



**HAL**  
open science

## **Les nouveaux enjeux juridiques des données (big data, web sémantique et linked data)**

Thomas Saint-Aubin

### ► **To cite this version:**

Thomas Saint-Aubin. Les nouveaux enjeux juridiques des données (big data, web sémantique et linked data). *Revue Lamy Droit de l'immatériel*, 2014, n° 102 Mars 2014 (102). <hal-01523989>

**HAL Id: hal-01523989**

**<https://paris1.hal.science/hal-01523989v1>**

Submitted on 17 May 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## Les nouveaux enjeux juridiques des données (*big data*, web sémantique et *linked data*)

### Les droits de l'opérateur de données sur son patrimoine numérique informationnel

Le cadre juridique des données numériques est essentiellement constitué d'obligations : obligations déclaratives liées au traitement des données personnelles, obligations d'acheminement des données, obligations de conservation et de communication des données pour le compte des autorités judiciaires et même, parfois, obligations de mise à disposition des données publiques.

Au niveau des droits, la capacité juridique à réutiliser, à agréger et à valoriser les gisements de données détenues ou produites fait débat. *L'universalité des données détenues et produites par une personne physique ou morale constitue un patrimoine numérique, composante de son patrimoine immatériel.* Il s'agit d'un ensemble de droits et de devoirs, d'actifs et de passifs, en mutation permanente en fonction des événements juridiques de la vie numérique d'une donnée.

Pour recenser son actif numérique à un instant donné et déterminer sa capacité à valoriser les données, il faut pouvoir connaître précisément le passif, principalement constitué des droits concurrents des tiers.

Pour les données collectées sur le réseau, c'est la problématique des droits concurrents de l'ensemble des acteurs de la chaîne numérique qui se pose. La problématique des données personnelles se situe au carrefour des possibles. Les possibilités d'agrégation et la granularité admise pour les croiser avec l'ensemble des autres ressources disponibles en interne ou en externe soulèvent des questions juridiques pour lesquelles le droit ne propose pas *a priori* de réponses immédiatement applicables.

Pour définir le champ des possibles, il faut déterminer le régime juridique des données à partir du droit existant, en dégagant *un ensemble de critères de qualification*. Pour chaque donnée, le régime juridique applicable peut être synthétisé autour des droits d'usage, des droits d'exploitation et des obligations de conservation sur ce patrimoine informationnel numérique. *À partir de cette catégorisation juridique des données, il faut recenser des variables* qui déterminent les règles de changement de régime et de reconnaissance de droits des tiers.

Par Thomas SAINT-AUBIN



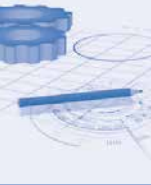
Université Paris-I  
Chercheur associé  
Laboratoire de normologie, linguistique et  
informatique juridique de l'Institut de recherche  
juridique de la Sorbonne  
Chargé d'enseignement.  
Master Droit du numérique administration-  
entreprises  
École de droit de la Sorbonne et DIU Informatique  
et droit

→ RLDI 3405

#### INTRODUCTION

Afin que le droit demeure un facteur d'innovations économiques et sociales, il est nécessaire que le juriste inscrive sa réflexion dans une démarche de prospective juridique.

Il est nécessaire d'appréhender le cadre juridique de la valorisation des données numériques sous l'angle des droits actuels et futurs dont disposent les collecteurs et les producteurs de données numériques.



Le droit est souvent considéré par l'homme de l'art comme un obstacle majeur à la faisabilité des projets. Le droit est-il compatible avec l'innovation économique et sociale ? Plus que jamais, le droit est au cœur de l'innovation : à partir de l'utilisation des données issues de l'*open data*, du web des données et du *big data*, le champ des possibles est très large.

Dans ce contexte prospectif et mouvant, en l'absence de règles immédiatement identifiables et transposables, le juriste doit faire preuve de créativité. Il ne s'agit plus seulement de confronter des usages au cadre juridique existant, il lui revient surtout de créer et d'optimiser le droit pour permettre à ces nouveaux possibles de se développer dans un cadre cohérent et dans un système de références connu et partagé de tous.

Selon Mireille Delmas-Marty, « *on peut répondre aux innovations technologiques par des innovations juridiques, mais il est très difficile de penser un droit évolutif* »<sup>(1)</sup>.

La définition de taxonomies juridiques pour opérer une catégorisation juridique des données s'inscrit ainsi dans un cadre innovant de sémantisation du droit pour redéfinir ses règles d'application dans l'environnement juridique.

La définition de métadonnées juridiques et leur rattachement à chaque donnée doivent permettre de définir un statut juridique des données.

Tout en conservant le double objectif de sémantisation et de simplification du droit, ce descriptif devra être en mesure d'accueillir des variantes juridiques importantes sur les données, en fonction de la granularité des conditions de collecte, des autorisations collectées et des croisements opérés.

Dans nos travaux de faisabilité juridique de la valorisation des données, qui se confrontent au « flou du droit » pour que celui-ci soit évolutif et innovant, nous pensons comme la présidente de la Cnil que les acteurs doivent faire preuve d'un « *surcroît de rigueur et de transparence [avec] des indicateurs sur la marge de variabilité acceptable, pour que celle-ci soit compréhensible et prédictible (...). Ainsi organisé, le droit peut être évolutif et innovant* ».

Les conclusions dégagées et les outils proposés au cours de nos recherches visent à favoriser les synergies entre le juridique, le marketing, l'informatique, les chercheurs, et globalement l'ensemble des acteurs de la chaîne numérique, pour tout ce qui relève des aspects juridiques de la valorisation des données.

### I. – STATUT JURIDIQUE ACTUEL DE LA DONNÉE

Les données ne sont pas reconnues en l'état comme des choses pouvant faire l'objet d'une appropriation. Il n'y a pas de régime juridique général de la donnée défini en Europe ou dans la loi française. Cependant, il existe en Europe un corpus de règles, en pleine évolution, sur certaines catégories de données<sup>(2)</sup>.

(1) Cnil, Cahiers IP n° 1, Vie privée à l'horizon 2020, <[http://www.cnil.fr/fileadmin/documents/Guides\\_pratiques/Livrets/Cahier-ip/cnil\\_cahieripn1/files/assets/basic-html/page57.html](http://www.cnil.fr/fileadmin/documents/Guides_pratiques/Livrets/Cahier-ip/cnil_cahieripn1/files/assets/basic-html/page57.html)>.

(2) En matière de données personnelles ou de données publiques, par exemple.

À l'analyse du droit applicable aux données numériques, on peut recenser quelques grandes règles :

– il n'y a pas de statut juridique de la propriété des données en France. En droit, la collecte des données n'emporte aucune consécration d'un droit de propriété ;

– une donnée est une information qui est de principe de libre parcours : la donnée brute rend simplement compte d'un moyen de mesure convenu : cours boursiers, données météo... ;

– le droit peut également consacrer des stratégies d'appropriation sur les données, notamment pour le producteur de base de données qui peut en interdire toute extraction substantielle.

On peut identifier d'autres droits réservés sur certaines catégories de données (secret de fabrique, *know-how*...) :

– les droits sont réservés sur les données confidentielles (secret des affaires, secret professionnel, secret de l'instruction...) ;

– enfin, il existe des droits de la personnalité sur certaines données qui relèvent de la vie privée. En Europe, la protection des données personnelles est un droit fondamental de l'individu. Il n'existe pas de « droit de propriété » des individus sur leurs données personnelles mais un simple « droit de contrôle » sur les traitements. C'est un droit personnel qui s'exerce par l'intermédiaire d'un tiers.

En résumé, nous ne relevons pas *ab initio* de droits d'exploitation pour les différents opérateurs de données : leur régime juridique est principalement constitué d'obligations (obligation de conservation, règles relatives au traitement, à la communicabilité, à l'effacement...).

### A. – Des obligations légales de conservation et de communicabilité pour le détenteur des données

À ce jour, le droit des données numériques est essentiellement abordé sous l'angle de la conservation, de la communication et de la réutilisation.

Depuis le 11 septembre 2001, on observe une prolifération de textes sur les obligations de conservation des données numériques, notamment à la charge des opérateurs.

Des droits de communication sur ces données conservées ont été progressivement consacrés au profit des pouvoirs publics, notamment au bénéfice de la justice, de la police, de l'administration fiscale, et plus récemment de la Hadopi. Ce cadre juridique évoluera encore en quantité et en qualité notamment avec l'automatisation des procédures de communication des données par les opérateurs de réseaux vers les autorités publiques.

À l'inverse, les collecteurs et les détenteurs n'ont pas connu de consécration légale d'un quelconque droit de réutilisation des données qu'ils conservent obligatoirement<sup>(3)</sup>.

Le droit a même reconnu dans certains cas des droits de réutilisation au profit des tiers, notamment vis-à-vis des informations pu-

(3) Seuls des droits à la compensation financière de l'État pour les surcoûts sont parfois consacrés : voir Arr. 22 sept. 2006, <<http://www.arcep.fr/fileadmin/reprise/textes/arretes/arr220806.pdf>>.

bliques comprises dans les documents administratifs détenues par des personnes morales de droit public ou de droit privé.



**À ce jour, le droit des données numériques est essentiellement abordé sous l'angle de la conservation, de la communication et de la réutilisation**

La confusion entre obligation de conservation et droit d'exploitation est fréquente. Les opérateurs des données peuvent légitimement s'interroger sur leur capacité juridique à réutiliser ces masses de données conservées.

## B. – Des droits de réutilisation pour le détenteur collectés *via* les contrats

Les opérateurs ont donc progressivement introduit dans leurs contrats d'abonnement des droits de réutilisation des données qu'ils détiennent, parfois même sur des données personnelles, pour des finalités explicites et pour une durée limitée.

### 1°/ Des droits individuels sur certaines catégories de données détenues

Le droit à la protection des données personnelles est un droit fondamental, prévu dans la Charte des droits fondamentaux de l'Union européenne. La simple mémorisation des données relatives à la vie privée d'un individu constitue une ingérence au sens de l'article 8 de la Convention européenne des droits de l'Homme (droit au respect de la vie privée).

C'est un droit personnel dont disposent les individus sur les données qui les concernent. Mais en l'absence de la détention matérielle de la chose, il ne dispose pas d'un droit réel susceptible de s'exercer directement sur les données. Ce droit personnel sur ses données personnelles s'exerce donc par l'intermédiaire d'un tiers.

Autrement dit, il s'agit surtout de s'intéresser aux droits d'usage dont disposent les parties sur cette catégorie de données plutôt que d'en rechercher un hypothétique propriétaire : c'est davantage un enjeu de souveraineté sur la donnée entre la personne concernée, qui dispose d'un droit personnel, et un détenteur, qui souhaiterait pouvoir disposer d'un droit réel pour l'exploiter.

Les détenteurs disposent donc de données qu'ils ne peuvent pas toujours réutiliser, à l'inverse, les particuliers ne disposent pas matériellement de leurs données qu'ils devraient en principe pouvoir utiliser.

### 2°/ Des droits collectifs sur les données publiques

Depuis la directive européenne n° 2003/98/CE du 17 novembre 2003 instaurant un régime juridique prévoyant la réutilisation des informations publiques et sa transposition dans les différentes législations nationales, le droit consacre progressivement un droit à la réutilisation des données publiques.

La construction de ce cadre légal alimente très largement le mouvement de l'ouverture des données. L'*open data*, tel qu'il est pratiqué par un nombre croissant d'acteurs publics et de façon encore marginale par des entreprises, conduit à un accroissement sans précédent des données externes disponibles librement. Ces en-

treprises peuvent dès à présent réutiliser ces nouveaux gisements afin de les croiser et d'enrichir leur propre patrimoine numérique.

Mais ils pourraient également en devenir prochainement les principaux contributeurs. En effet, dans certaines situations, l'ouverture de données par un opérateur<sup>(4)</sup> lui permettra de consacrer *de facto* son droit à la réutilisation, certes non exclusif, mais étanche à certaines prohibitions de principe (issues du droit de la concurrence notamment)<sup>(5)</sup>.

## II. – ÉVOLUTIONS ENVISAGEABLES DU STATUT JURIDIQUE DE LA DONNÉE

Le mouvement de l'*open data* contribue également à sacraliser le rapport juridique des entités à leurs données : entre données totalement ouvertes et données fermées, le champ d'investigation est large pour consacrer une granularité des droits d'exploitation. L'ensemble des entités publiques ou privées s'intéressent ou s'intéresseront à leurs données et à leurs prérogatives sur ce patrimoine numérique. Le droit est sollicité pour mieux protéger le patrimoine informationnel des entreprises et pour clarifier le rapport juridique de l'ensemble des acteurs numériques sur les données qui circulent.

Le *big data* pose de nouveaux défis aux juristes, notamment par la collecte de données produites directement ou indirectement par les individus (traces numériques) et donc de la possibilité de les exploiter. Ce nouveau cadre de stockage suppose de redéfinir les modalités technico-juridiques du partage des données.

### A. – Évolution du droit des données personnelles

La directive européenne du 24 octobre 1995, qui constitue le socle commun à tous les pays de l'Union européenne en matière de protection des données personnelles est en cours de révision. Un nouveau projet de règlement général sur la protection des données a été présenté le 25 janvier 2012 par la Commission<sup>(6)</sup>. Le projet, qui sera d'application directe dans les États membres, vise à mieux adapter le régime aux technologies du XXI<sup>e</sup> siècle, à mieux encadrer les transferts internationaux de données et à réduire les divergences d'application de la directive entre les États membres.

À moyen terme, certaines grandes tendances en matière de droit et de pratiques sont susceptibles de faire évoluer les concepts.

La reconnaissance et la consécration de nouveaux droits à l'édition, et à la portabilité des données<sup>(7)</sup>, vont contribuer à rapprocher le cadre juridique de la donnée de celui applicable à la propriété<sup>(8)</sup> : c'est la consécration d'un véritable droit réel de l'individu sur ses

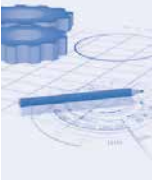
(4) Par exemple, une entreprise titulaire d'un marché public.

(5) Notamment en cas d'abus de position dominante reconnue dans l'utilisation exclusive de données.

(6) <[http://ec.europa.eu/justice/data-protection/document/review2012/com\\_2012\\_11\\_fr.pdf](http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_fr.pdf)>.

(7) Voir le projet de règlement européen.


(8) Et notamment de la propriété intellectuelle. Par exemple, le droit d'auteur aménage un régime dual entre droit moral, incessible et indissociable de la personne, et droits d'exploitation, qui peuvent faire l'objet de cession.



données, qui renforcera le degré d'appropriation. À l'image d'évolutions déjà connues par le passé pour d'autres droits<sup>(9)</sup>, au-delà de la simple reconnaissance de droits de la personnalité, les particuliers se verront reconnaître de véritables droits de réutilisation et d'exploitation sur leur patrimoine numérique.

Les usages relatifs à l'autoproduction et à l'autoédition des données personnelles, notamment en vigueur dans le web 2.0, renforcent l'exigence de protection juridique mais aussi technique pour mieux contrôler la circulation.

Avec la consécration de droits réels, nous observons donc un mouvement de patrimonialisation des données personnelles des individus.

 **Le mouvement de l'open data contribue également à sacraliser le rapport juridique des entités à leurs données**

Nous pouvons raisonner par analogie avec les personnes morales, particulièrement sensibles à la protection de leur patrimoine informationnel. Celles-ci disposent en Europe depuis 1996<sup>(10)</sup> d'outils juridiques pour garantir la rémunération de leurs investissements et l'appropriation de leurs données.

Le sentiment d'appropriation par l'entreprise des données comprises dans la base (pourtant seule concernée par la protection selon la directive) est tel que certains prestataires n'hésitent pas à annoncer à leurs clients dans leur communication qu'ils deviennent « propriétaire des données collectées » pour exploiter leur base PRM/CRM, laissant ainsi admettre de facto une appropriation par les entreprises sur les données personnelles collectées.

Afin d'aller jusqu'au bout de la logique, une partie de la doctrine propose de consacrer un véritable droit de propriété de l'individu sur ses données personnelles. À mesure que les législations reconnaissent des droits constitutifs d'une véritable identité numérique de l'individu<sup>(11)</sup>, l'humain-numérique pourrait rechercher un véritable droit de propriété sur ses données, en tant que créateur ou détenteur (stockage sur ses équipements personnels). A minima, le droit pourrait alors aligner le régime des droits patrimoniaux, dont dispose une personne sur ses données, sur celui progressivement consacré au bénéfice des entreprises sur leur patrimoine informationnel<sup>(12)</sup>.

Pourtant l'objectif du législateur européen n'est pas de consacrer un droit de propriété sur les données personnelles ; il s'agit de replacer l'individu au cœur de la régulation. Aussi séduisante soit-elle, la patrimonialisation soulève d'autres difficultés juridiques : pour Meryem Marzouki, cela reviendrait à dire « que la protection des données n'est plus un droit fondamental, ni un droit de la personnalité. Elle n'en serait pas facilitée. Un droit de propriété implique un droit d'être dépossédé, alors qu'un droit de la personnalité est inaliénable »<sup>(13)</sup>.

L'apparition de droits réels pour les individus sur leurs données ne permet donc pas de reconnaître l'avènement d'un véritable droit de propriété : s'il s'inscrit dans une démarche de patrimonialisation, il s'agit davantage d'un droit à la réutilisation, dont le législateur devra aménager les effets pour protéger les individus des risques liés à la monétisation.

Ce mouvement de patrimonialisation des données personnelles des individus entraîne parallèlement celui de la dépatrimonialisation des entreprises sur les données personnelles collectées, traitées et déclarées.

Dans ce nouveau paradigme, les entreprises, et plus globalement l'ensemble des acteurs de la chaîne numérique, doivent s'interroger sur leurs prérogatives d'exploitation sur les données numériques qu'elles détiennent. Parallèlement, elles sont invitées à mieux protéger leur patrimoine informationnel.

### B. – Consécration de nouveaux droits sur le patrimoine informationnel des entreprises

Au-delà des seules traces collectées à partir des communications électroniques, le droit s'adapte progressivement aux nouvelles exigences de l'intelligence économique et tend à consacrer des droits réservés aux entreprises sur leurs données. La protection du secret des affaires ou la notion de « patrimoine informationnel des entreprises » sont souvent exploités pour justifier ces droits exclusifs.

Aux États-Unis, le Cohen Act définit largement la notion de « secret des affaires » afin de mieux appréhender l'ensemble des informations disponibles, y compris lorsque celles-ci se présentent sous la forme immatérielle et qu'elles soient ou non stockées électroniquement. Les données sont donc expressément visées<sup>(14)</sup>.

En France, il a fallu plus de 15 ans pour envisager un outil juridique pour les entreprises dédié à la protection de leurs données par « le secret des affaires » via la notion d'« information à carac-

(9) Droits à l'image et sur l'image notamment.

(10) Consécration d'un droit *sui generis* sur les bases de données par une directive européenne, <[http://ec.europa.eu/internal\\_market/copy-right/prot-databases/index\\_fr.htm](http://ec.europa.eu/internal_market/copy-right/prot-databases/index_fr.htm)>.

(11) Voir, par exemple, la création en 2011 d'un délit d'usurpation de l'identité numérique. Article 226-4-1 du Code pénal : « Le fait d'usurper l'identité d'un tiers ou de faire usage d'une ou plusieurs données de toute nature permettant de l'identifier en vue de troubler sa tranquillité ou celle d'autrui, ou de porter atteinte à son honneur ou à sa considération, est puni d'un an d'emprisonnement et de 15 000 € d'amende. Cette infraction est punie des mêmes peines lorsqu'elle est commise sur un réseau de communication au public en ligne. »

(12) Même si le droit des producteurs de bases de données est ouvert aux personnes physiques, il a davantage été consacré pour rémuné-

rer l'investissement que pour permettre à un individu de protéger et de valoriser son patrimoine numérique.

(13) Cnil, Cahier IP n° 1, Vie privée à l'horizon 2020, précité, p. 55.

(14) <<http://www.economicespionage.com/EEA.html>>. La loi fédérale dénommée « Economic Espionage Act » ou « Cohen Act » définit le secret des affaires comme « toute forme et tout type d'information financière, commerciale, scientifique, technique, économique, industrielle, incluant modèles, plans, compilations, mécanismes, formules, dessins, prototypes, méthodes, techniques, procédés, procédures, programmes ou codes, qu'elle se présente sous forme matérielle ou immatérielle, qu'elle soit ou non stockée, compilée, ou mémorisée physiquement, électroniquement, graphiquement, ou par écrit ».

« *patrimoine économique protégée* » et le système de classification des données de la proposition de la loi « Carayon »<sup>(15)</sup>. Mais le texte a suscité de nombreuses critiques, moins sur sa légitimité que sur les procédures mises en place. Celles-ci, particulièrement lourdes pour aboutir à cette classification, semblaient réservées à des grandes entreprises qui disposent d'un savoir-faire en intelligence économique<sup>(16)</sup>. Surtout, il s'agissait d'une réponse pénale qui ne permettait pas d'aboutir à une saisie des données confidentielles qui auraient été divulguées et d'éviter la lourdeur des solutions existantes à l'international<sup>(17)</sup>.

Mais cette intention du législateur démontre bien que les entreprises, et *a fortiori* les acteurs du numérique, ont pris conscience des enjeux stratégiques liés à leur patrimoine informationnel et de la relative insuffisance des outils juridiques à leur disposition. Souhaitant désormais étendre le périmètre de la protection au-delà des solutions classiques offertes par la propriété intellectuelle, elles revendiquent une protection accrue de leurs données sensibles, dans des proportions parfois proches de celle reconnue aux individus sur leurs données personnelles. Va-t-on vers une sorte de « protection de la vie privée des entreprises » ? Dans un mouvement inverse de celui observé pour les individus, « *l'entreprise, qui exerçait un droit réel sur ses données, c'est-à-dire un droit immédiat sur les données qui étaient sur son ordinateur ou son serveur, va progressivement passer vers un droit personnel, supposé moins fort, qui s'exercera par un intermédiaire* »<sup>(18)</sup>.

On voit en tout cas que les évolutions du droit des données numériques pourraient se rejoindre et se compléter à moyen terme : les personnes morales vers davantage de protection de leur patrimoine informationnel, les particuliers vers la reconnaissance de droits patrimoniaux sur leurs données.

La notion de « *patrimoine informationnel* » tend à être reconnue en France. En effet, une proposition de loi « Carayon » a utilisé en 2012 cette notion dans son exposé des motifs<sup>(19)</sup> pour justifier l'extension du périmètre de protection du secret des affaires aux données numériques : « *La dématérialisation de l'économie rend plus diffus ce qui constitue le patrimoine d'une entreprise : ses hommes bien sûr, mais aussi leurs idées, leurs savoir-faire, leurs réseaux relationnels et commerciaux, leurs méthodes de gestion, son patrimoine informationnel, c'est-à-dire un ensemble de pratiques non brevetées, résultant de l'expérience, et testées (...). Or, l'utilisation croissante et les rapides progrès des nouvelles technologies de l'information et de la communication fragilisent ce patrimoine malgré l'amélioration des moyens de défense technique. C'est pourquoi une protection juridique adaptée à ce patrimoine*

*s'avère indispensable, l'atteinte et la révélation d'un tel patrimoine immatériel générant des conséquences dévastatrices auxquelles il convient d'apporter des réponses judiciaires adaptées* » (nous soulignons).

Mais la reconnaissance de droits privatifs sur des données numériques s'exerce régulièrement au détriment d'autres acteurs de la chaîne numérique. Le droit devra nécessairement adapter ses règles pour clarifier le rapport juridique de l'ensemble des intervenants aux données qui transitent sur les réseaux numériques.

## C. – Clarification des droits des autres acteurs numériques sur les données

En matière de communications électroniques, plusieurs acteurs sont en capacité de collecter des traces et de procéder à un enregistrement numérique.

L'ensemble des acteurs de la chaîne numérique (moteurs, navigateurs, éditeurs, CDN, régies pub, FAI...) revendiquent également des droits d'exploitation sur ces données. Par exemple, les fournisseurs de services en ligne revendiquent parfois une appropriation exclusive des données qu'ils collectent.

Pourquoi seraient-ils plus légitimes que l'opérateur de transport ou l'utilisateur lui-même à revendiquer un droit de propriété sur la donnée qui transite via son support ?

En principe ils peuvent *a minima* obtenir la possibilité de collecter des données de navigation sur leurs visiteurs en obtenant auprès d'eux le droit de stocker des cookies sur leurs disques durs. C'est une manière indirecte d'obtenir un retour sur la mise à disposition gratuite d'informations et de services en ligne et de faciliter la navigation. La collecte de données via des outils de type *toolbar* permet à leurs éditeurs de collecter des données comportementales parfois exhaustives<sup>(20)</sup>. Parfois, cela est fait de manière totalement illégale par des logiciels indésirables qui ciblent les recherches pour proposer de la publicité<sup>(21)</sup> ou même en toute légalité en se spécialisant dans ce type de services *B2B*<sup>(22)</sup>. Pour les FAI, l'intégration de clause de collecte des données de trafic dans leurs contrats d'abonnement rencontre plus de réticences, même lorsqu'elle s'accompagne de mesures *a priori* plus respectueuses de la vie privée et moins intrusives.

## III. – PARTI PRIS DOCTRINAL POUR L'ÉVOLUTION DU STATUT JURIDIQUE DE LA DONNÉE

### A. – Le rejet de la thèse de la propriété des données

Le statut juridique de la donnée transcende les concepts juridiques traditionnels comme le droit de propriété. En effet, l'une des com-

(15) <[http://www.assemblee-nationale.fr/13/dossiers/violation\\_secret\\_affaires.asp](http://www.assemblee-nationale.fr/13/dossiers/violation_secret_affaires.asp)>.

(16) <[http://www.lemonde.fr/idees/article/2012/02/27/la-loi-sur-le-secret-des-affaires-est-un-pansement-sur-une-jambe-de-bois\\_1648093\\_3232.html](http://www.lemonde.fr/idees/article/2012/02/27/la-loi-sur-le-secret-des-affaires-est-un-pansement-sur-une-jambe-de-bois_1648093_3232.html)>.

(17) Notamment en passant par la possibilité d'obtenir un accord judiciaire entre pays, prévue par le règlement européen de 2001 et par la Convention de La Haye au niveau international.

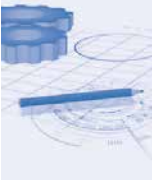
(18) Intervention de M<sup>e</sup> Dupichot lors des Journées européennes d'informatique juridique, nov. 2012, <<http://www.legalaccess.eu/?Video-des-journees-2012&lang=fr>>.

(19) <<http://www.assemblee-nationale.fr/13/propositions/pion3985.asp>>.

(20) Cas de l'application News Factory proposée par Google Chrome qui prévoit notamment de collecter « toutes les données sur tous les sites web », de « détecter la position géographique », de « lire et détecter l'historique de navigation » et d'accéder aux données que vous « copiez et collez ».

(21) Cas de l'indésirable PUPs/LPIs <[search-babylon.com](http://search-babylon.com)>.

(22) <<http://www.voxco.com/fr/pdf/collecte-donnees-multi-canaux-fiche-produit.pdf>>.



posantes essentielles de la donnée numérique est sa capacité à être clonée très rapidement. La question juridique que soulève la donnée est donc moins la question d'un droit réel sur la chose que celle d'un droit personnel sur l'usage de cette chose.

Juridiquement, les données sont des biens incorporels sur lesquels leur détenteur peut parfois se voir reconnaître des droits d'usage et d'exploitation, c'est-à-dire un droit réel. Inversement, le législateur peut également imposer certaines obligations de conservation au détenteur.

En effet, la notion de « propriété » comprend une dimension exclusive dans le rapport à la possession, ce qui caractérise mal le rapport des personnes à la donnée. La reconnaissance de droits personnels concurrents ne les prive pas de sa possession et donc de son usage. L'éventuelle appropriation définitive est sous-jacente à la notion de « propriété ». Cette exclusion des droits des tiers ne s'inscrit pas dans la même logique que la donnée, dont c'est justement la circulation et l'usage attendus qui créent la valeur.

Le droit de la propriété intellectuelle n'est pas non plus adapté car il s'inscrit dans cette même logique de reconnaissance d'un propriétaire, susceptible de céder ou de concéder des licences sur son bien mais sur lequel il conserve un lien. De plus il vise à protéger les créations intellectuelles et est attribué à un inventeur ou à un auteur par exemple, ce qui n'est pas le cas du droit de la donnée, qui cherche surtout à définir les droits d'exploitation de son détenteur ou de son producteur. Les solutions dégagées par les régimes de protection de la propriété intellectuelle peuvent néanmoins constituer des solutions pour protéger ces biens incorporels.

### B. – L'universalité des données : un patrimoine numérique

Afin de ne pas entrer dans un débat juridico-juridique sur l'opportunité de consacrer un véritable droit de propriété sur les données, pour décrire les droits et les obligations dont dispose un opérateur sur l'ensemble des données qu'il détient, nous proposons d'utiliser la notion de « patrimoine ».

Cette notion est plus souple et plus adaptée à la mutation continue des droits et des obligations de l'opérateur sur les données qui transitent sur son SI.

Ce qui est important dans le statut juridique de la donnée, c'est essentiellement le droit de la mobiliser, de l'enrichir et de la distribuer. Il s'agit davantage de définir la capacité juridique à exploiter des données plutôt que d'en rechercher un hypothétique propriétaire.

Mais avant de pouvoir valoriser ses données, l'entreprise doit pouvoir déterminer les droits dont elle dispose avant de les inscrire, dans l'idéal, à l'actif immatériel de son bilan comptable.

Pour cette raison également, le recours à la notion de « patrimoine », et plus particulièrement à celle de « patrimoine immatériel », semble plus adapté.

Juridiquement, le patrimoine peut être défini comme l'ensemble des biens qui appartiennent à une personne physique ou morale. Le patrimoine inclut les droits et actions s'y rapportant. C'est l'enveloppe (fictive) qui a vocation à recueillir les droits, les biens et

les obligations d'une personne. Le patrimoine comporte un actif et un passif.

Le patrimoine immatériel est constitué de l'ensemble des actifs immatériels d'un patrimoine. Si un opérateur est bien matériellement détenteur de ses données, pour déterminer sa capacité à les utiliser et les valoriser, il doit pouvoir évaluer ses droits et ses actions s'y rapportant.



**Ce qui est important dans le statut juridique de la donnée, c'est essentiellement le droit de la mobiliser, de l'enrichir et de la distribuer**

Sur son patrimoine immatériel de données, il dispose de droits et d'obligations : il y a bien un actif et un passif.

Le patrimoine immatériel comprend notamment le capital technologique, le capital système d'information et le capital informationnel.

Nous regroupons ces trois approches dans le patrimoine numérique, sous-classification du patrimoine immatériel, et qui comprend notamment l'ensemble des droits et des obligations d'une entité sur ses données.

Il s'agit donc d'une approche moins sacralisée que la propriété, dont l'objectif est d'énumérer les droits et des actions envisageables sur des biens.

### C. – Le patrimoine numérique : une approche tangible et concrète des droits et des obligations sur les données

La notion de « patrimoine numérique » est donc très largement une représentation à un instant T des droits et des obligations d'une entité sur ses données et leurs contenants, une copie fidèle de sa capacité à valoriser à un moment précis.

De plus, la notion de « patrimoine numérique des données » s'intègre également dans la notion plus vaste de « patrimoine informationnel » ci-dessus exposée.

Les données peuvent donc être rattachées simultanément à plusieurs patrimoines : leur régime juridique est donc fondé sur le partage réciproque des droits sur les données entre les différents acteurs de la chaîne.

Il est possible d'auditer les actifs et de définir les droits et les obligations des parties à leur endroit en analysant rigoureusement l'usus, le fructus et l'abus dont chacune dispose, en fonction du droit applicable et des contrats qui régissent leurs relations.

Pour réaliser l'audit juridique du patrimoine de données d'une entité publique ou privée, il faut donc d'abord envisager des critères de différenciation pour consolider des régimes juridiques avant d'envisager les variables pour déterminer les évolutions des actifs.

Voici quelques exemples de critères de différenciation.

#### 1°/ Distinction utilisateur/abonné à un service en ligne

Il importe de considérer le statut de la personne à l'origine de la donnée collectée pour définir les prérogatives de l'opérateur. Par

exemple, si la donnée est générée par un abonné, l'opérateur dispose alors nécessairement de certains droits (en fonction du contenu du contrat d'abonnement à ses services) et d'obligations (par exemple, en matière de conservation). En toute hypothèse, dans le croisement des données collectées et l'identification des données mobilisables, il faut distinguer les utilisateurs des abonnés. Pour ces derniers, l'opérateur disposera d'une base CRM qualifiée permettant d'aller jusqu'à la mobilisation de certaines données pour proposer des services à valeur ajoutée et des offres de partenaires pour les abonnés.

## 2° Distinction : données personnelles nominatives ou anonymisées

Les traces numériques collectées peuvent être nominatives<sup>(23)</sup> ou simplement individuelles et peuvent avoir fait l'objet d'un procédé d'anonymisation, irréversible ou non. Toutes ces variantes sont susceptibles d'avoir des conséquences juridiques. Dans le croisement des données, il faut distinguer le croisement des traces avec l'autorisation des personnes concernées, qui permet de mobiliser des données personnelles, du croisement des traces sans consentement qui nécessite *a priori*, sauf exception, d'effectuer une anonymisation préalable. *A priori*, la valorisation des données sera davantage orientée B2C pour le croisement des traces nominatives et B2B pour le croisement des traces anonymes. Mais nous pouvons envisager d'autres solutions autour du croisement de traces nominatives avec des traces anonymisées et des gisements externes de données (issues de l'*open data* ou des données personnelles des réseaux sociaux notamment). Cela peut également rejoindre les préoccupations du marketing<sup>(24)</sup>, dans le cadre de stratégies contributives sur les données CRM. Il est également important d'identifier avec précision les procédures d'anonymisation et la granularité du croisement des données afin de définir si elles mettent en œuvre ou non un traitement automatisé de données personnelles et à quel stade du croisement. En particulier, il importe de définir précisément les hypothèses dans lesquelles le recoupement de données individuelles peut aboutir à la réidentification de la personne.

## 3° Distinction : traces numériques avec obligation de conservation ou sans obligation

Depuis le 11 septembre 2001, le législateur a multiplié les obligations de conservation des données notamment envers les fournisseurs d'accès à internet. Les textes mettent à la charge de ces opérateurs des obligations de conservation des données. À l'inverse, aucun texte ne crée de droits d'exploitation au bénéfice de l'opérateur sur les données qu'il doit conserver. Il est pertinent de catégoriser précisément les données avec obligation de conservation, qui limitent l'abus de l'opérateur.

## 4° Distinction : données publiques, données privées et données personnelles

En l'état du droit ne sont pas considérées comme des informations publiques, au sens de l'article 10 de la loi n° 78-753 du

17 juillet 1978, les informations publiques produites ou reçues dans l'exercice d'un service public industriel et commercial (Spic). Par exemple, les délégations de service public de télécommunications sont des Spic<sup>(25)</sup>. Il s'agit donc bien de données privées dont l'exploitation n'est pas *a priori* envisagée dans les contrats publics et pour lesquelles la législation spécifique aux informations publiques ne s'applique pas obligatoirement, la mise à disposition de ces « données privées » en *open data* reste une opportunité pour les opérateurs concernés. Mais toutes les activités de service public de télécommunications ne sont pas exercées à titre industriel et commercial. En tant que données détenues par les collectivités territoriales ou des personnes privées chargées de la gestion d'un service public administratif (SPA), certaines données produites ou détenues sont susceptibles de répondre à la qualification juridique de données publiques. La nouvelle directive ISP tend à transformer les anciens principes en exceptions et à généraliser la « disponibilité et la réutilisation des informations du secteur public à des fins privées et commerciales ». Comme en France, la philosophie du projet se rapproche davantage d'une logique *open data*, y compris pour certaines données culturelles.

## 5° Distinction : granularité des droits collectés pour l'agrégation des données publiques, données privées et données personnelles

En France et plus généralement dans les pays qui ne pratiquent pas l'autorégulation, l'agrégation des données personnelles détenues par une entreprise sur ses clients est rarement envisagée du fait de la complexité juridique de sa mise en œuvre.

↳ Les traces numériques collectées peuvent être nominatives ou simplement individuelles et peuvent avoir fait l'objet d'un procédé d'anonymisation, irréversible ou non

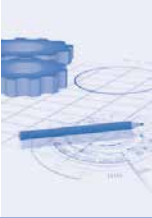
Les autres acteurs, notamment américains, ont largement envisagé des hypothèses d'agrégation des données collectées, notamment Google. Il s'agissait même de l'axe majeur de stratégie juridique de la nouvelle version de ses CGU du 1<sup>er</sup> mars 2012 (fusion de 60 CGU en une seule) qui lui permettent désormais d'agréger l'ensemble des informations disponibles pour un utilisateur. Google prévoit même d'utiliser des informations auxquelles ont accès les « amis » de l'utilisateur s'ils utilisent des applications. Les nouvelles CGU de Google ne vont toutefois pas jusqu'à envisager une revente des données agrégées à des tiers.

Si les services en ligne de gestion des données personnelles améliorent l'expérience client, ils ne constituent pas une finalité. Néanmoins, la consécration prochaine d'un système d'autorégulation

(23) C'est-à-dire qu'une personne physique est directement identifiable.

(24) Voir notamment Étude de Markess International : <<http://blog.markess.fr/2012/04/page/2/>>.

(25) D'après la jurisprudence, deux critères permettent de qualifier ce régime juridique : l'indice de l'objet du service (activités économiques comparables au secteur privé : production, distribution, services) et l'indice du mode de financement (dépenses couvertes par des recettes d'exploitation, prix payés par les usagers).



lation en Europe pour la gestion des données personnelles invite les opérateurs à s'interroger sur les modalités de mise en œuvre.

Il pourrait leur permettre de proposer de nouveaux services à leurs clients tout en leur garantissant un niveau élevé de protection de leurs données personnelles.

### CONCLUSION

Au cours de nos différents travaux, nous avons réussi à accueillir dans des taxonomies l'ensemble des critères susceptibles de définir le statut juridique d'une donnée à partir de l'approche simple et universelle des variables communes au droit des biens de l'*usus*, du *fructus* et de l'*abusus*.

Chaque donnée comprend ainsi des critères susceptibles de déterminer son statut juridique et sa capacité à être réutilisée et associée à d'autres catégories de données. Ces métadonnées constituent une synthèse du régime juridique applicable aux différentes données issues du SI.

Afin de favoriser les croisements, d'harmoniser le cadre contractuel de collecte des consentements et d'assurer une meilleure traçabilité des autorisations, nous avons recensé, tout au long de nos recherches, différents droits d'exploitations susceptibles d'être concédés sur l'ensemble des *data*.

Nous proposons donc en complément l'ébauche d'un générateur de licence dont l'objectif est de constituer un outil d'aide à la décision, de garantir juridiquement les exploitations internes et externes et, afin d'optimiser l'interopérabilité juridique des données dans le SI, d'affecter des métadonnées juridiques.

La même démarche pourrait être proposée en mettant à la disposition des clients de l'opérateur le référentiel de licences *Design Your Privacy*<sup>(26)</sup> que nous avons développé parallèlement à la rédaction de ce rapport.

Ce générateur de licences et de métadonnées juridiques tend à « sémantiser » le statut juridique des données afin de favoriser leur interopérabilité juridique. Il doit permettre de définir les droits d'accès, de conservation, de suppression, de diffusion et de la réutilisation des données.

Pour industrialiser les processus et initier de nouveaux services autour des gisements du *big data* et de l'*open data*, il faudra maintenant en interne développer l'ingénierie de croisement et « éduquer » la machine pour qu'elle apprenne à gérer l'interopérabilité juridique des données entre elles<sup>(27)</sup>.

L'objectif est de relier l'ensemble des données soumises à un statut juridique comparable afin de certifier la faisabilité de l'agrégation de données de nature différente et définir des politiques de réutilisation par catégories de données (recours aux taxonomies).

Nous pensons que l'implémentation de ces métadonnées qui tendent à garantir l'interopérabilité juridique des données au sein du SI mais aussi, plus globalement dans le *linked data* permettra de créer de nouvelles données consolidées, dynamiques, enrichies, agrégées, susceptibles d'apporter de l'intelligence dans les outils, à cette « nouvelle classe d'outils » évoquée par Tim Berners-Lee<sup>(28)</sup>.

Voir le projet de générateur de licences et de référentiel de métadonnées juridiques, <<http://www.patrimoine-immateriel.fr/>>. ■

(26) <<http://www.internetactu.net/2012/06/22/design-your-privacy-pour-une-licence-de-partage-des-donnees-personnelles/>>.

(27) Des travaux en ce sens ont été menés dans le cadre du projet ANR, <DataLift <http://datalift.org/>>.

(28) Berners-Lee T., *Semantic web roadmap*, Draft, sept. 2008, <<http://www.w3.org/DesignIssues/Semantic.html>>.