un consentement éclairé. Cette finalité n'est pas toujours simple à définir à l'avance dans le cadre de l'innovation ou de la R&D. Le recours à une définition de cette finalité de manière relativement générique permet de se laisser une marge de manœuvre suffisante pour valoriser au mieux les expérimentations. De plus, tant que le contexte n'est pas lié au business, les données personnelles peuvent faire l'objet d'analyses réalisées en silos, et permettre par exemple d'améliorer la qualité d'autres données, d'affiner un modèle algorithmique, et ainsi de perfectionner les savoir-faire internes des entreprises. Certaines méthodologies *Big Data* permettent en effet d'extraire des modèles pertinents sur la base de données personnelles. Les modèles prédictifs, fréquents dans les approches *Big Data*, peuvent ainsi catégoriser/segmenter une base clients sans recourir à l'identification personnelle.

2.1.4. Marketing

La personnalisation des services et des produits à des fins marketing est un pilier de l'économie des données personnelles. De nombreuses plateformes Web sont fondées exclusivement sur la vente de *retail*. L'analyse comportementale peut porter sur des données non-structurées et permettre de développer des modèles prédictifs pour être proactif dans la relation et la connaissance client.

De nombreux modèles existent, mais nous pouvons citer en particulier les résultats rendus publics des chercheurs iraniens des universités Azad et Alzahra à Téhéran : ils ont mis au point un algorithme de modélisation (par apprentissage semi-automatisé) des comportements clients par l'intégration de données non balisées à un classificateur de données balisées. Ce modèle permet, selon les chercheurs « *d'affiner les tendances et les modèles de prédiction, en vue d'un usage marketing* ». Ce modèle d'analyse des données et de comportements clients est publié dans *Advances in Information Sciences and Service Sciences* :

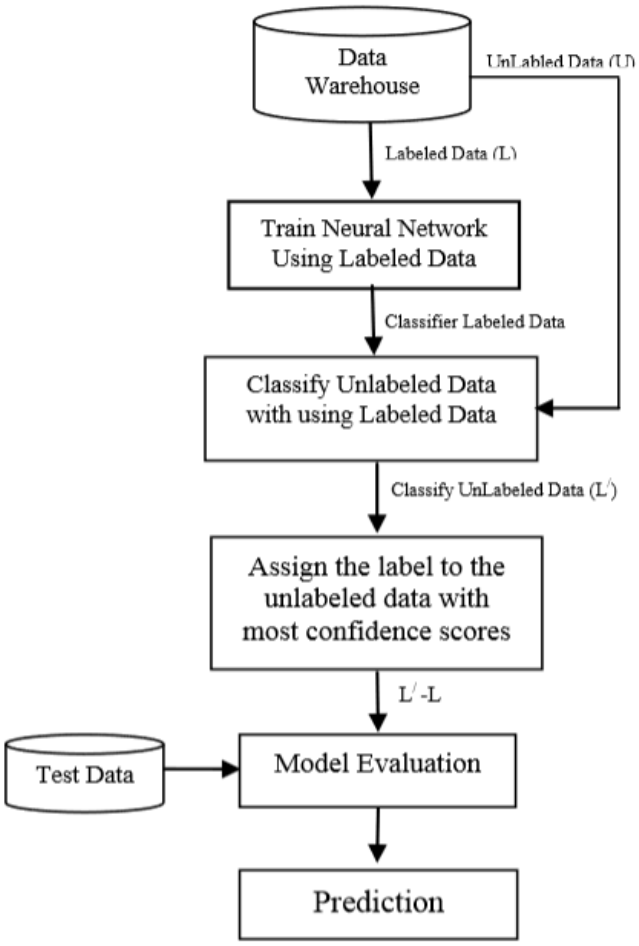


Figure 1. An overview of the Algorithm

Source: ***Customers Behavior Modeling by Semi-Supervised Learning in Customer Relationship Management*** [10]

Comme le montre ce schéma, l'algorithme repose sur un réseau neuronal utilisant des données déjà annotées. Les données non balisées sont alors traitées et classifiées par rapport au classificateur initial de données balisées. Un nouveau classificateur est ainsi créé, regroupant les données balisées et celles nouvellement balisées. Ce système permet d'acquérir un meilleur degré de confiance et de précision pour l'évaluation des modèles, notamment des modèles prédictifs.

L'usage marketing des données personnelles ne se fera pas

[10] http://www.globalcis.org/aiss/ppl/AISS%20vol.3%20.no.9_31.pdf

cependant sans le concours et le consentement des individus. Selon une enquête menée en février 2014 par serdaLAB[11] auprès de 533 individus de 15 à 65 ans, **90% des personnes acceptent les pratiques commerciales à partir de leurs données personnelles**, à condition **d'exprimer au préalable leur consentement** et de **pouvoir maîtriser les conditions de cette réutilisation des données.** Par ailleurs, les données de consommation sont jugées par les trois quarts des répondants comme étant les moins sensibles pour une réutilisation commerciale, tandis que les données de géolocalisation, de navigation internet, des réseaux sociaux, de santé et les données bancaires sont considérées comme étant les plus confidentielles. En outre, « *les individus acceptent d'être scrutés et analysés en tant eu consommateurs, beaucoup moins en tant que citoyens (opinions, réseaux sociaux, géolocalisation)* »[12].

2.1.5. Le cas de l'administration

Dans le cadre de la réforme de l'Etat, ont été créées chacune par un décret du 21 février 2011 la Direction Interministérielle des SIC (DISIC) par décret du 21 février 2011 et la mission ETALAB. Ces deux structures sont placées sous l'autorité du Premier ministre et sont toutes deux intégrées au Secrétariat Général pour la Modernisation de l'Action Publique (SGMAP).

La DISIC répond à la volonté d'assurer une meilleure gouvernance des systèmes d'information au sein de l'administration. De son côté la mission ETALAB coordonne les administrations en matière de mise à disposition de leurs données. La création d'un portail permet d'en faciliter l'accès et donc leur réutilisation par la communauté des entrepreneurs et des startups. De plus, le Premier ministre a nommé par décret du 16 septembre 2014 un responsable interministériel des données, placé sous son autorité. Au sein du Secrétariat Général pour la Modernisation de l'Action Publique (SGMAP), cet administrateur général des données est chargé de coordonner l'action des administrations en matière d'inventaire, de gouvernance, de production, de circulation et d'exploitation des données.

Ces deux structures viennent d'être réunies par décret du 21 septembre 2015 en une unique Direction Interministérielle du Numérique et du Système d'Information et de Communication de l'Etat (DINSIC), afin d'assurer une meilleure coordination entre les services liés au développement numérique et les services d'appui et d'accompagnement des acteurs publics.

S'agissant des données personnelles, l'administration travaille sous deux angles. D'une part dans le cadre de l'ouverture des données publiques, elle met à disposition les données de l'Etat mais aussi celles des collectivités territoriales qui le souhaitent à des fins statistiques mais aussi de partage des connaissances des informations, tout en garantissant par la stricte anonymisation la protection des données personnelles. D'autre part, elle accélère la mise en

[11] « Données à caractère personnel, enjeux, menaces et nouvelle donne », étude réalisée par serdaLAB, en partenariat avec le cabinet Lexing – Alain Bensoussan Avocats, 2014 (Dossier de presse)

[12] Id. p.16

place de téléservices et permet à chaque administré l'accès à ses données personnelles dans le cadre du projet France Connect, créé par arrêté du 24 juillet 2015.

L'initiative France Connect

L'ambition de « France Connect » est de déployer un système d'authentification et d'identification unique pour toutes les démarches administratives en ligne. Des expérimentations sont menées dès cette année auprès d'institutions volontaires sous le regard vigilant de la CNIL. Ainsi, lorsqu'un internaute souhaitera se connecter à un site adhérant à ce programme (celui de la CAF par exemple), un bouton lui proposera de s'identifier à partir des codes lui ayant été fournis par un autre participant à France Connect, comme les impôts. Le dispositif est également couplé à un autre projet : « Dites-le nous une fois ». L'usager pourra aussi choisir de faire profiter une administration d'informations ayant déjà été fournies à un autre organisme public (RIB, extrait Kbis...).
Si la CNIL se montre globalement favorable à la mise en œuvre de France Connect, elle en appelle néanmoins à la plus grande prudence. « *Le développement de l'administration électronique ne passe pas nécessairement et ne doit pas conduire à la création d'un identifiant unique des administrés, au plan local comme au plan national* » souligne ainsi l'institution. Avant de prévenir : « *Les traitements de données mis en œuvre dans ce cadre ne doivent pas être utilisés à d'autres fins que l'accomplissement de certaines démarches administratives, et tout particulièrement aux fins d'alimenter d'autres fichiers ou de constituer un fichier de population* ».

2.2. Les modèles d'affaires autour des données personnelles

Les données ont un impact sur les modèles d'affaires, qui peuvent ou doivent dans certains cas se redéfinir ou se réinterroger par rapport à ce nouveau contexte/environnement. L'exploitation des données personnelles a permis à certains *business models* de se réinventer, voire de se créer, impulsant ainsi une nouvelle dynamique dans l'économie numérique. En témoigne en particulier l'explosion de deux modèles économiques[13] :

- **Les modèles bifaces :** Google et Facebook sont typiquement des entreprises bifaces, c'est-à-dire qu'elles font de **l'intermédiation entre deux types d'acteurs**, **les utilisateurs et les annonceurs** : « *c'est l'existence d'une face qui représente une valeur pour l'autre face*. »[14] Ce modèle économique n'est pas propre au numérique, il existait déjà dans l'économie des médias et dans la banque de détail. De nouveaux acteurs, les ***data brokers***, ont su saisir l'opportunité de ces modèles servant d'intermédiaires à la valorisation d'informations issues d'analyses très précises sur de grandes masses de

[13] Décrit dans le livre de S.Chignard et L.D. Benyayer, *Datanomics, les nouveaux business models des données*, Fayard, 2014
[14] Id, p. 80

données (faites sur plus de 700 millions de personnes dans le monde entier). Les courtiers de données collectent, analysent et revendent leurs résultats à des sociétés visant essentiellement des fins marketing.

- **Les modèles serviciels** : les données participent grandement au mouvement de *servicialisation*, c'est-à-dire le fait de « *vendre un produit sous la forme d'un service* »[15]. L'exemple des assurances est instructif : plutôt que de spéculer sur l'accidentologie pour établir leurs tarifs, les sociétés se basent désormais sur l'analyse des usages réels des dispositifs, grâce aux capteurs et objets connectés qui transmettent les données de conduite, de kilométrage etc. Le temps réel est une vraie révolution dans la conduite des *business models* autour des données. Les modèles serviciels permettent également d'enrichir l'expérience utilisateur : il y a un « *troc implicite, données contre service* […] *les données ne servent pas uniquement à l'entreprise qui les collecte, elles bénéficient également indirectement aux utilisateurs par un meilleur service* »[16], c'est-à-dire plus personnalisé, plus efficace et pertinent. L'exemple de l'automobile est certainement le plus parlant : les projets autour de la voiture connectée supposent une alliance entre trois acteurs : un client, un garagiste, une marque de production. Google souhaite faire les OS (*Operating System*) des voitures. Les géants du web interviennent désormais dans des secteurs traditionnels et deviennent de nouveau des acteurs qui seront incontournables, à moins que les entreprises n'anticipent suffisamment tôt la construction de partenariats de confiance avec des plateformes indépendantes. **En outre, il y a une opportunité pour les acteurs européens de se démarquer des acteurs américains en développant des partenariats ainsi que des stratégies de communication basées sur la confiance. Il faut arriver à développer des services à forte valeur ajoutée en intégrant cette confiance** qui fait aujourd'hui défaut, réussir à créer la même valeur de service, mais de façon différente.

2.3. Anonymisation *versus* valorisation ?

Comment valoriser une donnée personnelle si celle-ci est anonyme ? Cette question illustre l'opposition fréquente que les entreprises établissent entre protection des données personnelles et valorisation. Il existe pourtant des méthodes qui permettent de rendre les données cohérentes, statistiquement significatives, sans être nécessairement nominatives : c'est ce qui est attendu par les entreprises et cadré par la législation.

[15] Id., p. 81
[16] Id., p. 83

Pour aider à évaluer une bonne solution d'anonymisation, le G29 définit trois critères qui déterminent les risques de réidentification :

- **L'individualisation** : est-il toujours possible d'isoler un individu ?
- **La corrélation** : est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?
- **L'inférence** : peut-on déduire de l'information sur un individu ?

A titre de compléments, la CNIL anglaise, l'ICO, a également défini un guide complet de ce processus d'anonymisation[17]. En France, la CNIL a travaillé par exemple avec de nombreux acteurs autour du *Pass Navigo*, des compteurs Linky, des téléphones *Android* dans des objectifs de conformité et d'innovation. L'exemple du *Pass Navigo* permet d'illustrer la démarche de la CNIL : dès 2009, les *Pass Navigo* étaient tous dotés d'un numéro unique, personnel et infalsifiable, permettant d'effectuer des statistiques fines sur la fréquentation du réseau. La puce du *Navigo* contenaient des informations sur le « qui », le « quand » et le « où ». La CNIL et la STIF (Société de Transport d'Ile de France) ont donc décidé ensemble d'anonymiser ces données, sans que cela n'entrave la finalité poursuivie qui était l'optimisation du réseau. La STIF pouvait donc travailler sur des données qui n'étaient plus soumises à la loi « Informatique et Libertés » car anonymes.

Il doit être cependant souligné que **l'anonymisation sera chaque jour plus complexe, notamment avec l'émergence du temps réel**. Les progrès techniques et la multiplicité des sources des données fragilisent en effet la durabilité de l'anonymisation des données et mettent en cause les investissements qui seraient faits dans ce sens. Il faut notamment prendre garde à la notion de « jeu de données anonymes » car il toujours possible de le « désanonymiser » dans un contexte *Big Data*. Un jeu de données anonymes peut donc être de nouveau soumis à la loi Informatique et Libertés. **L'anonymisation est quelque chose qui doit être sans cesse révisé dans le temps.**

Au regard de ces nouvelles complexités, **assurer une bonne gestion des données personnelles par l'exercice du droit des personnes semble une approche beaucoup plus réaliste et économique**. Rappelons en effet que l'usage de données personnelles n'est pas interdit, il est juste soumis à la loi Informatique & Libertés. Cela nécessite un consentement éclairé du client, établissant un premier contrat dans cette relation. Respecter la loi Informatique & Libertés n'est pas synonyme de contraintes insurmontables : la CNIL dispose d'un pouvoir d'autorisation, non d'interdiction. Elle autorise bien plus souvent qu'elle ne sanctionne, et s'interroge avant tout sur **la proportionnalité de la collecte et la capacité de choix de l'utilisateur de s'y opposer.** Si ces conditions sont réalisées, les données personnelles pourront être valorisées sans avoir perdu leur noyau « identifiant ». En outre, comme le souligne Ann Cavoukian dans le Livre Blanc « *Using Privacy By Design to Achieve Big Data Innovation* » : *« **Savez-vous pourquoi on ne peut pas avoir des données de qualité avec le*** **Big**

[17] Le guide est disponible ici : https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/

Data *sans* privacy *? Parce que le contexte est la clé du* Big Data ». Les données personnelles font partie de ce *contexte* qui va permettre de valoriser tout un écosystème de données. **Les données personnelle sont bien un facteur de réussite du *Big Data*.** Voilà pourquoi, le consentement des personnes et la mise en œuvre de dispositifs leur permettant d'exercer leurs droits seront certainement toujours plus rentables.

Néanmoins, les conditions du consentement pourraient sembler être elles aussi de plus en plus complexes avec le *Big Data* : « *A l'ère du* Big Data*, une grande partie de la valeur des données naît d'utilisations secondaires qui peuvent avoir été inimaginables lorsque les données ont été recueillies, ce qui signifie que le mécanisme de « notification et de consentement » pour assurer la confidentialité n'est plus adapté* »[18].

Si les finalités évoluent dans le temps, la question est de savoir **comme adresser cela à l'utilisateur ?** Une solution pourrait être de réfléchir **à des plateformes**, à des canaux de communication privilégiée qui vont **permettre aux personnes de solliciter elles-mêmes leurs droits** (informations, recueillement du consentement etc.).

2.4. Les innovations technologiques : quel avenir pour les données personnelles ?

Certaines innovations technologiques qui s'intègrent peu à peu sur le marché s'orientent de plus en plus vers la quantification et la mesure des activités humaines, du corps, et vont également impacter les modèles décisionnels : Internet des objets, intelligence artificielle, *machine to machine.* C'est un marché qui confirme la tendance généralisée de collecte de données qui entrent de plein pied dans la sphère intime des individus. Les données personnelles vont-elles finalement *driver* l'économie numérique ?

[18] Viktor Mayer-Schönberger et Kenneth Cukier, *Big Data* : une révolution qui va transformer notre façon de vivre, de travailler et penser, Houghton Mifflin, 2013

Bracelets connectés, T-Shirts connectés, lunettes intelligentes, semelles connectées, casques connectés, intégrant chacun des capteurs de données, des *Brain Computer (interfaces neuronales)*, permettant l'analyse en temps réel. Dans ce contexte, une question demeure : l'individu sera-t-il autre chose que des données ? Telle est l'interrogation que soulève un article du Financial Times « *We are data : the Future of machine intelligence* », daté du 19 juillet 2015. C'est en tout cas une question légitime à se poser lorsque l'on observe le « sur-homme » de l'affiche « Amazing French Tech » : la technologie devient de plus en plus miniaturisée et permet de vêtir l'ensemble du corps de capteurs et d'augmenter l'homme. La captation de l'individu et de son environnement ouvre la porte au développement des modèles serviciels, qui seront davantage tournés vers le prédictif. Quoiqu'il en soit, il faut s'attendre à ce que tous les secteurs, toutes les industries soient touchées par cet outillage connecté de l'homme. Les hôpitaux auraient-ils par exemple vocation à se transformer en immense *datacenter* analysant les données des patients, gérées par un système d'IA, offrant ainsi des suivis personnalisés ?

Pour aller un peu loin, la fusion du cerveau et des machines est également un scénario du futur très probable. C'est en tout cas ce qu'affirme Joël de Rosnay, dans « 2020, les scénarios du futur » en parlant de l'*affective computing* : « *on pense que l'ordinateur va apprendre à ressentir les émotions des personnes avec lesquelles il est en contact […] On va vivre une communication différente entre l'homme et les machines et ce bouleversement est pour demain […] Des centaines de laboratoires dans le monde travaillent sur ces projets. Ils tentent de faire fusionner le cerveau et les machines* »[19]. L'interaction homme-machine serait ici à son paroxysme ; l'homme n'existerait plus sans ses outillages technologiques qui l'habillent, le

[19] de Rosnay, Joël, 2020, Les Scénarios du Futur, Comprendre le Monde qui Vient, Fayard, 2008, p.227

guident, surveillent sa santé, son bien-être, sa performance etc. Les investissements technologiques semblent bien être tournés de plus en plus vers cette personnalisation exacerbée de l'homme et de son augmentation.

Quoiqu'il en soit, le défi majeur pour les entreprises et acteurs du numérique sera dans « *la détermination de la responsabilité [qui] nécessite autant la compréhension que l'intégration du futur, sans savoir véritablement ce qu'il sera… et c'est bien là où se trouve la complexité.* »[20]

[20] Pavie, X., Innovation responsable, Stratégie et levier de croissance des organisations, Eyrolles, Paris, 2012.

3. LES ENJEUX DE LA RÈGLEMENTATION SUR LA PROTECTION DES DONNÉES PERSONNELLES FACE AU MARCHÉ

Pour les entreprises, l'investigation des données fait partie intégrante du processus de valorisation. Une chose est sûre, la loi doit être appliquée à toutes les étapes de l'utilisation des données personnelles. Si la loi Informatique & Libertés est perçue comme étant contraignante pour le développement économique, elle semble pourtant bien plus souvent converger avec les intérêts des organisations, au regard de ce qu'il est possible de définir à l'avance, comme un contrat avec l'utilisateur, comprenant des règles de bon sens et de sécurité. Par ailleurs, l'adaptation de la règlementation à l'évolution technologique est une problématique à part entière : de nombreuses solutions technologiques et pratiques interpellent tous les jours la règlementation, comme par exemple le *fast data* qui consiste à faire des traitements très courts, sans que les données ne soient conservées une fois le traitement exécuté. Pour répondre à cet enjeu, la règlementation commence à se réinventer autour de principes ou d'outils plus agiles par rapport au caractère dynamique des *Big Data* et à l'impératif de compétitivité qu'ils représentent.

3.1. Innovation et conformité

Pour que le V de Valeur soit au rendez-vous, il y a donc quelques incontournables à respecter et avant toute autre chose, quelques idées fausses à combattre. Première idée fausse : « il est impossible de faire du *Big Data* avec la loi Informatique & Libertés ». Le *Big data* interroge le principe de finalité, qui semble contradictoire avec la nature même du *Big Data*. Pourtant, vus les outils et les investissements déployés pour un projet *Big Data*, il est évident qu'un objectif est défini. Il y a nécessairement *a minima* des hypothèses qui sont formulées. **La logique de finalité n'est donc pas incompatible avec le *Big Data*.**

Isabelle Falque-Pierrotin, Présidente de la CNIL, lors d'une conférence donnée en juin 2014 disait à ce propos : *« L'idée que la CNIL s'opposerait aux aspects positifs en termes d'innovation qu'apporte le* Big Data*, n'est pas étayée par la réalité* (...) *Ces sujets (l'interconnexion de fichiers, la segmentation toujours plus fine des clients), ne sont pas en soi contraires à la Loi Informatique et Libertés. La seule chose que demande celle-ci, ce sont des garanties en termes de liberté des individus* ».

Développer des règles et des lois est indispensable mais il faut être en mesure de les appliquer avec l'agilité et la rapidité des entreprises numériques. Les autorités de régulation sont bien au fait de cet enjeu et proposent des solutions pertinentes. La CNIL a par exemple mis en place **des « boîte à outils conformité » qui répondent à des objectifs d'innovation**. Cela permet **une innovation plus durable** selon Isabelle Falque-Pierrotin : « *Respecter la Loi Informatique et Libertés est **un moyen de construire quelque chose de plus solide.** On a construit un pacte de conformité permettant à l'industrie d'être en conformité de façon constante avec la loi. **On réinvente la régulation non pas dans les principes mais dans les outils qu'on met à***

disposition des entreprises ». En outre, l'accompagnement de l'innovation est le maître mot du régulateur, qui est passé « d'une logique de formalités administratives à une logique de conformité tout au long de la vie du traitement de données. »[21] A titre d'exemple, la constitution par la CNIL d'un « pôle innovation et prospective » en 2012 permet d'être « *en contact avec de très nombreuses start-up ou pépinières de talents, [que la CNIL] conseille dès la conception de leurs produits pour une protection optimale des données* »[22]

Finalement le régulateur permet beaucoup de choses et s'adapte de la manière la plus agile possible face aux enjeux de l'innovation.

3.2. Convertir les contraintes en avantages concurrentiels

Si le droit ne peut ignorer la révolution numérique, il est clair que la révolution numérique ne peut ignorer le droit. Selon Thomas Beaugrand, Avocat chez Staub & Associés, la réticence des acteurs économiques concernant la réglementation sur les données personnelles s'explique, selon les propos issus du terrain, par le fait qu'elle :

- entrave certains modèles
- complique la prospection
- complique les projets de transition numérique de l'entreprise
- ralentit les entreprises EU face aux entreprises US

Pourtant, la protection des données personnelles correspond à des principes fondamentaux et devient de plus en plus importante dans une société en cours de numérisation complète. Depuis plusieurs années, le débat fait donc rage sur l'équilibre à trouver entre modèles innovants (dynamiques mais parfois intrusifs) et protection de la vie privée (essentielle mais parfois pesante).

En effet, l'économie numérique étant plus que jamais mondiale, compétitive et faisant entrer toujours plus de « disrupteurs » en jeu, les entreprises européennes ont peur de se retrouver hors-jeu. Maître Beaugrand les enjoint cependant à convertir ces contraintes en avantages concurrentiels : *« les exploitants de la donnée sont confrontés à la concurrence étrangère (notamment nord-américaine) qui ne s'embarrasse pas de telles précautions, mais ils bénéficient d'une notoriété forte et d'une pénétration marché importante. Mais puisqu'ils ne peuvent pas s'en affranchir,* ***les nouveaux modèles économiques européens doivent convertir les contraintes liées à la législation de protection des données, en avantages marketing et compétitifs****, sous forme de « label qualité » par exemple.* »

Quelques associations se forment dans ce sens, telles que Cloud Confidence, qui fournit un cadre de transparence sur les règles de sécurité et de confidentialité des données, ou encore

[21] Source : http://www.cnil.fr/linstitution/ip/actualite-innovation-et-prospective/article/enjeux-2015-2-la-protection-des-donnees-cle-de-voute-de-linnovation/

[22] ibid

le label « Gouvernance Informatique & Libertés » proposé par la CNIL, qui sont à la fois des aides et des marques de confiance. Par ailleurs, des initiatives de type « *privacy by design* » sur lesquelles nous reviendrons plus loin dans ce livrable, portent sur la conception de solutions qui protègent intrinsèquement les données personnelles, mais sont encore à éprouver au contact d'une législation changeante.

3.3. Les bonnes pratiques pour l'accompagnement juridique du marché

Selon Thomas Beaugrand, Avocat chez Staub & Associés, il convient de suivre quelques bonnes pratiques pour se prémunir des risques juridiques et développer un climat de confiance avec les clients :

- **Transformer en argument de vente le fait de se conformer à une législation** : à côté de la rentabilité économique et de l'exploitabilité technique des données, la validation des exigences légales permet d'augmenter la valeur du fichier par validation de la confiance associée. En clair, puisque les exigences légales européennes s'imposent, les entreprises doivent s'assurer de « *l'exploitabilité juridique* » des données qu'elles collectent, et apporter ainsi une certification complémentaire à leurs clients.
- **Miser sur *l'empowerment***, la responsabilisation ou l'implication des individus : ce sont des notions qui redonnent une forme de contrôle des personnes sur leurs données personnelles. Cela nécessite avant tout une meilleure connaissance de ce qu'est une donnée personnelle et de ce qui peut en être fait. **Le marché est davantage sur le contributif, sur le partage de valeur**. Si les données personnelles sont une émanation de l'individu, celui-ci devient partie prenante de la valeur créée. C'est ce paradigme qui va redonner de la confiance, encore faut-il lui donner une application efficace et équitable sans tomber dans les impasses que connait par exemple le droit d'auteur, ni dans la captation généralisée pratiquée par certains leaders du *cloud* grand public.
- **Avoir une approche pragmatique** : il s'agit de vérifier si les traitements concernent des données identifiantes, et d'évaluer les bénéfices apportés au client final (faire le bilan bénéfices/risques) pour mettre en place une législation moins monolithique, qui prenne en compte les spécificités des modèles économiques fondés sur le *big data* et les données purement techniques tant qu'ils n'aboutissent pas à la constitution de fichiers dangereux ou trop vastes...
- **Se poser la question de la légitimité et de la proportion** : cette question est prise en charge au niveau de la loi : est-on légitime, au regard des finalités spécifiées, de collecter telles ou telles données ? (services publics, institutions de santé, de recherches…) Les données recueillies sont-elles proportionnées au but poursuivi, ou exposent-elles trop la vie privée de la personne physique au regard de l'enjeu pour l'entreprise ?

- **L'essentiel pour les entreprises européennes est d'assurer un traitement « *adéquat* » des données, conforme aux exigences légales européennes, où que soit hébergée la donnée dans le monde :**
  - Recourir à des prestataires « conformes » (*Safe Harbor*, clauses contractuelles types). Certaines associations se sont formées pour labelliser certaines offres et/ou développer un cadre de confiance entre clients et prestataires. L'association Cloud Confidence par exemple a pour objectif de mettre à disposition des prestataires Cloud et des clients un cadre de transparence sur la protection des données clients.
  - Déployer une politique homogène à l'échelle du groupe quelles que soient ses implantations (via les règles internes de l'entreprise ou « *Binding Corporate Rules*)
  - Prévoir des procédures de signalements en interne.

3.4. Le projet de règlement européen sur la protection des données personnelles

La Directive 95/46/CE[23] actuelle sur la protection des données à caractère personnel fait l'objet d'une révision par la Commission européenne. Elle vise l'établissement d'un règlement ayant pour objet de s' «*adapter aux nouvelles exigences liées au développement rapide des nouvelles technologies et à la mondialisation tout en veillant à ce que les individus gardent un contrôle effectif sur leurs propres données : la proposition de la Commission concernant un règlement général sur la protection des données permettra d'harmoniser davantage les règles en la matière existant dans l'UE et, par l'établissement d'un niveau élevé et uniforme de protection des individus, garantira la sécurité juridique aux entreprises et suscitera une confiance accrue dans les services […] en ligne*. »[24]

3.4.1. Cadre général

Le futur Règlement Européen signifie entre autres, selon Isabelle Falque-Pierrotin, Présidente de la CNIL :

- **un cadre juridique unifié**, **un marché européen de la donnée**, **une « Europe de la donnée »**. Sur ce point, Günther Oettinger, Commissaire pour l'économie et la société numériques, précise : « *L'Europe ne pourra pas être à l'avant-garde de la révolution numérique si 28 réglementations différentes continuent à coexister dans le domaine des services de télécommunications, des droits d'auteur, de la sécurité informatique ou de la protection des données. Nous avons besoin d'un marché européen qui permette à de*

[23] Directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, JO L 281 du 23.11.1995, p.31

[24] Commission Européenne, *Livre vert sur la santé mobile*, 2014, p.9

nouveaux modèles économiques de s'épanouir, aux start-ups de se développer et à l'industrie de tirer parti de l'internet des objets »[25] .

- la capacité de **mettre au même niveau de concurrence les acteurs européens et les acteurs internationaux**, car ces derniers seront soumis au droit européen quand ils mettront en place un service à destination d'un citoyen européen.

Le régulateur qu'est la CNIL doit changer de posture pour accompagner ces évolutions : « *Avant, le métier de la CNIL était un métier d'encadrant des fichiers. Son activité était ponctuelle. Mais si les données sont partout, et si on veut intégrer le respect des personnes dans ce système transversal et universel de la donnée,* ***il faut une relation avec le régulateur qui soit plus continue et plus relationnelle****. La CNIL a une mission de conseiller et déploie certaines équipes spécialisées pour accompagner les projets qui posent question sur le respect des principes Informatique et Libertés. Avec ce règlement, il faudra trouver un équilibre entre deux ambitions* a priori *contradictoires : d'un côté* ***mieux protéger les données*** *numériques des personnes. Et de l'autre,* ***favoriser la compétitivité des entreprises****, notamment de celles dont le modèle économique est basé sur l'exploitation de ces données*»[26] .

Parmi les principes fondateurs de ce projet de règlement, il faut noter :

- Le principe *d'Accountability :* il s'agit de s'engager sur la conformité de l'organisme en répondant aux exigences de la règlementation sur la protection des données personnelles, d'assurer la traçabilité des actions des entreprises et d'expliquer les mesures mises en œuvre pour s'y conformer. Le RSSI sera particulièrement sollicité pour documenter les démarches de protection, *a posteriori*. L'idée est d'assurer un suivi de la traçabilité des données personnelles et une véritable transparence vis-à-vis de la règlementation.
- Le *Privacy Impact Assessement* ou l'Etude d'Impact sur la Vie Privée : ce principe est déjà intégré dans la loi Informatique & Libertés. Il vise à évaluer *a priori* les risques associés au développement d'un produit/service en matière de protection de la vie privée. Identifier les risques à l'avance permettra d'élaborer ensuite les moyens adéquats à mettre en œuvre pour les parer.
- La *Privacy by Design :* il s'agit d'intégrer le respect du cadre règlementaire dès la création et tout au long de la chaîne d'un service/produit reposant sur un traitement de données personnelles. Les moyens de consentement des personnes doivent par exemple être prévus dès le départ.
- Le *Data Protection Officer* (DPO) qui étend les fonctions et responsabilités actuelles du CIL en passant, selon Alain Bensoussan, Avocat « *d'une obligation d'assurer la conformité (article 22 de la Loi française) à une obligation de contrôler au sens de surveiller et d'auditer l'entreprise. Ce sont les premiers pas d'un commissaire aux*

[25] Cité sur : http://europa.eu/rapid/press-release_IP-15-4653_fr.htm

[26] Isabelle Falque-Pierrotin, Matinée de l'IFA

données. Il s'agit là d'une modification fondamentale du règlement qui fait du DPO, un relais des autorités de protection des données à caractère personnel »[27] .

3.4.2. Discussions et positionnement des entreprises utilisatrices

Traiter de la donnée personnelle se faisant de moins en moins seul au sein de sa propre infrastructure privée, mais plutôt via des partenaires (par exemple des fournisseurs *Cloud* ou de solutions *Big Data*), la notion de la co-responsabilité est cruciale dans la mesure où elle permet de préciser ensuite qui est responsable de quoi (par exemple, dans le cadre de la fourniture de services *cloud* public ou *Big Data*, où les moyens de sécurité sont conçus et mis en œuvre par le fournisseur de manière uniformisée pour l'ensemble de ses clients.) Une entreprise ne devrait pas être sanctionnée d'une amende de 5% de son CA si c'est un tiers qui commet l'erreur de sécurité contribuant à la révélation de données personnelles. Se tromper de fautifs en matière de sanctions fragiliserait grandement l'écosystème. De ce fait, la notion de « *joint controller* » (article 24) reste clef dans l'impact de ce futur Règlement sur l'économie de la donnée personnelle, tout comme la propriété de la donnée et la propriété intellectuelle que la Commission souhaiterait revoir au risque de fragiliser des pans entiers de l'économie européenne sous le *lobbying* des géants qui tirent profit financier de la donnée la plus libre possible.

Les obligations en matière de transfert de donnée personnelle hors de l'UE méritent aussi d'être clarifiées et simplifiées, tout en renforçant leur sécurité.

Enfin, il est fondamental que les contraintes que le futur Règlement fera peser sur les uns et les autres soient véritablement négociables notamment par les entreprises utilisatrices, ce qui est souvent compliqué face aux géants de l'industrie numérique. Une autre question porte sur le concept de « Responsable de traitement » de la donnée ; en effet si la responsabilité continue à peser sur l'entreprise utilisatrice, pourquoi le fournisseur ferait-il des efforts pour aligner son organisation et ses contrats sur des éléments qui ne pèsent pas sur lui ? L'entreprise utilisatrice se retrouverait alors coincée entre ses obligations et l'incapacité de les imposer contractuellement, mais le futur Règlement Européen pourrait anticiper cela et le résoudre.

3.5. Le cas particulier des données de santé et des données de bien-être

Les données de santé sont considérées par le droit européen comme des données sensibles (avec les données raciales, d'opinions politiques, religieuses, philosophiques, syndicales ou relatives à la vie sexuelle), et sont par conséquent soumises à un régime spécial. Les données

[27] Source : http://www.videosurveillance-infos.com/Tout-sur-le-projet-de-reglement-europeen-sur-la-protection-des-donnees-et-le-role-du-DPO_a1472.html

de santé personnelles faisant appel au secret médical, sont en effet interdites de collecte et de traitement par la loi, sauf si celles-ci sont anonymisées en prenant en compte les moyens possibles de réidentification. Certains traitements peuvent également être autorisés dans certaines conditions (s'ils relèvent de l'intérêt général, si la personne a donné son consentement explicite, s'ils correspondent à un objectif de recherche ayant une mission de service public…). La collecte et le traitement de ces données doivent être traités au cas par cas en fonction des objectifs poursuivis.

3.5.1. Données de bien-être ou données médicales, quelles frontières ?

Avec l'émergence du marché des objets connectés et des pratiques du *quantified self*, un nouveau type de données apparaît, se rapprochant des données de santé, mais n'étant pas tout à fait similaires : ce sont les « données de bien-être et de mode de vie » issues notamment des bracelets connectés, podomètres, tensiomètres, moniteurs de fréquences cardiaques, ou tout autre capteur mesurant une activité corporelle. Le statut de données de bien-être est actuellement en débat au sein de la Commission européenne. L'hébergement de ces données n'est en effet pas soumis actuellement à la législation sur le droit des données de santé qui, elle, exige que les données soient gérées par un prestataire agréé par l'Agence des systèmes d'information partagés de santé.

« *A l'heure actuelle, ces informations [données de bien-être] sont considérées comme des données ne relevant pas du domaine médical même si elles permettent d'enregistrer des constantes pouvant donner lieu à des analyses et diagnostics précis. En effet les informations de « fitness » ou de « wellness » ne sont pas soumises aux règles applicables aux données médicales* »[28]

Cependant, la Commission met à disposition des développeurs et fabricants de logiciels des orientations, ou « normes harmonisées » non obligatoires pour le traitement des données de bien-être, qui sont régulièrement mises à jour afin de leur permettre de déterminer si leurs produits relèvent ou non de la directive sur les dispositifs médicaux[29] ou de la directive portant sur les dispositifs médicaux de diagnostic *in vitro*. Selon les recommandations de la Commission, cette distinction doit se faire sur la base des finalités définies par les applications. Si celles-ci répondent à la définition de « dispositif médical » ou d'un « dispositif médical de diagnostic *in vitro* », elles devront se conformer aux directives susmentionnées.

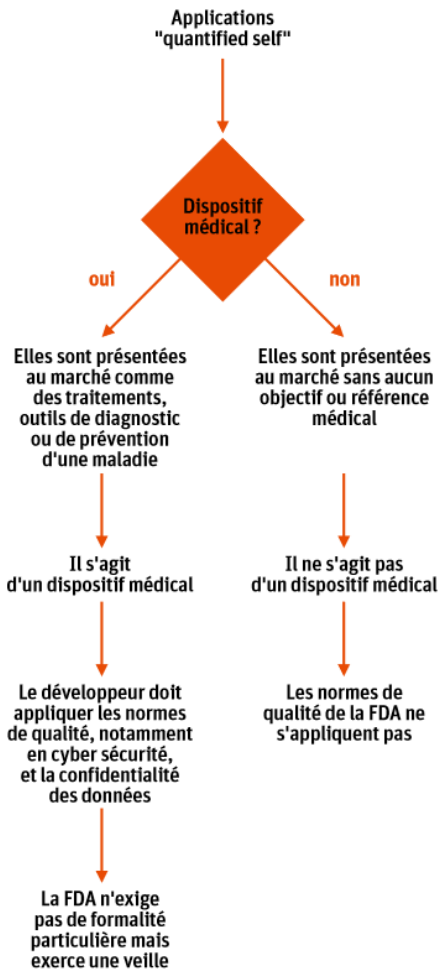
Pour être exploitées conformément aux « normes harmonisées » de la Commission, les données de bien-être devraient en outre se limiter à des données physiologiques brutes ne permettant pas d'aboutir à des informations ou diagnostics médicaux. Les enjeux techniques

[28] Maxime SALAUN, Source : http://blog.economie-numerique.net/2015/07/24/donnees-de-sante-et-donnees-de-bien-etre-orientations-europeennes/

[29] Directive 93/42/CEE relative aux dispositifs médicaux, JO L 169 du 12.7.1993. Cette directive est en cours de réexamen pour devenir un règlement.

et organisationnels sont forts, puisqu'il s'agit d'empêcher certains croisements qui permettraient, via le *data mining* par exemple, de mener au diagnostic médical.

Comme le montre le schéma ci-contre, tiré du Cahier Innovation & Prospective n° 2 de la CNIL « Le Corps, nouvel objet connecté », la FDA (*Food and Drug Administration*), aux Etats-Unis, s'est fixé cette grille d'analyse afin de déterminer les conditions selon lesquelles une application peut être considérée comme un dispositif médical ou non. La FDA a clarifié ces règles notamment parce qu'en 2013 elle a dû demander à la société génomique *23andMe* de cesser la commercialisation d'un de ses kits, qui avait pour objectif premier de faire des recherches généalogiques à partir des échantillons de salive des utilisateurs. Mais ce kit s'est ensuite réorienté vers un objectif dédié à la santé des utilisateurs : ceux-ci ayant donné leur génome, il était donc possible de déterminer leur prédisposition à certaines maladies, c'est pourquoi « […] *la FDA a décidé de stopper la commercialisation du kit, en l'absence d'informations suffisantes sur la fiabilité des diagnostics et conseils de prévention donnés aux usagers. En effet, les résultats transmis pouvaient leur faire courir de sérieux risques en cas de mauvaise interprétation* ».



3.5.2. La m-Santé, au cœur de l'économie numérique

Par ailleurs, la Commission européenne encourage et reconnait les opportunités en termes économiques de la valorisation des données de santé mobiles : celles-ci permettraient, grâce à une plus forte implication des personnes *via* leurs applications de santé, une meilleure prévention (permettant de « *mieux planifier les interventions, d'éviter les consultations inutiles et de mieux préparer les professionnels* »[30]), ce qui allègerait les charges financières des systèmes de santé de l'UE. De plus, selon une étude de GSMA et PwC, citée dans le « livre vert sur la santé mobile » de la Commission européenne : « *le marché mondial de la santé mobile atteindra l'équivalent de 23 milliards de dollars en 2017, l'Europe représentant 6,9*

[30] Livre vert sur la santé mobile, p.5

milliards et l'Asie-Pacifique 6,8 milliards, devant le marché américain estimé à 6,5 milliards »[31]. Cet essor économique est dopé par le marché des applications mobiles et les *smartphones*. D'après certaines estimations récentes[32], près des trois quarts des applications existantes sur le marché sont consacrés au bien-être et un quart est destiné aux professionnels de santé afin de faciliter le suivi des patients et l'accès à leurs données.

Il est à noter que, selon ce rapport de la Commission européenne, les principaux acteurs de cette économie sont des petites structures : « *30% des sociétés de développement d'applis mobiles sont des entreprises unipersonnelles et 34,3% sont de petites entreprises (définies comme ayant de 2 à 9 employés)*».[33]

Cependant, l'essor de ce marché économique ne pourra se faire qu'en tenant compte de la confiance qu'auront les utilisateurs dans les traitements effectués par les entreprises ou autorités publiques, en mettant en place des dispositifs adaptés de transparence et de sécurité des systèmes : en effet, 45% des consommateurs, soit presque la moitié de la population, se disent inquiets concernant l'utilisation de leurs données de santé collectées *via* un dispositif mobile.[34]

3.6. La localisation des données personnelles

Les données personnelles sont des actifs sensibles – comme toute information confidentielle de quelque nature qu'elle soit – au point d'impacter l'encadrement légal sur leur localisation territoriale.

Comme le souligne l'AFNOR dans son Livre Blanc *« Données Massives* – Big Data – *Impact et Attentes pour la Normalisation »* paru en juin 2015 *: « le cadre règlementaire qui s'applique en matière de localisation des données complique leur exploitation et leur stockage. En effet, la réglementation peut différer d'un pays à l'autre, même en Europe. Il convient de noter toutefois que le projet de règlement en matière de données personnelles vise à uniformiser les règles au sein de l'Union Européenne »*[35].

Le récent projet de loi de la Douma russe a relancé le débat sur la localisation des données personnelles des entreprises mondiales. Cette loi exige que les données personnelles des clients russes ainsi que tous les traitements réalisés sur ces données soient stockées sur le territoire russe, et ce à compter du 1er septembre 2016. Or, à ce jour, les données des clients

[31] GSMA et PwC, *Touching lives through mobile health – Assessment of the global market opportunity*, Février 2012.

[32] Research2Guidance, *The mobile health global market report 2013-2017: the commercialisation of mHealth apps*, vol.3, 2013.

[33] IDC, Worldwide and U.S. Mobile Applications, Storefronts, Developer, and In-App Advertising 2011-2015 Forecast : Emergence of Postdownload Business Models, In « Livre vert de la santé mobile », 2014, p.7

[34] Blue Chip Patient Recruitment, Leveraging Mible Health Technology for Patient Recruitment, octobre 2012.

[35] AFNOR, Livre Blanc *« Données Massives* – Big Data – *Impact et Attentes pour la Normalisation » Juin 2015, p.45*

russes de nombreux grands groupes français sont centralisées en France. Les enjeux financiers, industriels et techniques sont gigantesques.

Mais cette loi s'inscrit dans un mouvement général : en Allemagne, en Suisse et en France, de nombreux opérateurs et services d'hébergement (Deutsche Telekom, Swisscom, Numergy, Cloudwatt…) soutenus par les gouvernements et les agences de sécurité informatiques (l'ANSSI pour la France) ont créé des solutions de ce type. La garantie d'un stockage et d'un traitement sur le territoire des clients est devenue un gage de respect de la protection (confidentialité et sécurité) des données personnelles et de confiance. Mais c'est aussi une opportunité majeure pour la création d'emplois qui se développent localement pour subvenir aux besoins des services et/ou produits en question.

Néanmoins, la localisation territoriale ou « souveraine » des *datacenters* ou *Cloud*, soulève plusieurs questions :

- Dans le cadre du *FISA Amendments Act* de 2008 aux Etats-Unis et de la loi de programmation militaire n°2013-1168 du 18 décembre 2013 en France, qui facilitent les modes d'accès aux données personnelles sans contrôle d'un juge, le protectionnisme des données personnelles ne va-t-il pas à terme favoriser le contrôle des autorités sur la vie privée des citoyens ?
- Les possibilités de *hacking* s'étendent aussi via les flux de données transitant par le réseau ; ainsi la localisation territoriale n'est pas une garantie optimale de sécurité ;
- Il faut également s'interroger sur « la conformité de la localisation territoriale des données à la Directive 95/46/CE du Parlement européen et du Conseil qui énonce dans son préambule que « (…) *l'établissement et le fonctionnement du marché intérieur dans lequel, conformément à l'article 7 A du traité, la libre circulation des marchandises, des personnes, des services et des capitaux est assurée, nécessitent non seulement que des données à caractère personnel puissent circuler librement d'un Etat membre à l'autre, mais également que les droits fondamentaux des personnes soient sauvegardées* » et que « (…) *le renforcement de la coopération scientifique et technique ainsi que la mise en place coordonnée de nouveaux réseaux de télécommunications dans la Communauté nécessitent et facilitent la circulation transfrontalière de données à caractère personnel* »[36].

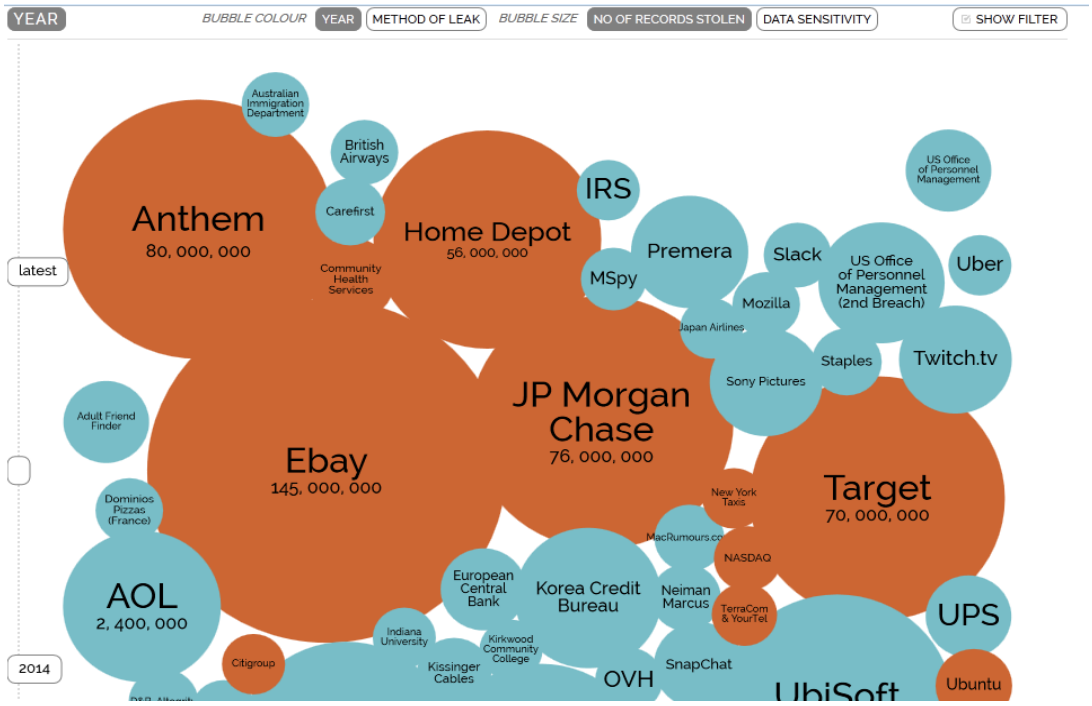
Si certains principes semblent être encore imprécis dans le paysage de la réglementation sur les données personnelles, qui soulèvera toujours des contraintes techniques et économiques, il faut tout de même noter la volonté affirmée du cadre législatif à accompagner du mieux possible les exigences du marché en adaptant ses méthodes pour soutenir l'innovation tout en étant conforme. Quoiqu'il en soit, respecter la loi relative à la protection des données

[36] Source, August & Debouzy avocats, « La localisation territoriale des données : quelles réalités pour quelles garanties » : http://www.august-debouzy.com/fr/flash/00980-la-localisation-territoriale-des-donn-es-quelles-r-alit-s-pour-quelles-garanties.html

personnelles concourt au maintien de la confiance dans l'économie numérique. La confiance, devient une valeur centrale dans le monde numérique caractérisé par l'incertitude, la dématérialisation et la vitesse. Mais la confiance est un capital qui s'accumule lentement, et qui est extrêmement volatil. En somme, elle doit également se construire dans la bonne gouvernance de l'entreprise, et en particulier la gouvernance de la donnée qui devient un incontournable.

4. LA GOUVERNANCE DE LA DONNÉE PERSONNELLE

Si l'utilisation des données personnelles pour créer de nouveaux services constitue pour les entreprises un vaste champ de développement et de création de valeur, exploiter ces données peut donc exposer l'entreprise à des risques et être source de destruction de valeur. Les risques concernant les données personnelles sont des risques beaucoup plus importants que par le passé et très diffus : l'actualité avec ses faits et ses méfaits vient pointer du doigt les faiblesses d'une société numérique notamment avec les violations et malversations liées au piratage des données personnelles.



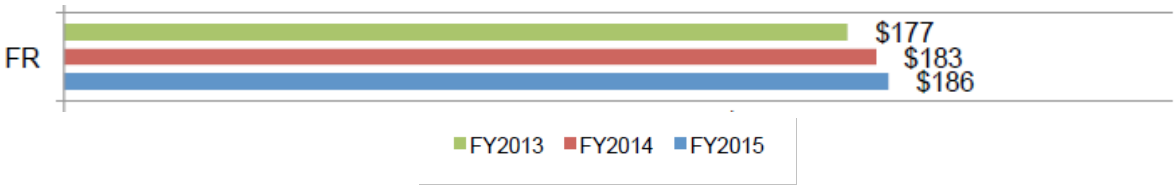
VALEUR

Nombre de données personnelles compromises (*data breaches*) par organisme 2014/2015[37]

L'Institut américain Ponemon a analysé le coût de ces violations (schéma ci-dessus). Les résultats de cette étude sont publiés annuellement et en accès libre. Selon la dernière étude publiée, le coût moyen, pour un organisme, d'une violation des données s'élève à 3,8 millions de dollars, ce qui représente une augmentation de 23% depuis 2013.

[37] Source : http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

Pour la France, le coût moyen par donnée ainsi compromise s'est élevé à 186 $ en 2015.



Coût moyen en France d'une violation de données personnelle, par année

Les conséquences pour les entreprises sont catastrophiques également en termes d'image. Les entreprises sont bien conscientes que leurs erreurs ont un coût et que la pérennité de leur activité économique impose une meilleure gouvernance des données de leurs clients.

4.1. Les référentiels existants

4.1.1. Assurer la conformité : Label de gouvernance des données - CNIL

La CNIL a adopté le 11 décembre 2014 un nouveau référentiel s'adressant à tous les organismes doté d'un CIL, et qui délivre les procédures d'une « bonne » gouvernance des données personnelles. Cette certification permettra entre autres de préparer les détenteurs au principe *d'accountability*[38] présent dans le projet de règlement européen sur la protection des données personnelles. L'obtention du label est un indicateur de confiance pour les clients. En pratique, les 25 exigences de ce nouveau référentiel sont organisées en trois thématiques qui concernent :

- **l'organisation interne** mise en place pour assurer un déploiement optimal de la politique de protection des données ;
- **la méthode** de vérification de la conformité des traitements à la loi Informatique et Libertés ;
- **les processus de gestion** des réclamations et incidents.

Les entreprises qui répondront à ces exigences seront ainsi pourvues du label « Gouvernance Informatique et Libertés ».

4.1.2. Evaluer les projets : Charte Ethique et *Big Data* – Alliance *Big Data*

La charte « Ethique et *Big Data* » développé par l'Alliance *Big Data* consiste à proposer aux pôles de compétitivité et à toute organisation réalisant des projets *Big Data*, un outil descriptif sur les données (source, traçabilité, réglementations..) leur permettant d'évaluer leurs projets. C'est un outil pédagogique qui vise non pas à dicter une ligne de conduite à suivre mais à aider les responsables des projets à y voir plus clair sur **la nature des données** qui sont

[38] Dans le contexte de la protection des données à caractère personnel, on entend par *Accountability* une obligation de rendre compte et d'expliquer, avec une idée de transparence et de traçabilité permettant d'identifier et de documenter les mesures mises en œuvre pour se conformer aux exigences issues de la réglementation sur la protection des données personnelles. (http://www.cil.cnrs.fr/CIL/spip.php?article1890)

utilisées dans un projet, sur **leur provenance**, **leur destination**, **les finalités envisagées**, **les procédés de sécurité et de confidentialité mis en œuvre**, **les divers responsables** concernés. Cette charte est disponible en pdf à cette adresse : http://wiki.ethique-big-data.org/index.php?title=Accueil

4.2. L'urbanisation des données personnelles

L'urbanisation des données personnelles est un enjeu crucial dans la gouvernance des données. Les questions techniques sont au premier plan pour assurer un bon suivi des questions règlementaires et un bon *data management*. La complexité et la taille des SI est un point problématique : la provenance, la circulation, les entrecroisements des données sont de plus en plus difficiles à identifier et à localiser. Cela devient problématique lorsqu'il est question de données personnelles. On s'aperçoit que techniquement, il est parfois compliqué d'assurer totalement le droit à l'oubli. Définir un cadre méthodologique pour répondre à cet enjeu est donc crucial.

Rappelons que la loi encadre le traitement des données personnelles et que **la notion de traitement englobe tant les activités de collecte, que celles d'exploitation, de consultation et de stockage**. Il convient donc non seulement d'identifier les données personnelles dans les SI mais également d'identifier tous les traitements auxquels elles se rattachent. Si, de toute évidence, il existe toujours des congruences entre les données personnelles et leurs traitements, ou inversement, seule la cartographie des traitements permettra de remonter jusqu'à certaines sources de données personnelles qui n'auraient pas pu être recensées autrement.

La méthodologie de cartographie doit reposer sur quelques principes ou actions [39]:

- produire un référentiel de définition des données personnelles à destination des équipes ;
- construire un *datamart* dédié aux données personnelles : il s'agit d'un ensemble de données ciblées, organisées, regroupées et agrégées pour répondre à un besoin spécifique à un métier dans un domaine donné ;
- développer un outil de visualisation des processus concernés par les données personnelles, permettant d'identifier rapidement le parcours des données dans l'entreprise : les composants SI, les lieux de stockage ou de traitements des données personnelles et leur flux ;
- mettre à disposition des Métiers un référentiel de finalités et de traitements, afin d'être en phase avec les exigences de la CNIL plus facilement. Ceci est un outil efficace qui viendra supporter le travail des CIL/DPO.

39 Selon les enseignements tirés de Luc Delpha, « Cartographie des données à caractère personnel en entreprise : quid des bonnes pratiques » Source : http://www.journaldunet.com/solutions/expert/60166/cartographie-des-donnees-a-caractere-personnel-en-entreprise---quid-des-bonnes-pratiques.shtml

- identifier les flux de données issus des partenaires ou en interaction avec eux.

4.3. La bonne gouvernance des données personnelles

Afin d'assurer une bonne gouvernance des données personnelles, il convient d'être attentif à de nombreux *process* de l'entreprise ayant recours ou étant susceptibles de recourir à des bases de données sensibles et/ou personnelles. La désignation d'un responsable des données (*data owner*), comme le suggère le rapport CIGREF « Enjeux business des données » (2014), est essentiel. Il convient tout d'abord d'assurer l'intégrité et les qualités des données dès leur source. Le travail sur l'amélioration des métriques permettra d'assurer de manière constante le niveau de qualité exigé. Dès que les données personnelles sont classifiées (selon les catégories définies par la CNIL) et que les autorisations sont accordées pour des besoins métiers ou des projets R&D, il convient d'assurer un travail de *monitoring* permanent. Le travail de cartographie des données, tenant compte de la complexité des SI, est également essentiel pour avoir une vision d'ensemble du cycle de vie des données personnelles dans l'entreprise (leur provenance, leur partage avec des tiers ou en interne, leur stockage, leur usage, leur traitement et leur suppression).

Les enjeux de cybersécurité s'intègre naturellement dans cette gouvernance. La *privacy* est en effet fortement rattachée à une approche par le risque : risque de piratage, de malversation, de divulgation ou de perte de données personnelles, qui sont tant de sujets essentiels rattachés aux enjeux de cybersécurité.

4.3.1. Penser conjointement *security by design* et *privacy by design*

Il existe diverses formes de « gestion » de la *privacy* dans la gouvernance des données :

- par les risques (méthode EBIOS, sécurité des systèmes, des méthodes)
- par l'éthique (règles internes, culture, responsabilisation)
- par la technologie (évaluer leurs potentialités et leurs impacts sur la vie privée –en réalisant des EIVP[40])

Ce sont toutes ces formes qu'il faut intégrer de manière transversale dans la gouvernance des données personnelles. Nous avons choisi de présenter deux méthodes qui permettent de répondre aux enjeux soulevés par les risques sur la *privacy*, qui sont la *security by design* et la *privacy by design*.

Les principes de la *security by design* sont définis dans le livre blanc « *Privacy and Security by Design : An Entreprise Architecture Approach* », coréalisé par Ann Cavoukian, Commissaire à l'information et à la protection de la vie privée de l'Ontario, et Mark Dixon, Architecte d'Entreprise chez Oracle. Les enjeux de *privacy* et de sécurité sont inséparables. Il serait inapproprié de parler de *privacy by design* si d'ores et déjà les enjeux de sécurité ne sont pas

[40] EIVP : Etude d'Impact sur la Vie Privée, cf art 33 du RGPD (Règlement Général sur la Protection des Données)

pris en compte dans l'architecture des systèmes d'information. La notion de *design* renvoie d'ailleurs à celle d'architecture. Comme le précise Ann Cavoukian, de même que « *l'architecture conventionnelle définit les règles et les normes pour la conception et la construction des bâtiments, l'architecture des systèmes d'information adresse ces mêmes questions pour la conception et la construction des ordinateurs, des réseaux de communication et les systèmes d'entreprise distribués qui sont implémentés en utilisant ces technologies* »[41]. Voici, dans le tableau ci-dessous, les 7 principes de *privacy by design* et de *security by design*, qui sont analysés en fonction de leurs enjeux respectifs :

***Privacy by Design* Foundational Principles**	**Privacy**	**Security**
	Respect and protect personal information.	Enable and protect activities and assets of both people and enterprises.
1. Proactive not Reactive; Preventative not Remedial	Anticipate and prevent privacy-invasive events *before* they happen. Do not wait for privacy risks to materialize.	Begin with the end in mind. Leverage enterprise architecture methods to guide the proactive implementation of security.
2. Default Setting	Build privacy measures directly into any given ICT system or business practice, by default.	Implement "Secure by Default" policies, including least privilege, need-to-know, least trust, mandatory access control and separation of duties.
3. Embedded into Design	Embed privacy into the design and architecture of ICT systems and business practices. Do not bolt it on after the fact.	Apply Software Security Assurance practices. Use hardware solutions such as Trusted Platform Module.
4. Positive-Sum	Accommodate all legitimate interests and objectives in a positive-sum "win-win" manner, not through a zero-sum approach involving unnecessary trade-offs.	Accommodate all stakeholders. Resolve conflicts to seek win-win.
5. End-to-End Security	Ensure cradle-to-grave, secure life-cycle management of information, end-to-end.	Ensure confidentiality, integrity and availability of all information for all stakeholders.
6. Visibility and Transparency	Keep component parts of IT systems and operations of business practices visible and transparent, to users and providers alike.	Strengthen security through open standards, well-known processes and external validation.
7. Respect for the User	Respect and protect interests of the individual, above all. Keep it user-centric.	Respect and protect the interests of all information owners. Security must accommodate both individual and enterprise interests.

Privacy by Design and Security by Design[42]

Ces dispositions ne sont pas seulement prises pour assurer la protection des *assets* des entreprises, **elles sont aussi des « facilitateurs de *business* »**. L'entreprise peut avoir des

[41] Ann Cavoukian, Mark Chanliau, Privacy and Security by Design : A convergence of Paradigms, Janvier 2013, p.3

[42] Source : Ann Cavoukian, Mark Chanliau, *Privacy and Security by Design : A convergence of Paradigms*, Janvier 2013, p.9

objectifs *business* améliorés, que ce soit avec ses partenaires commerciaux, ses fournisseurs ou ses clients, par le simple fait de garantir la mise en place de plateformes sécurisées et d'assurer le respect des trois fondamentaux de la sécurité de l'information qui sont **l'intégrité** (veiller à l'authenticité de l'information et se prémunir des risques de modification ou de destruction), la **confidentialité** (les données ne sont accessibles qu'aux personnes autorisées) et la **disponibilité** (assurant un accès rapide et fiable aux informations mises à disposition dans les SI). **Toute réflexion sur l'architecture de sécurité doit se construire à partir d'une analyse des besoins *business***. Cela permet en effet de définir plus rapidement les priorités et les actifs à protéger en fonction des objectifs *business*.

4.3.2. La *security by design*

Nous avons choisi le cas de ***Cityzen Data***, société née dans le monde de l'internet des objets, afin d'illustrer la méthode de *security by design*.

La protection des données personnelles est un sujet critique qui concerne pratiquement tous les *business model* visant l'expérience client. C'est pourquoi il est vital pour certains acteurs, de mettre rapidement le focus sur la *security by design*. ***Cityzen Data*** utilise pour répondre à cet enjeu, une technologie disruptive qu'elle a nommé « **séries spatio-temporelles »** (*Geo-Time series©*): il s'agit d'une **nouvelle architecture particulièrement adaptée à l'Internet des Objets** qui permet une **structuration des données intégrant au même niveau l'espace et le temps** : c'est une évolution conséquente des SI. Les données recueillies sont ensuite envoyées sur les serveurs de ***Cityzen Data***, puis, une fois manipulées, corrélées et analysées, sont renvoyées aux applications métiers (énergie, télécoms, transport, industrie, santé …). **Ce modèle permet de traiter des données sensibles avec une architecture qui répond techniquement aux exigences de la loi**. Il s'appuie notamment sur un langage qui techniquement protège la source de la donnée et les règles s'appliquant aux différents ayant-droits. Il **rend ainsi inopérant tout risque de réidentification comme tout risque de détournement des algorithmes**. La *security by design* est bien un prérequis essentiel à la *privacy*.

4.3.3. La *privacy by design*

La protection intégrée de la vie privée (PIVP) (ou le respect de la vie privée dès la conception) en anglais «*Privacy by Design*», (PbD) est une idée développée durant les années 1990 par la Commissaire à l'information et à la protection de la vie privée de l'Ontario (Canada) Ann CAVOUKIAN.

Partant du principe que le cadre légal ne serait pas suffisant pour assurer la protection de la sphère privée, elle a proposé d'intégrer le respect de la vie privée directement dans la conception et le fonctionnement des systèmes et réseaux informatiques, mais également dans l'élaboration de pratiques responsables.

Le respect de la vie privée dès la conception signifie prendre en compte dès le début les exigences en matière de protection de la sphère privée/protection des données et intégrer les outils de protection directement dans le produit, au lieu de les ajouter ultérieurement sous forme de compléments.

La protection intégrée de la vie privée repose sur **sept principes fondamentaux**, qui peuvent s'appliquer à toutes les catégories de données à caractère personnel[43], à toutes les mesures organisationnelles, ainsi qu'à tous les dispositifs techniques nécessaires à leur mise en œuvre. Ces derniers devront être en adéquation avec la sensibilité des données traitées.

Il s'agit de rendre à la personne concernée le contrôle sur ses données. En respectant les sept principes fondamentaux développés par Ann CAVOUKIAN, l'ensemble des entreprises qui traitent des données à caractère personnel pourront alors s'inscrire dans une **démarche responsable et durable tout en bénéficiant d'un avantage concurrentiel**. On pourrait même considérer cette démarche comme la **naissance d'une forme de responsabilité sociétale numérique des entreprises.**

Selon Alessandro FIORENTINO du Cabinet Infhotep, Ambassadeur du Privacy by Design depuis Mai 2013, le concept de *Privacy by Design* est la clé, il est bien plus qu'un article de loi ou de règlement, bien plus qu'une nouvelle norme reconnue à l'échelle mondiale. Le concept de *Privacy by Design* est un ensemble de principes et de pratiques qui régissent les rapports entre les hommes et l'information. Il s'agit d'un paradigme, une manière d'appréhender les choses, une vision cohérente du système qui repose sur une base définie, une matrice disciplinaire, un modèle théorique. C'est une **forme de rail de la conception dont les lois sont destinées aux concepteurs pour exprimer la façon dont un système d'informations *ab initio* doit être conçu et pensé dans ses grandes lignes afin d'assurer la conformité informatique et libertés**.

[43] La CNIL classe les données à caractère personnel en six catégories : Les données d'Etat-civil (identité, données d'identification), les données liées à la vie personnelle (habitudes de vie, situation familiale), les données de la vie professionnelle (CV, scolarité formation professionnelle, distinctions), les informations d'ordre économique et financier (revenus, situation financière, situation fiscale), les données de connexion (adress IP, logs) et les données de localisation (déplacements, données GPS, GSM).

4.3.3.1. LES SEPT PRINCIPES FONDAMENTAUX

Ci-dessous, voici les sept principes fondamentaux publiés en aout 2009[44], ces principes sont suivis de la lecture personnelle de chacun d'Alessandro FIORENTINO[45] :

1. **Prendre des mesures proactives et non réactives, des mesures préventives et non correctives**

La protection intégrée de la vie privée (PIVP) se caractérise par des mesures proactives et non réactives. Elle consiste à prévoir et à prévenir les incidents d'atteinte à la vie privée avant qu'ils ne se produisent. En effet, la PIVP n'attend pas que des risques pour la vie privée se concrétisent, et elle ne propose aucune solution pour résoudre les cas d'atteinte à la vie privée qui se sont déjà produits. Elle vise plutôt à les prévenir. Bref, la protection intégrée de la vie privée vient avant et non après de tels incidents[46].

Ce principe souligne que le concept de *Privacy by Design* n'offre aucune solution corrective en cas d'atteinte à la vie privée, sa mise en œuvre doit intervenir dans le cycle de vie d'un projet numérique dès sa conception. Il s'agit d'anticiper les incidents d'atteinte à la vie privée avant qu'ils ne se produisent, ce premier principe place au premier plan l'importance d'agir en amont. Ce principe est une forme d'indicateur temporel, il situe dans le temps la place de la protection des données à caractère personnel. Il révèle la nécessité pour l'équipe dirigeante de l'organisme de s'engager réellement dans la prévention des atteintes à la vie privée.

2. **Assurer la protection implicite de la vie privée**

On peut être sûr d'une chose : la protection intégrée de la vie privée est implicite. Elle vise à procurer le maximum de vie privée en veillant à ce que les renseignements personnels soient systématiquement protégés au sein des systèmes informatiques ou dans le cadre des pratiques internes. Donc, la vie privée d'un particulier est protégée même si ce dernier ne pose aucun geste, car la protection de la vie privée est intégrée dans le système, implicitement.

Il s'agit d'offrir le maximum de vie privée à l'utilisateur, la protection de la vie privée n'est pas optionnelle, en effet un utilisateur doit bénéficier d'une protection maximale sans aucune intervention de sa part. La notion de « protection implicite » définit la protection des données à caractère personnel comme une convenance qui vise à protéger l'intérêt individuel de chacun, une forme de savoir-vivre ou de bienséance numérique.

Ce second principe est aujourd'hui connu sous le nom de *Privacy by Default* et repris à l'alinéa 2 de l'article 23 du futur règlement européen.

«*Le responsable du traitement met en œuvre des mécanismes visant à garantir que, par défaut, seules seront traitées les données à caractère personnel nécessaires à chaque finalité spécifique du traitement, ces données n'étant, en particulier, pas collectées ou conservées au-delà du minimum*

[44] Les sept principes fondamentaux : http://www.privacybydesign.ca/index.php/about-pbd/7-foundational-principles/

[45] Alessandro Fiorentino est Ambassadeur du Privacy by Design depuis Mai 2013 : sa lecture des sept principes est issue de sa thèse « Privacy by Design à la lumière du règlement européen » (2013) : http://www.fiorentino-consulting.com/fiorentino-privacy-by-design-isep2013.pdf

[46] Les sept principes fondamentaux en français http://www.privacybydesign.ca/content/uploads/2009/08/7foundationalprinciples-f.pdf

nécessaire à ces finalités, pour ce qui est tant de la quantité de données que de la durée de leur conservation. En particulier, ces mécanismes garantissent que, par défaut, les données à caractère personnel ne sont pas rendues accessibles à un nombre indéterminé de personnes physiques.[47]»

Il est de la responsabilité des développeurs, des chefs projets et des responsables de programmes de mettre en application ce principe.

3. Intégrer la protection de la vie privée dans la conception des systèmes et des pratiques

La protection intégrée de la vie privée, comme son nom le suggère, est intégrée dans la conception et l'architecture des systèmes informatiques et des pratiques des organismes; elle n'y est pas greffée après coup. La protection de la vie privée devient donc un élément essentiel des fonctionnalités de base. Elle fait partie intégrante du système, sans porter atteinte à ses fonctions.

Ce principe définit la protection de la vie privée comme un élément à intégrer dans la conception et l'architecture même des systèmes informatiques mais également dans les pratiques des organismes. La protection des données à caractère personnel doit être une composante intégrante du système et de la stratégie organisationnelle de l'organisme sans pour autant porter atteinte à ses fonctions principales. Pour une application optimale du second principe fondamental, il sera de la responsabilité des développeurs, des chefs projets et des responsables de programmes ainsi que du responsable de la conformité de mettre en application ce principe.

4. Assurer une fonctionnalité intégrale selon un paradigme à somme positive et non à somme nulle

La protection intégrée de la vie privée vise à tenir compte de tous les intérêts et objectifs légitimes en cause selon un paradigme à somme positive et non selon une approche périmée à somme nulle, qui nécessite des compromis inutiles. La protection intégrée de la vie privée évite ces fausses dichotomies, par exemple celle qui oppose la protection de la vie privée à la sécurité, en démontrant qu'il est vraiment possible de réaliser ces deux objectifs à la fois.

Ce principe a pour vocation d'imposer une mise en œuvre du concept de *Privacy by Design* sans porter atteinte au business, en tenant compte de tous les intérêts et objectifs légitimes de l'organisme. La protection des données ne devra pas prendre le pas sur le bon déroulement des affaires. Le concept de *Privacy by Design* n'est pas un adversaire du *business*, ils sont tous deux partenaires et complémentaires selon un paradigme à somme positive.

La protection intégrée de la vie privée doit être considérée par tous les effectifs de l'organisme et plus particulièrement par la fonction commerciale comme une valeur ajoutée inestimable,

[47] Proposition de RÈGLEMENT DU PARLEMENT EUROPÉEN ET DU CONSEIL 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 (règlement général sur la protection des données) http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_fr.pdf

elle permet de conserver le paramètre confiance des utilisateurs ou des clients, un élément majeur dans la relation client.

La dichotomie opposant la protection des données à caractère personnel à la sécurité selon un paradigme à somme nulle est une approche erronée du concept. En effet cette vision des choses est restrictive et ne peut être provoquée que par un déni total des responsabilités qui incombent au responsable de traitement.

Il est de la responsabilité des développeurs, des chefs projets, des responsables de programmes et des directeurs commerciaux d'assurer la bonne diffusion de ce principe.

5. Assurer la sécurité de bout en bout, pendant toute la période de conservation des renseignements

La protection intégrée de la vie privée, lorsqu'elle est intégrée dans le système avant que l'on ne commence à recueillir les renseignements qu'il contiendra, persiste de façon sécurisée pendant toute la période de conservation de ces renseignements; ainsi, des mesures de sécurité essentielles à la protection de la vie privée sont mises en œuvre du début jusqu'à la fin. Cela permet d'assurer la conservation sécurisée des données, puis leur destruction sécurisée à la fin de leur période de conservation. Ainsi, la protection intégrée de la vie privée assure une gestion intégrale, sécurisée et de bout en bout des renseignements pendant toute leur période de conservation.

Ce principe impose d'assurer la sécurité de l'information tout au long de son cycle de vie, on parle alors d'une gestion intégrale visant à garantir une conservation sans risque pour les données et d'en assurer la destruction à la fin de la période de conservation.

La mise en application de ce principe devra être sous la responsabilité du Correspondant Informatique et Libertés, le futur « Data Protection Officer » ainsi que de toutes les équipes de développement informatique. Le respect de ce principe permettra également d'être en conformité avec l'article 14 du futur règlement européen.

6. Assurer la visibilité et la transparence

Grâce à la protection intégrée de la vie privée, tous les intervenants seront assurés que sans égard aux pratiques ou aux technologies employées, le système fonctionne conformément aux promesses et aux objectifs établis, sous réserve d'une vérification indépendante. Les éléments et le fonctionnement du système demeurent visibles et transparents, tant pour les utilisateurs que pour les fournisseurs. La vérification permet d'établir un climat de confiance.

Le respect de la vie privée dès la conception garantit au responsable de traitement que le système fonctionne conformément aux promesses et aux objectifs établis.

Chacun des éléments intégrés aux systèmes inhérents à la protection des données à caractère personnel doit rester visible et transparent en cas de vérification indépendante, ce principe vise à conserver un haut climat de confiance. La mise en application de ce principe devra être sous la responsabilité du « Data Protection Officer » ainsi que de toutes les équipes de développement informatique, chefs projets, responsables de programmes et tous les architectes des systèmes d'information. Le DPO sera le mieux placé pour anticiper les attentes de transparence, en effet il sera lui-même amené à effectuer des audits internes, et à répondre en cas de contrôle de la CNIL.

7. **Respecter la vie privée des utilisateurs**

Avant tout, la protection intégrée de la vie privée oblige les concepteurs et utilisateurs à privilégier les intérêts des particuliers en prévoyant notamment des mesures strictes et implicites de protection de la vie privée, des exigences appropriées quant aux avis et des fonctions habilitantes et conviviales, axées sur l'utilisateur.

Ce principe impose aux concepteurs qui développent le projet et aux utilisateurs qui seront amenés à utiliser le projet une fois développé, de toujours privilégier les intérêts du particulier c'est-à-dire des personnes concernées par les données gérées via le projet ou y ayant accès via webservice. Le principe place la protection des données à caractère personnel de l'utilisateur au centre de toutes réflexions.

Ce principe n'est pas indépendant, il est entrelacé aux six autres, il impose aux concepteurs de munir les systèmes ou les produits qu'ils développent d'une certaine «adequacy[48]», conforme aux attentes des utilisateurs et aux exigences légales. Comme dans le second principe on peut y voir une forme de savoir-vivre ou de bienséance numérique qui doit être présente au sein du système, ce principe nécessite l'engagement de tous les effectifs de l'organisme

4.3.3.2. L'APPROCHE *PRIVACY BY DESIGN* DÉVELOPPÉE EN ENTREPRISE

Selon Sanaa Ait-Daoud, Responsable R&D à ***Digital & Ethics***, la méthode de PbD (*Privacy by Design*) est basée sur une **approche pragmatique**, qui consiste à **se projeter dès la phase de conception sur ce qui va se passer sur le terrain** en termes de protection des données personnelles et de la vie privée. C'est un travail continu, qui permet de **revenir sur les résultats en fonction des règles et revenir sur les règles en fonction des résultats.** Il faut mettre en place une sorte de **méthode agile pour les données personnelles**, afin d'avoir l'idée la plus précise possible des finalités. Ceci soulève un défi majeur puisque la *privacy by design* voudrait, en théorie, que l'on prenne en compte les exigences en matière de protection des données personnelles dès le début de la conception des solutions, sachant qu'en pratique, l'environnement est plutôt dynamique et les besoins peuvent évoluer. **La *privacy by design* nécessite donc une certaine agilité et capacité d'adaptation**.

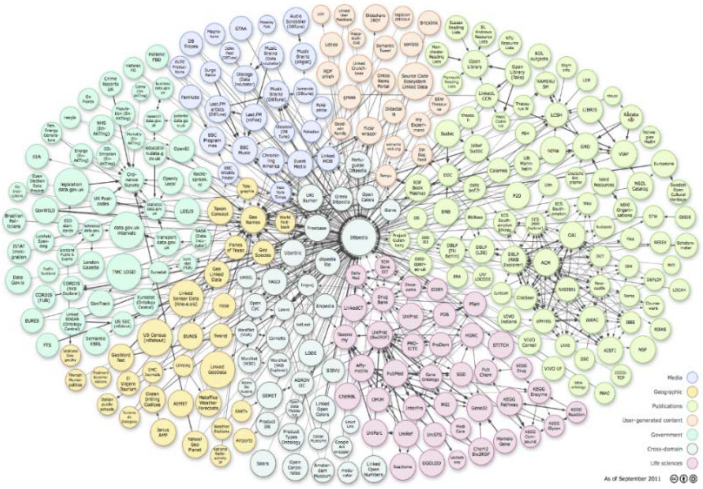
La méthodologie de la *privacy by design,* dont il est fait état dans le Règlement Général sur la Protection des Données en cours de finalisation, s'appuie en grande partie sur les recommandations de la CNIL. Cette méthodologie s'appuie notamment sur trois points :

- **L'analyse préalable** : il s'agit de questionner le système d'information ou l'application que l'on souhaite mettre en place :
    - Traite-t-on de données personnelles ?
    - L'application en crée-t-elle de nouvelles ?
    - Les réponses à ces questions vont pouvoir guider une **analyse d'impact complète sur tout le cycle de vie de la *data*.**

[48] *Adequacy* est un terme qui est le résultat de la concaténation d'«adéquate» et de « privacy», il définit le niveau de conformité adéquate aux attentes européenne et aux CNILs

- **L'évaluation des risques** : il s'agit de calculer le risque pour la vie privée en déterminant le ratio gravité / vraisemblance. C'est le principe de la méthode EBIOS.
- **L'élaboration d'un rapport EIVP** (Evaluation d'Impact sur la Vie Privée) : conduire une étude d'impact sur la vie privée permet, lorsque l'on travaille sur des projets *Big Data*, d'envisager les moyens adéquats à déployer pour prévenir les risques liés à la vie privée. Néanmoins, cette méthode a ses limites car certaines étapes de la méthodologie ne sont pas toujours applicables. En effet, les méthodologies existantes sont adaptées pour des SI « simplifiés », dont les contours sont bien délimités et pour lesquels tout est prédéfini à l'avance. Aujourd'hui ce n'est plus le cas : nous sommes passés d'une application verticale à une application plutôt horizontale de la méthode, par *process.* C'est sur ce type de modèle horizontal que sont basés par exemple les outils de *linked open data* (cf schéma ci-dessous). On s'aperçoit qu'entre 2008 et 2014, les interactions des données sur le web se sont démultipliées de manière impressionnante :

2008



2014

Source : ***Wikipedia "Linked open data" - Instance linkages within the linking open data datasets***

Nous voyons bien que l'évolution croissante des interconnexions entre les données ne permet pas d'interroger la *privacy* à chaque étape. C'est pourquoi l'un des défis primordiaux de la *privacy by design* est de **s'adapter de manière agile dans un environnement dynamique par une ré-interrogation permanente et une redéfinition des règles en fonction des contextes et objectifs**.

Même si ces méthodologies ne font pour l'instant pas partie de la culture actuelle des entreprises, elles font tout de même leur chemin. Certains acteurs n'hésitent pas à affirmer que ceux qui prendront en compte ces questions rapidement bénéficieront d'une longueur d'avance en termes de gain compétitif et d'images. **Parier sur la *privacy by design* et la *security by design*, c'est assurer une réduction des risques et donc des coûts associés.**

De plus, l'innovation peut se servir des contraintes pour développer sa créativité. Comme le précise Ann Cavoukian : «*La vision selon laquelle la vie privée étouffe l'innovation reflète un paradigme daté. L'idée selon laquelle la vie privée doit être sacrifiée pour l'innovation est une fausse dichotomie, composée de compromis inutiles. En fait, le contraire est vrai :**la vie privée stimule l'innovation. Elle oblige les innovateurs à penser de façon créative pour trouver des solutions qui serviront de multiples fonctionnalités*** »[49].

[49] Livre Blanc "Using Privacy by Design to Achieve Big Data Without Compromising Privacy", Ann Cavoukian & Deloitte Consulting, 2014

5. INNOVATIONS D'USAGES ET BONNES PRATIQUES POUR CONSTRUIRE UNE RELATION DURABLE AVEC LE CLIENT

C'est un fait, le contexte actuel autour de l'usage d'internet n'est pas favorable au développement de la confiance que porte le client à l'usage que l'on peut faire des données qui le concernent[50]. La défiance des citoyens et internautes n'a jamais été aussi forte. Et pourtant, la confiance est un élément fort de **différenciation concurrentielle**.

5.1. Le pilotage des données par le client : la théorie

Dans la même lignée du consentement éclairé du client que nous avons décrit plus haut comme première étape d'un usage contractualisé et sain des données clients, l'avenir est à la possibilité donnée aux clients de viser, contrôler et administrer par eux-mêmes les données que les entreprises détiennent sur lui.

Nombre d'initiatives visent le rétablissement de la confiance, en affichant en toute transparence ce qui est conservé lors de la relation avec tel ou tel client. Il ne s'agit pas, en tant que personne, de choisir entre nouveaux services et perte de contrôle de ses données : il est parfaitement possible de concilier les deux. Il s'agit bien en effet pour les entreprises d'avoir plus de transparence vis-à-vis des clients et de leur permettre de maîtriser l'utilisation de leurs données personnelles.

Information, consentement et autodétermination sont les trois piliers sur lesquels repose le traitement des données personnelles :

- **L'information** : il est clair que les clients, même avertis, n'ont pas l'expertise nécessaire pour protéger et contrôler leurs données. C'est à l'entreprise de trouver le meilleur moyen de communiquer avec ses clients pour leur donner une information suffisamment claire et digeste.
- **Le consentement** : bien informé, le client sait ce à quoi il consent dans le traitement et l'utilisation de ses données.
- **L'autodétermination** : c'est ce que préconise le Rapport du Conseil d'Etat de 2014 «Le numérique et les droits fondamentaux » publié en septembre dernier et qui fait 50 propositions pour mettre le numérique au service des droits des individus. La proposition numéro 1 définit un droit à l'auto-détermination des personnes sur leurs données plutôt qu'un droit de propriété : un droit à l'autodétermination, c'est-à-dire le droit de décider de la communication ou non de ses données. C'est ce que reprend le

[50] Plusieurs études — *Future of Digital Trust*, réalisée par Orange en 2014 et la 3ème vague du Baromètre *La confiance des Français dans le numérique*, réalisée par l'IDATE (Institut De l'Audiovisuel et des Télécommunications en Europe) — montrent que la confiance des clients/consommateurs est aujourd'hui en crise. D'après leurs études, plus de trois quarts des clients/ consommateurs se révèlent être méfiants vis-à-vis de la gestion et de l'utilisation de leurs données personnelles par les entreprises.

rapport du CNNum (Conseil National du Numérique) publié en Juin 2015 avec son droit à l'autodétermination informationnelle et sa proposition n°4 : « favoriser la maîtrise et l'usage de leurs données par les individus ».

Ce que l'entreprise doit proposer à son client se résume en 3 points :

- **Une information suffisante et pédagogique** dans sa formulation pour obtenir un véritable consentement des individus sur l'usage de leurs données personnelles.
- **Un suivi transparent de l'usage** fait de ses données, tout au long de leur cycle de vie.
- **Des modalités simples et faciles d'opposition** – le règlement pour la protection des données personnelles y concourt.

Plusieurs entreprises mettent en place des **ateliers pédagogiques de réalisation d'interface par le réseau**, **rendant possible ce « pilotage » par le client**. Liste est faite des données concernées sur chaque client et la possibilité d'en autoriser ou proscrire l'usage, le client peut également faire disparaitre radicalement la donnée.

Ces approches visent notamment à **bénéficier de données plus pertinentes pour une réelle performance de la relation client**. En ouvrant le champ à de nouveaux types d'échanges fondés sur une relation équilibrée entre toutes les parties prenantes, **le client contribue lui-même à l'amélioration des services rendus** et devient force de proposition dès la conception des nouveaux produits et services. La démarche VRM vient se substituer à celle du CRM.

5.2. Le VRM (*Vendor Relationship Management*) ou *Self Data* : la pratique

Le concept de VRM *(Vendor Relationship Management)* est une véritable innovation d'usage[51], théorisé par Doc Searls, co-auteur du « Manifeste des Evidences » (1999). Cet ouvrage traite du marketing à l'ère numérique, dont la première thèse, devenue célèbre, est : « *les marchés sont des conversations*. » Une conversation suppose une réciprocité d'écoutes et d'échanges. La théorie du VRM s'inscrit dans cette logique en prônant l'établissement d'une relation plus symétrique avec le client.

Le paradigme du VRM défend l'idée selon laquelle le client peut devenir lui-même *trader* de ses données, ou gérer lui-même les usages qu'il souhaite en faire (ce que la Fing appelle plus précisément le *« Self Data »)*, renversant ainsi le paradigme du CRM (*Customer Relationship Management*) dans lequel toutes les données des clients sont centralisées et gérées par l'entreprise. Plusieurs expérimentations (au Royaume-Uni avec le projet MiData, puis en France avec la Fing et le projet « MesInfos ») vont dans ce sens.

[51] L'innovation d'usage est le changement introduit dans la manière d'utiliser le produit ou service, la mise en place d'une nouvelle facilité d'usage pour répondre à des besoins du marché ou anticiper sur des futurs besoins.

En 2014, la Fing a mis en ligne les résultats issus de « MesInfos » : ont participé activement à l'expérimentation six entreprises Orange, Axa, Les Mousquetaires, La Poste, Société Générale, Crédit Coopératif) et un panel de 300 consommateurs clients dans au moins deux de ces enseignes. Les données personnelles ont été reversées dans un espace privé (créé par une start-up : *Cozy Cloud*). Il s'agissait de tickets de caisse, de factures avec les détails de communications, de données de géolocalisation, de relevés bancaires etc…

L'enjeu ensuite était de mobiliser des développeurs et designers dont le rôle était de concevoir des **applications dédiées à la gestion individuelle des données personnelles**. Les clients pouvaient choisir eux-mêmes les applications qu'ils voulaient activer ou non selon l'usage qu'ils souhaitaient faire de leurs données. La raison d'être de ce projet est de pouvoir rendre compte de **la valeur d'usage des données personnelles**, c'est-à-dire d'évaluer le bénéfice d'usage que cela apporte à l'individu.

Les deux principaux enseignements de cette expérimentation sont que :

- Tout d'abord, il y a un immense potentiel pour la création de nouveaux services : en quelques mois, la Fing a mobilisé une cinquantaine de concepteurs de services, des start-up, des développeurs etc. **Les données ne sont plus seulement dans les SI des entreprises, mais dans les mains des consommateurs** : **c'est un nouveau marché d'applications et de services qui s'ouvre**.
- Deuxièmement, **il est urgent de trouver des modèles économiques qui n'ont pas en leur cœur un conflit d'intérêt entre consommateur et entreprise.** L'expérimentation a démontré que c'est au travers de services et d'applications que nous pourrons percevoir la valeur d'usage. Et cette valeur d'usage va permettre **d'exprimer au marché les attentes de consommateurs**. La relation entre marché et consommateur n'est-elle pas de ce fait rééquilibrée ?

Pour résumer, les modèles du VRM ou du *Self Data* offrent de réelles opportunités pour le futur.

<u>Les opportunités</u>

- Améliorer la qualité des bases de données marketing : les clients mettent à jour eux-mêmes leurs données puisqu'ils en font usage ;
- Création de nouveaux services qui sont une opportunité de croissance : **type *« services personnels de gestion des données » ;***
- Recréer la confiance et la fidélité à travers le renversement du paradigme dans la relation client. L'individu est acteur de ses données, et ne pourra pas contester ce que lui-même approuve. Nous allons vers une économie plus symétrique, une économie du consentement.

Mais ces modèles supposent également de relever un certain nombre de défis face à certains risques.

Les risques

- Les individus ne seront pas prêts à exploiter ces informations et s'en remettront à des tiers, qui ne seront pas forcément plus transparents ou vertueux dans leurs pratiques que les entreprises. Le risque de désintermédiation est donc important.
- La question de la sécurisation des données est primordiale. Les données personnelles qui transitent par les applications proposées sur les plateformes ne vont-elles pas circuler sur le web ? Peut-on inventer des dispositifs pour que les données ne voyagent pas ?
- La création de services complètement nouveaux qui proposent de rémunérer les internautes en échange de leurs données personnelles (comme le font par exemple YesProfile ou Datacoup) vient concurrencer ce modèle fondé sur la valeur d'usage, pour une valeur financière.

Si la confiance constitue une préoccupation majeure de la protection des données. Elle est non seulement nécessaire pour lier des communautés de travail en réseau mais aussi indispensable aux transactions avec les clients et les autres partenaires. La confiance est également partie prenante de l'e-réputation.

Le modèle du VRM permet de comprendre en quoi l'instauration d'une relation équilibrée et de confiance entre le client et l'entreprise constitue une réelle valeur ajoutée.

CONCLUSION

Ce document a pour ambition de rappeler le contexte dans lequel peut se développer l'économie des données personnelles et de proposer les briques nécessaires à la construction d'un *business* éthique. Dans un monde de données personnelles partagées, où l'entreprise établit avec ses clients une relation symétrique, la confiance devient l'arme fatale de ce nouvel équilibre.

Alors qu'un marché est en train de se construire exclusivement autour de la valorisation des données personnelles, les grandes entreprises ont aussi leur rôle à jouer dans cette économie *via* la redéfinition de leurs modèles d'affaire. Les secteurs traditionnels sont parfois les premiers à développer des modèles disruptifs en partenariat avec des plateformes web. L'économie des données personnelles est bien une économie du partage et de la coopétition. Mais c'est aussi et surtout une économie de la confiance.

Les enjeux techniques, juridiques, éthiques et de sécurité vont de pair avec la construction de cette confiance. De nombreuses méthodologies permettent de guider le cheminement vers une bonne gouvernance, et vers des innovations durables et responsables. Bénéficier ainsi des nombreux potentiels de valorisation liés aux données personnelles est tout à fait possible au regard de la réglementation.

Enfin, les innovations d'usage comme le *Self Data* ou le VRM dessinent un nouveau paradigme, dont certaines entreprises pourront être les pionnières : celui d'une économie symétrique, gagnant-gagnant, où les consommateurs/clients deviennent les premiers partenaires des entreprises.

CIGREF

21 avenue de Messine

75008 PARIS

Tel. : +33 1 56 59 70 00

cigref@cigref.fr

www.cigref.fr