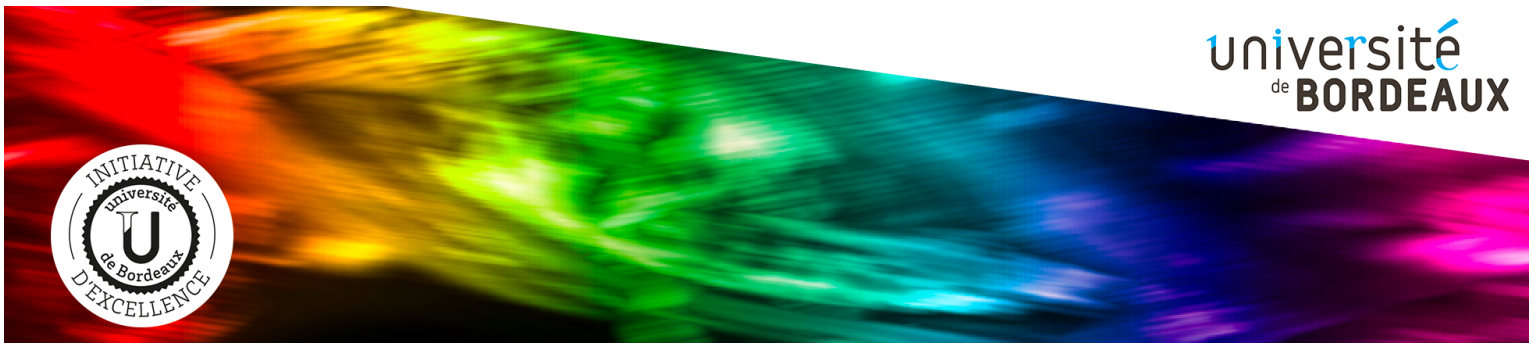
Actes du colloque des Convergences du Droit et du Numérique

François Pellegrini



Convergences du droit et du numérique

Actes du colloque

11 – 13 septembre 2017

Version 1.0 / 26 juillet 2019

Colloque organisé avec le soutien de la Fondation Anthony Mainguené.

Colloque organisé avec le soutien du GIP Mission de recherche Droit et Justice.
En association avec le débat public sur les enjeux éthiques des algorithmes lancé par la CNIL.

La Fondation Anthony Mainguené, reconnue d’utilité publique et sous égide de la Fondation de France, a été créée en 2015 afin de poursuivre les idées, actions et espérances d’Anthony Mainguené, responsable de la sécurité technique des systèmes d’information du groupe Bouygues Construction.

Dans notre société en mutation, la Fondation Anthony Mainguené, à vocation universelle, œuvre à promouvoir une réflexion éthique au moyen de colloques, conférences, modules d’enseignement et remise de prix au sein de l’enseignement supérieur, de la recherche, de la formation des futurs dirigeants et au cœur des entreprises. Elle entend ainsi créer des synergies entre les organismes et institutions vers un futur plus éthique.

« L’Humanité est une aventure responsable et solidaire qui passe par l’autre »

Anthony Mainguené

Pourquoi des « Convergences du droit et du numérique » ?

En consacrant son étude annuelle de 2014 au thème « numérique et droits fondamentaux », le Conseil d'État avait engagé une vaste réflexion sur des questions complexes jusque-là assez inédites. Dans une société dont tous les champs sont désormais investis par le développement du numérique, il s'agit de prendre l'exacte mesure d'une mutation radicale – de type systémique – qui n'affecte pas seulement le contenu des droits fondamentaux, mais qui oblige à tout repenser : le régime juridique applicable à ces droits, afin d'en assurer la reconnaissance éventuelle lorsque des droits nouveaux émergent, et leur protection effective, lorsque les droits existants sont menacés. Mais cette mutation impose aussi de réinventer la réponse juridique elle-même qui doit pouvoir emprunter d'autres voies que celles de règles rigides, qui peuvent n'être pas suffisamment adaptées au rythme des innovations que le numérique porte et nous impose. Nous sommes en présence d'une révolution agissant sur nos modes de vie, nos comportements, notre univers de communication : le droit doit impérativement être revisité s'il entend toujours réguler les rapports sociaux.

Les pistes de réflexion qu'avait alors dégagées le Conseil d'État s'inscrivaient déjà dans la nécessité de rechercher des formes de convergences : de même que le droit devait éviter, par l'édiction de normes excessives, de faire peser sur le développement du numérique des contraintes aussi inhibitrices qu'inutiles, il était attendu du numérique qu'il se mette au service des droits individuels et de l'intérêt général. Ainsi, le Conseil d'État proposait-il, par exemple, de définir les obligations des plateformes envers les utilisateurs au regard du principe de loyauté, ou de rééquilibrer la gouvernance de l'Internet afin de mieux y faire valoir les intérêts généraux que la société entend protéger.

Le colloque « Convergences du droit et du numérique » se propose de poursuivre cette réflexion, en lui donnant un prolongement original et inédit. Non seulement la communauté des juristes et les acteurs du numérique sont invités à dialoguer ensemble, mais ils sont encouragés à le faire en se remplaçant mutuellement : les juristes parleront informatique, cependant que les informaticiens parleront de droit. Déjà, les ateliers préparatoires qui se sont tenus les 8, 9 et 10 février sur les thématiques « Algorithmes et autonomie » et « Numérique et pratique juridiques » ont rendu possible ce partage des savoirs, cette mise en commun des expériences. Il est attendu du prochain colloque qui se tiendra à Bordeaux des 11 au 13 septembre qu'il franchisse une étape supplémentaire dans le partage entre la communauté des juristes et des acteurs du numérique : à la première, je dirai « appropriez-vous le numérique pour repenser le droit », aux seconds « apprivoisez le droit pour faire avancer le numérique » ! ,

Anne Guérin

Conseiller d'État,
Présidente de la cour administrative d'appel de Bordeaux

Présentation

La « révolution numérique », provoquée par la diffusion massive des technologies numériques au sein de la société, bouleverse tant les modèles organisationnels et économiques que les catégories juridiques. De ce constat découle la nécessité de faire collaborer juristes et acteurs du numérique, en suscitant une réflexion commune sur l'évolution du droit et l'encadrement des pratiques informatiques à l'aune de la révolution numérique.

C'est dans cet objectif que l'université de Bordeaux s'est investie dans la création d'un événement dédié à favoriser cette collaboration. Pour autant, cet événement ne peut prendre une forme traditionnelle ; ce serait nier le bouleversement de la pensée et des modes d'action induits par la révolution numérique. Le format proposé est donc celui d'un événement co-construit de façon participative, destiné à créer des ponts durables entre les communautés juridique et numérique.

L'objet numérique que vous tenez entre vos mains est donc original à plusieurs égards, car issu d'un processus itératif, agile et collaboratif.

C'est sur la base des contributions reçues, sans restriction de sujet, qu'ont été dégagés les quatre thèmes des ateliers, qui font émerger les questionnements actuels sur l'informatique : les données et les traitements qui les manipulent. Les ateliers, qui se sont déroulés du 8 au 10 février 2017, ont permis de premiers échanges fructueux entre participants, débouchant sur la production de synthèses pour chaque thème, rédigées de façon collaborative et publiées dans les actes des ateliers. Ces échanges ont permis d'entamer la constitution des binômes (parfois trinômes) de travail, qui ont collaboré tout l'été en préparation des interventions au colloque des 11 au 13 septembre 2017, sachant que lors de ces interventions à plusieurs, les juristes devaient exposer sur les aspects 'informatiques, et les informaticiens sur les aspects juridiques.

Ces actes reprennent, par thème, les contributions produites par les binômes à cette occasion, ainsi que certaines interventions des tables rondes.

François Pellegrini

Professeur d'informatique, université de Bordeaux

Table des matières

Thème A : « Algorithmes et loyauté »

Thème B : « Systèmes autonomes et décision, droits fondamentaux

Table ronde : « L'encadrement juridique des traitements automatisés est-il adéquat ? »

Thème C : « Numérique et pratiques juridiques »

Table ronde : « L'évolution des pratiques juridiques »

Thème D : « Droit des données à caractère personnel »

Table ronde : « Les traitements de données personnelles à l’heure du RGPD »

Sous la présidence de Clotilde Cazamajour, avocat à la Cour, professeur ICH-CNAM, avec :

- Sarah Cadiot, avocat à la Cour
- Clémence Scottez, cheffe du service des affaires économiques à la CNIL
- Moufid Hajjar, CIL-DPO, Responsable de l’unité IAM au CHU de Bordeaux

N.B. : Chaque texte remis par les orateurs du colloque étant indépendant, la numérotation des pages de ces actes suit la logique suivante :

- lettre_de_thème ;
- numéro d’ordre dans le thème, par ordre croissant d’apparition dans le programme du colloque, commençant pour chaque thème à partir de 1 ;
- numéro de page du document, commençant pour chaque document à partir de 1.

Par exemple, la troisième page du premier document du thème A doit être référencée en tant que : « p. A-1-3 ».

Les communications du colloque auxquelles ne correspondent pas (ou pas encore) de document ne sont pas numérotées. Ceci est susceptible d’introduire des « trous » dans la numérotation au sein de chaque thème, qui pourront être comblés à mesure que les documents seront livrés. Cela pourra conduire à la production de versions successives de ces actes, à chaque fois plus complètes.

Une pagination numérique incrémentale est fournie à titre d’aide pour naviguer au sein du document, mais est susceptible d’être modifiée à chaque version.

Thème A

« Algorithmes et loyauté »

La transparence des algorithmes

Céline Teyssier[1], Philippe Roose[2], Diane Adrianirina[3]

[1] CDRE, /Université de Pau et des Pays de l'Adour

[2] LIUPPA/Université de Pau et des Pays de l'Adour

[3] Université de Bordeaux

Le terme l'algorithme nous est devenu familier. Jusqu'à lors réservé aux mathématiciens, statisticiens et informaticiens, la société de l'information et de la communication a contribué à la diffusion de ce mot sur la place publique.

Déjà utilisé sous l'antiquité, l'algorithme est un programme de calcul qui permet à partir des données d'entrée de trouver un résultat. Aujourd'hui les algorithmes servent des usages très variés. Ils permettent d'identifier le meilleur candidat lors d'un recrutement ; ils aident les entraîneurs sportifs dans le choix des tactiques de jeux ; sur les marchés financiers, ils passent des ordres d'achat et de vente ; ils vous proposent la musique que vous aimez ; ils calculent l'itinéraire le plus rapide et vous trouvent le ou la partenaire sexuel(le) idéal(e).

L'algorithme a pris une part considérable dans la société grâce à deux phénomènes : d'une part, la miniaturisation de l'électronique permettant un développement exponentiel des puissances de calcul des ordinateurs et des capacités de stockage des serveurs et d'autre part, la circulation instantanée et la collecte à grande échelle des données, démultipliées par l'utilisation des Smartphones.

Les recherches sur l'ordinateur quantique permettant de dépasser les limites de la physique laissent penser que les vitesses de calcul vont encore augmenter et ce de manière très importante. Il faut conjuguer à ceci l'explosion des données personnelles qualifiées « d'or noir du XXIème siècle ».[1] Nouvelle manne financière dont les experts s'accordent cependant, à dire qu'elle n'est rien sans une méthode de tri efficace et pertinente, c'est-à-dire rien sans un bon algorithme. Ainsi, selon la société de conseil Gartner « les données sont intrinsèquement passives. Elles ne font et ne représentent rien sauf si vous savez comment les utiliser, comment agir sur ces données, car la véritable valeur réside dans les algorithmes »[2] qui réalisent des corrélations entre elles.

[1] http://www.challenges.fr/high-tech/20140926.CHA8245/vos-donnees-personnelles-sur-internet-peuvent-valoir-de-l-or.html

[2] http://www.zdnet.fr/actualites/sans-algorithmes-le-big-data-ne-sert-a-rien-explique-le-gartner-39832842.htm

Vitesse de calcul des ordinateurs et explosion des données personnelles ont permis l'arrivée d'une nouvelle catégorie d'algorithme : les algorithmes prédictifs.

Un algorithme prédictif est un programme mathématique permettant de calculer des scores prédictifs en fonction des différents types de données disponibles sur les individus étudiés et sur leur comportement de consommation ou d'usage[3].
Les algorithmes prédictifs sont présentés par les experts comme des outils révolutionnaires et incontournables. Tous les secteurs d'activité sont concernés. Donnons quelques exemples. Les entreprises commerciales s'en servent pour ajuster au plus près l'offre, pour définir des profils de consommateurs en fonction de leurs goûts, de leurs centres d'intérêt, de leur appartenance à une communauté. Les banques et assurances les utilisent pour prédire les risques de fraude ou d'impayés. Les réseaux sociaux en font un grand usage pour mettre en relation les amis et « les amis de vos amis ». Les objets connectés dans le secteur de la santé fournissent une manne d'informations permettant tris, recoupements et ciblages en tout genre. Les services de police en font également usage pour prévoir notamment les zones de rassemblement lors de manifestations.

De cette énumération pourtant non exhaustive, il est possible de tirer l'enseignement suivant : le chiffre a pris une place immense dans notre société. Il est également possible d'identifier deux risques : une croyance inconditionnelle de la société dans le résultat mathématique ; l'opacité du fonctionnement des algorithmes. Or leurs usages peuvent conduire à influencer des comportements sans même que l'individu concerné en ait conscience ; ou peuvent conduire à des décisions discriminatoires.

Le traitement des données par les ordinateurs grâce aux algorithmes s'apparente pour un néophyte à de la magie. Or, la magie fait référence à des puissances cachées, à des croyances. Dès lors, comment accepter l'utilisation des algorithmes si on n'en comprend le fonctionnement ? Dans ce contexte, méfiance et suspicion sont de mise ; pourtant, il est certain que leur utilisation à grande échelle requiert un capital confiance. La transparence peut-elle être une réponse face aux difficultés de compréhension d'un algorithme ?

[3] https://www.definitions-marketing.com/definition/algorithme-predictif/

Aujourd'hui, **la transparence dans la société de l'information** a pris une place majeure. La quête de transparence a pris de multiples facettes et ce besoin s'est répandu dans toutes les strates de la société.

L'État et l'administration s'y soumettent. Inhérente à la démocratie, la transparence contribuerait à la confiance des citoyens dans les institutions faisant échec aux privilèges liés à l'asymétrie de l'information. Ainsi Jeremy Bentham estime que « l'œil du public rend l'homme d'État vertueux » (J. Bentham, The works of Jeremy Bentham, J. Bowring, Edimbourg, W. Taitn, 1838-1843, vol. X, p145) [5]

Dans l'entreprise, la transparence n'est pas une donnée naturelle. Opacité, secret d'affaires, secret de fabrication ont longtemps dominé le monde de l'entreprise. Le capitalisme financier et la responsabilité sociale des entreprises sont à l'origine d'une plus grande diffusion des informations financières et comptables.

Pour les citoyens, le règne de la transparence trouve son point d'orgue avec Internet. Les réseaux sociaux permettent le développement du culte de la transparence. Avec la société de l'information on assiste à un renversement de principe. Alors que traditionnellement le secret prime et la publicité nécessite une action délibérée, les technologies de l'information et de la communication inversent le principe, (comme le montre l'accroissement des logiciels open-source ainsi que l'intérêt des utilisateurs envers eux). La transparence est devenue la règle (tracer, suivre, localiser, collecter, trier, traiter les données) au point que le secret nécessite une action. L'individu lui-même est devenu transparent.

Alors que l'absence de transparence est source de suspicion, de collusion, de risque de détournement de l'intérêt général au profit d'intérêts particuliers, la recherche de transparence absolue porte a contrario les germes du totalitarisme permettant l'avènement d'une société du contrôle où tout est su sur tout le monde. C'est tout le paradoxe de la transparence. Le bon équilibre est donc la conciliation des informations qui peuvent être partagées et celles qui doivent rester secrètes.

Un algorithme doit-il être transparent ? Peut-il être transparent ? Quels en sont les enjeux ?
Pour comprendre les enjeux, il faut d'abord définir la transparence et les algorithmes.

[5] J.-F. Kerléo, « la transparence en droit », recherche sur la formation juridique d'une culture juridique, mare&martin, bibliothèque des thèses, 2015, 995p

I - Transparence et algorithmes : définir pour mieux comprendre les enjeux

A - Définition de la transparence

La transparence résulte d'un souci de moraliser les relations et de protéger les intérêts individuels . Moraliser implique la diffusion d'informations pertinentes. Or, la publication d'informations peut aller à l'encontre du secret professionnel qui entoure les algorithmes ; elle est cependant indispensable car elle permet de trouver un équilibre entre les intérêts des usagers et le propriétaire de l'algorithme. Elle permet de lutter contre l'asymétrie informationnelle, source de suspicion.

Protéger ? Dans cette société gouvernée par le chiffre, que devient le citoyen ? Quelle liberté de pensée pour le citoyen dont l'information lui parvient en fonction de ses centres d'intérêt eux-mêmes identifiés grâce aux calculs algorithmiques ? Quel libre arbitre pour le consommateur ? Quand l'internet est construit initialement comme un espace de liberté basé sur l'échange collaboratif ; l'algorithme prédictif n'est-il pas susceptible de conduire à l'enfermement de la personne ? De manière plus large ce sont les questions relatives à la protection de la vie privée et à la (nécessaire ?) anonymisation des données qui se posent.

Deux exemples illustrent parfaitement le besoin de transparence qui donne aux usagers un pouvoir de contrôle.

Le premier exemple porte sur les conclusions d'une étude Suédoise[7] ayant révélé que sur 6000 candidatures étudiées, les individus âgés de plus de 40 ans étaient exclus des procédures d'embauche ; ce qui aboutit à un processus de recrutement discriminatoire. Bien évidemment l'algorithme utilisé n'était pas public, et donc non transparent

Le second exemple concerne l'algorithme d'Admission Post-Bac (APB). En 2016, la communauté lycéenne a réclamé plus de transparence en ce qui concerne son code pour savoir sur quels critères les bacheliers étaient départagés, notamment pour les filières universitaires ayant recours au tirage au sort ou ayant une capacité d'accueil incompatible avec le nombre de demandes, et pour vérifier *in fine* si cette sélection était loyale. En effet, l'algorithme n'était, au départ, pas du tout transparent pour les utilisateurs, chose assez intrigante pour eux, étant donné l'impact que peuvent avoir ces résultats sur leurs vies. La transparence demandée passait évidemment par la publication du code de

l'algorithme mais aussi par sa spécification (c'est-à-dire la traduction et le détail du cahier des charges, des jeux et des résultats des tests utilisés) et même par une possibilité d'analyse, donc de tests de l'algorithme avec n'importe quelles données. Ces informations ont permis de mettre en lumière que certaines modalités de sélection ne respectaient pas les dispositions du code de l'éducation.

Le centre National des Ressources Textuelles et Lexicales définit la transparence comme la « propriété qu'a un corps, un milieu, de laisser passer les rayons lumineux, de laisser voir ce qui se trouve derrière ».

Au sens figuré c'est « la qualité d'une personne dont les pensées et les sentiments sont faciles à comprendre, à deviner » ; c'est aussi « la qualité d'une institution qui informe complètement sur son fonctionnement, ses pratiques » ; c'est encore « la qualité de ce qui est facilement compréhensible, intelligible ».

En droit, la transparence est associée à la notion de loyauté. Agir loyalement, c'est agir sans tricher. Et ne pas tricher consiste à ne pas dissimuler ; ce qui revient à être transparent.
La transparence est source de confiance ; l'absence de transparence source de suspicion. Selon Desmoulins, « la publicité est la sauvegarde du peuple » (maxime de Bailly reprise par Desmoulins, « le vieux cordeliers, n°7 p108) tandis que Bentham estime que « là où il n'y a pas de publicité, il n'y a pas de justice » (J. Bentham, Constitutional code, The works of Jeremy Bentham, 1843. J. Bowring, New-York, Russell & Russell, 1962, p.463, vol. IX) »[10].

Un certain nombre de textes de loi fait référence à la notion de transparence sans que l'on dispose pour autant d'une définition uniforme.

Ainsi l'article 1er de la loi n° 2006-686 du 13 juin 2006 relative à la transparence et à la sécurité nucléaire définit la transparence comme « l'ensemble des dispositions prises pour garantir le droit du public à une information fiable et accessible en matière de sécurité nucléaire ».

Ce droit à l'information est également présent dans les relations des administrations avec le public. Les administrations garantissent le droit de toute personne à l'information en ce qui concerne la liberté d'accès aux documents administratifs (art. L.300-1 code des relations entre le public et

[10] J.-F. Kerléo, « la transparence en droit », recherche sur la formation juridique d'une culture juridique, mare&martin, bibliothèque des thèses, 2015, 995p

l'administration). Or, depuis la loi n°2016-1321 du 7 octobre 2016 pour la République numérique, les codes sources sont des documents administratifs[11].

Cette même loi fait peser sur les opérateurs de plateforme une obligation de délivrer du profit du consommateur « une information loyale, claire et transparente sur :
1° Les conditions générales d'utilisation du service d'intermédiation qu'il propose et sur les modalités de référencement, de classement et de déréférencement des contenus, des biens ou des services auxquels ce service permet d'accéder […]» (art. L. 111-7 code de la consommation).

En l'absence de définition unique, il est néanmoins possible de relever que la notion de transparence en droit est liée au devoir d'information : délivrer une information claire et complète.

En informatique, c'est la notion de responsabilité qui prévaut. Un algorithme est responsable s'il respecte la confidentialité des données, la non-discrimination ET s'il se conforme à certaines règles éthiques comme la neutralité et la loyauté (*Nozha Boujemaa, DR/INRIA, Advisor to the CEO in Big Data*).

Si la transparence nécessite la diffusion d'informations claires afin de susciter la confiance ; nous verrons que le caractère protéiforme et évolutif d'un algorithme rend particulièrement complexe l'objectif de transparence des algorithmes.

B - **Définition et typologie des algorithmes**

Un algorithme est défini par la CNIL comme « la description d'une suite d'étapes permettant d'obtenir un résultat à partir d'éléments fournis en entrée ». De manière simpliste, une recette de cuisine est un algorithme permettant obtenir un plat à partir d'ingrédients.

Il est possible de classer les algorithmes en deux grandes catégories : les algorithmes déterministes et non déterministes.

[11] Article L300-2 code des relations entre le public et l'administration Modifié par LOI n°2016-1321 du 7 octobre 2016 - art. 2 « Sont considérés comme documents administratifs, au sens des titres Ier, III et IV du présent livre, quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources et décisions ».

Un *algorithme déterministe* se comporte de façon prévisible en accomplissant un processus défini au préalable par le concepteur pour résoudre un problème donné. Il donne une « valeur unique » pour n'importe quelle entrée dans son « intervalle de définition ». De ce fait on dit qu'il calcule une fonction mathématique. On l'assimilera naturellement à un automate avec un nombre fini d'états (un état étant ce qui décrit l'action de la machine à un moment donné). En effet, dans le cas déterministe, la transition entre deux états se fait de manière discrète et précise ; l'état dit courant étant entièrement déterminé par le précédent. Par ailleurs ici, l'état initial est unique. Cependant, le déterminisme peut être contraignant et restrictif, d'où la mise en œuvre d'algorithmes non-déterministes.

Un *algorithme non-déterministe* se comporte, de façon non-prévisible : il peut produire des résultats différents pour une même entrée. C'est un automate fini non déterministe : le chemin d'états que l'algorithme va suivre pour une entrée donnée n'est pas prédéfini et est juste une possibilité parmi d'autres. Il y a en effet une part plus ou moins importante d'aléa. Certains algorithmes non déterministes ne sont pas basés sur l'aléa, mais leur complexité combinatoire rend difficile voire impossible la prédiction des résultats.

À partir de ces définitions, il est possible à dire qu'un algorithme déterministe est transparent. Lorsqu'il est rendu public la connaissance des données d'entrée et de l'algorithme permet d'en déduire son résultat. S'il est privé, des tests peuvent quand même « l'éclaircir » car le caractère reproductible des résultats (les mêmes entrées produisent les même sorties) offre un certain degré de compréhension.

Concernant les algorithmes non déterministes les choses sont beaucoup plus compliquées. Pour les algorithmes Evolutifs, Génétiques, Chimiques, de type Machine Learning la structure même des algorithmes évolue, ce qui revient à dire que l'automate évolue en permanence.

On comprend que transparence et algorithme ne sont pas toujours compatibles. Pour autant peut-on se satisfaire de ce constat et accepter que certains algorithmes restent des boites noires alors même qu'ils calculent des résultats qu'il faut considérer vrais. La transparence n'est-elle pas la garantie contre l'arbitraire ?

II - Transparence et algorithmes : un mariage difficile

Ainsi, la structure de l'algorithme conditionne sa transparence rendant le degré de transparence très variable voire même nul dans certains cas (A).
En fonction du degré de transparence, il convient de s'interroger sur les garanties offertes au citoyen(B)..

A - Un niveau de transparence nécessairement lié à la structure de l'algorithme

Si l'on reprend notre classification initiale des algorithmes, nous avons les algorithmes déterministes et non-déterministes. Quand un algorithme déterministe est public, il est transparent à condition d'avoir accès à la fois au code, à l'automate et à l'algorithme lui-même. A contrario, un algorithme non déterministe et non public n'est pas transparent.
Il est alors possible d'établir un premier tableau :

	Public	Non Public
Déterministe	Transparent	
Non déterministe		Non transparent

Entre ces deux bornes, transparence acquise d'une part et transparence impossible d'autre part, une hiérarchie de transparence peut être établie entre les différents types d'algorithmes.

Pour les algorithmes non déterministes, il est nécessaire de faire plusieurs distinctions qui vont déterminer un degré d'opacité. Il faut distinguer les algorithmes évolutifs (également appelés évolutionnistes) et probabilistes.

Un algorithme évolutif ou dit évolutionniste s'inspire de la théorie de l'évolution de Darwin. En partant d'une population déterminée par avance ou en partant de données indéterminées, on définit des opérateurs de variation qui vont engendrer de nouvelles configurations. Un opérateur de sélection élimine la configuration la moins bonne. Les opérateurs de sélection sont variables. Ils peuvent être fondés sur l'élitisme (on garde le meilleur individu d'une population), sur la probabilité (on calcule un coefficient de probabilité permettant une sélection d'individus), sur le hasard ou encore sur le rang (ranking).
Le caractère évolutif de l'algorithme s'explique par la redéfinition de la population opérée à partir des opérateurs de sélection ; ce qui influent sur l'algorithme lui-même, c'est-à-dire sur son fonctionnement intrinsèque.

À cela, il convient d'ajouter que l'algorithme peut être supervisé permettant plus ou moins de contrôler les évolutions de l'algorithme lui-même afin d'éviter les dérives non souhaitées.

Concernant les algorithmes probabilistes, deux grandes familles existent :

- Las Vegas : pour une entrée donnée, le résultat attendu doit être correct ; la probabilité porte sur le temps de calcul. L'algorithme calcule une réponse correcte en un temps a priori aléatoire.
- Monte Carlo : pour une entrée donnée, le résultat peut différer et même être faux (pour une certaine probabilité)

Les deux algorithmes ont un caractère probabiliste par définition. Cependant pour le premier, il apparait clairement qu'il peut être comparé à un algorithme déterministe dans le sens où pour une donnée fournie, le résultat est toujours le même. Le second ne peut l'être car pour une même entrée, le résultat peut différer.

On obtient alors le tableau de transparence suivant :

Non déterministe/Las Vegas	Non déterministe/Monte Carlo	Non déterministe/Evolutif Supervisé	Non déterministe/Evolutif non supervisé
Assez transparent	Peut être transparent	Difficilement transparent	Non transparent

À partir de ce tableau, on comprend que la transparence dépend principalement de la typologie de l'algorithme et il apparaît clairement que le degré de transparence de l'algorithme a un impact direct sur les garanties envisageables.

B - Un niveau de garantie nécessairement lié au degré de transparence des algorithmes

Quand l'algorithme peut être transparent, il faut s'interroger sur les critères qui garantissent l'effectivité de la transparence et lorsque l'algorithme n'est pas transparent, il convient de s'interroger sur les garanties envisageables.

1. Les critères garantissant l'effectivité de la transparence

Rappelons en préalable que la recherche de transparence a une vocation protectrice des droits de la personne concernée par la mise en œuvre d'un algorithme. La transparence peut aussi servir l'intérêt général en favorisant la concurrence et l'innovation. Néanmoins, il ne faut pas occulter que

la transparence peut s'opposer aux intérêts économiques de la personne ou la société détentrice de l'algorithme.

En droit, la transparence peut s'opérer à plusieurs niveaux :

1. Etre informé de l'usage d'un algorithme ;
2. Avoir accès aux « modalités de fonctionnement ».

La première étape prend la forme d'un droit à l'information ; la deuxième étape est un droit d'accès. On comprend aisément que le second est conditionné par le premier. Étant informés de la mise en œuvre d'un algorithme, les individus concernés sont à même de demander un droit d'accès.

Être informé de la mise en œuvre d'un algorithme nécessite de définir les renseignements pertinents à transmettre. Connaître la finalité de l'algorithme est notamment un critère déterminent.

Avoir accès à l'algorithme mis en œuvre suppose d'identifier les informations pertinentes permettant d'assurer la transparence d'un algorithme.

Pour un programme informatique donné, si l'algorithme sur lequel il se base est déterministe, il possède d'un point de vue informatique toutes les caractéristiques pour être transparent. Néanmoins, cela n'est pas suffisant car il existe différents niveaux de description d'algorithme, des plus généraux aux plus détaillés. Ainsi, pour atteindre une transparence effective, la publication du code et d'une version détaillée et/ou explicite de l'algorithme est une nécessité pour « comprendre » ce que réalise et comment fonctionne le programme informatique. Un algorithme trop général ou un code trop complexe ne sont pas suffisants. En effet, la complexité peut être telle qu'une absence de détail (ou au contraire d'abstraction) rend l'algorithme (ie. le programme) incompréhensible et donc…non transparent.

Dans certains cas, l'absence de publicité de l'algorithme n'est pas un obstacle à sa compréhension. Ainsi, Qwant, le moteur de recherche Européen est un algorithme neutre et impartial, ne tenant pas compte des historiques, du profil psychologique de l'utilisateur, des annonceurs, des opinions, etc. Son code (pour une grande partie) est accessible à tous mais les algorithmes ne sont pas publics. Néanmoins la grande communauté fédérée autour de Qwant ainsi que le nombre de publications relatif à son fonctionnement donnent autant de renseignements que la publication de l'algorithme lui-même, ce qui le rend…transparent.

En résumé, le simple accès à l'algorithme n'est pas une garantie de transparence. Ainsi pour garantir une transparence effective les informations pertinentes sont l'accès au code, à l'algorithme dans sa version détaillée.

Cela correspond à la description d'une situation idéale mais comme nous l'avons vu précédemment un certain nombre d'algorithmes n'offre pas les critères de transparence. Quelles sont alors les garanties envisageables ?

2. Les garanties envisageables en l''absence de transparence

La finalité de la transparence face aux algorithmes est la protection du droit des personnes : garantie d'égalité face à un traitement algorithmique, absence de discrimination.
Or nous l'avons vu certains types d'algorithmes ne sont pas transparents.
D'une part, le caractère évolutif d'un algorithme rend impossible la compréhension du résultat fourni. Ainsi, dans le domaine du « morpho engineering », bien que l'algorithme de duplication des cellules soit connu (incluant la réplication d'ADN avec des erreurs statistiques), le nombre de possibilités est tel qu'il est impossible de prévoir les chimères résultantes. Il s'agit ici d'algorithmes évolutifs non supervisés. D'autres algorithmes liés au Big Data et relatifs à l'identification de traitements contre le cancer, intègrent des milliers de médicaments, étudient les interactions entre eux afin de proposer de nouveaux traitements contre le cancer [Marr, 2015]. Ils utilisent la plateforme d'IA Watson d'IBM utilisant le Big Data et ses algorithmes statistiques, sans pour autant en maitriser les algorithmes utilisés.
D'autre part, certains algorithmes font l'objet de brevet ; ils sont donc couverts par le secret des affaires.

Dans les cas où la transparence n'est pas possible (ou difficilement atteignable) il faut imaginer des jeux de tests permettant de vérifier le respect du droit des personnes. Cela ne rendra certes pas l'algorithme transparent mais permettra de diminuer le degré d'opacité de ce dernier, et par là même d'augmenter le degré de confiance que nous pouvons en avoir.

Conclusion : Quel avenir pour la transparence ?

La « tendance » actuelle est à l'utilisation de données massives (« Big Data »). Il s'agit de récupérer tout ce que l'on peut récupérer comme données (structurées ou non), d'y appliquer des formules mathématiques, statistiques ou algorithmes, de constater le résultat puis d'éventuellement l'exploiter. De très nombreux logiciels – tous domaines confondus - aujourd'hui utilisent ces « Big Data ». Face à la volumétrie, les algorithmes développés sont pour la plupart statistiques, à base de réseaux de neurones et/ou d'apprentissage rendant leur transparence difficile voire impossible (algorithme évolutifs supervisés ou non).

Associé à ces masses de données se pose un autre problème lié à l'anonymisation. En effet, la plupart des traces numériques que nous laissons peuvent être exploitées, décuplant les risques de discrimination.

Le défi des années à venir se situe très certainement autour de ces algorithmes et de ces données. D'ores et déjà de nombreux travaux de recherches visent à faire en sorte que les données « laissées » sur le web soient « anonymisées » afin d'éviter les recoupements et ainsi de faciliter la loyauté. Mais cela ne résout qu'un seul problème laissant la transparence s'éteindre de plus en plus.

Il semble urgent pour garantir la protection du droit des personnes et peut-être même la démocratie de réfléchir à une gouvernance des algorithmes dans laquelle le poids des pouvoirs publics soit fort afin d'éviter la main mise (si c'est encore possible) de sociétés dont les intérêts économiques sont évidents.

Il est certain que l'utilisation de données massives (« Big Data ») associée aux algorithmes évolutifs rendra aigue la question de la transparence des algorithmes.

L'algorithme a une dimension technique mais les choix techniques posent toujours des questions de sens, c'est-à-dire des questions politiques et ce serait une grave erreur de voir dans un algorithme la simple mise en œuvre d'une solution technique.

Sitographie

https://www.scriptol.fr/programmation/algorithmes-classification.php

https://www.scriptol.fr/programmation/liste-algorithmes.php

http://www.meta-media.fr/2017/01/31/algorithmes-lheure-de-la-grande-regulation.html

http://www.internetactu.net/2016/03/16/algorithmes-et-responsabilites/

http://internetactu.blog.lemonde.fr/2016/12/03/comment-rendre-les-algorithmes-responsables/

https://www.labri.fr/perso/duchon/Teaching/ENSEIRB/AlgoProba/Poly.pdf

http://ufrsciencestech.u-bourgogne.fr/master1/mi1-tc5/CM2009/Genetiques/Genetique_1.pdf

Bibliographie

S. Arpin, « vers le règne de la transparence », revue-le banquet, n°25, fev. 2008

O. Cognasse, O. james, H. Meddah, M. Moragues, J. Thoin-Bousquié « intelligence artificielle, cerveau à saisir », L'usine nouvelle, n°3517, 25 mai 2017

J.-M. Sauvé, « Culture du secret contre transparence sans limite : quel équilibre pour garantir l'intérêt général » exposé d'ouverture, colloque organisé par Transparence International France, Assemblée Nationale, 5 Jull. 2011

L. Cohen-Tanugi, « le clair-obscur d'internet », pouvoirs, n°97, 2001

G. Carcassonne, « le trouble de la transparence », pouvoirs, n°97, 2001

D. Kessler, « l'entreprise entre transparence et secret », pouvoirs, n°97, 2001

B. Marr - How Big Data Is Transforming The Fight Against Cancer - https://www.forbes.com/sites/bernardmarr/2015/06/28/how-big-data-is-transforming-the-fight-against-cancer/#3230fdd31d4f

De la preuve et de l’utilisation des Systèmes Inéquitables Numériques (SIN)

Anaïs Danet
(MCF en droit privé et sciences criminelles, Université de Reims Champagne Ardenne)

Chantal Enguehard
(LS2N, UMR CNRS 6004, Université de Nantes)

Le numérique se développe de plus en plus dans notre société, et s’immisce dans une grande partie des rapports sociaux : vie quotidienne, vie professionnelle, vie politique (citoyenneté), vie scolaire ou estudiantine, vie culturelle, vie amicale et familiale, etc., ce qui engendre une multiplication des systèmes numériques, c’est-à-dire une multiplication des systèmes fondés (au moins en partie), sur des processus dématérialisés et numériques de traitement d’informations. Ces systèmes numériques réalisent l’enregistrement d'informations, leur transformation, leur transmission ou encore leur publication.

Or, il arrive que des dysfonctionnements surviennent et que des informations soient mal enregistrées, perdues ou non accessibles. Les applications numériques peuvent également souffrir d'indisponibilité parce qu'un dispositif tombe en panne (par manque d'alimentation électrique par exemple) ou que la transmission d'informations via un réseau de télécommunication ne fonctionne pas. Nous examinons ici quelques cas ayant des implications juridiques et caractérisons ainsi les Systèmes Inéquitable Numériques.

La notion étant nouvelle, il est toutefois nécessaire de commencer par proposer une définition des systèmes inéquitables numériques. Pour ce faire, il faut sans doute revenir sur le concept d’inéquité à travers celui d’équité, avant de confronter ce concept aux systèmes numériques.

L'équité

Les juristes ont davantage tendance à définir l’équité que l’inéquité, l’équité étant d’ailleurs elle-même une notion fuyante. Du latin « *aequitas* », ayant la même racine qu’« *aequus* » qui signifie « égal », l’équité revêt un sens différent, selon que l’on parle d’équité substantielle ou d’équité processuelle. Mais qu’il s’agisse de l’une ou de l’autre, l’idée d’égalité est toujours présente. L’**équité substantielle** renvoie à la valeur qui permet au juge de corriger ou de compléter le droit lorsque les solutions résultant de son application stricte sont injustes. L’idée est que le juge doit « *peser les intérêts en présence afin de les rééquilibrer, de les égaliser au sens de l’égalité*

proportionnelle »[1]. L'**équité processuelle**, quant à elle, convoque immédiatement le concept globalisant de « procès équitable », tel que protégé par l'article 6 de la Convention européenne de sauvegarde des droits de l'Homme et des libertés fondamentales, qui renvoie schématiquement à l'idée que, pour qu'un procès soit équitable, les parties doivent être traitées sur un pied d'égalité. On remarque d'ailleurs que le respect de l'équité processuelle implique l'impartialité du juge, qui doit se trouver à « *égale distance* » de chacune des parties, ainsi que le respect de *l'égalité des armes,* et fait également écho au principe du contradictoire qui permet à toutes les parties d'avoir accès aux mêmes informations.

Système Inéquitable Numérique

Nous définissons un **système numérique** comme un assemblage de différents composants : numériques (logiciels, système d'exploitation, réseau), matériels (ordinateur, ordiphone, lecteur de carte, distributeur de tickets), auxquels il faut ajouter l'organisation humaine qui l'entoure (service de recueil des doléances, mises à jour logicielles, etc.).

Différents sujets de droit interviennent lors de l'usage d'un système numérique. *A minima* l'utilisateur d'une part et le détenteur/responsable/propriétaire du système numérique d'autre part

Il arrive que des systèmes numériques s'immiscent dans une relation juridique entre deux sujets de droit et soient utilisés notamment pour matérialiser l'exécution d'une obligation et, par suite, pour l'enregistrer. Or, l'utilisation d'un système numérique comme source de preuve de l'exécution de cette obligation peut engendrer plusieurs types de difficultés, la première étant que le système numérique est bien souvent mis en place, géré et contrôlé par une seule des deux parties en cause. Dans cette configuration, apparaît donc une inégalité entre les différents acteurs du système numérique. Voilà pourquoi il est possible de parler à leur égard de **Systèmes Inéquitables Numériques** ou **SIN**.

Voici deux exemples de systèmes numériques apparus lors de la dernière décennie et qui ont attiré notre attention du fait de la position asymétrique des acteurs (usager / société) quant à l'accès aux traces numériques et de la difficulté de l'usager à prouver sa bonne foi en cas de dysfonctionnement.

Le service de **location de vélo** de Nantes Métropole est bicloo[2] de la société Cyclocity, opératrice du groupe JCDecaux. Cette société gère des locations de vélos en libre service dans douze villes de France outre Nantes : Rouen, Créteil, Lyon, etc. En 2014, on dénombrait plus de 400 000 abonnés et une moyenne de 150 000 utilisations par jour.

---

1 L. Cadiet J. Normand, S. Amrani-Mekki, *Théorie générale du procès,* 2e édition, PUF, 2013, Coll. Thémis Droit, n°21.

2 http://www.bicloo.nantesmetropole.fr - Consulté le 3 décembre 2017

Les informations affichées sur le site web[3] indiquent : « *Lorsque vous déposez votre vélo attendez quelques minutes, un signal sonore et un voyant lumineux vous confirment que votre vélo est bien verrouillé.* ». Ces traces, sonores et lumineuses, sont fugitives au sens où l'utilisateur ne peut les conserver pour ensuite les produire et prouver ainsi sa bonne foi. Par ailleurs, lorsqu'un vélo n'est pas retourné sur une borne de location, le compte bancaire du client est débité d'une somme de 150 euros (montant du dépôt de garantie). Or, il peut arriver que le système de raccrochage fonctionne mal, ce qui aboutit à libérer un vélo, et que le dernier client l'ayant loué reçoive une pénalité de 150 euros alors même qu'il ne peut produire une preuve de sa bonne remise du vélo. En 2014, le motif le plus courant de saisine du médiateur de JCDecaux a été la demande de remboursement de cette pénalité de 150 euros[4].

La Société d'économie mixte des transports en commun de l'agglomération nantaise (SEMITAN) est l'exploitant du réseau de transport en commun de Nantes Métropole. **Des titres de transports dématérialisés** ont fait leur apparition sous deux formes. La première forme est une carte à puce nommée **LiberTan** qui peut héberger, entre autres, des tickets à l'unité donnant le droit de voyager durant une heure. Pour valider un voyage et, donc, consommer un titre de transport, l'usager passe sa carte devant une des bornes de validation disponibles dans le bus ou le tram qu'il emprunte. Lors de la validation la borne émet un son et affiche un message, mais l'usager ne dispose d'aucune trace tangible de la validation de son ticket de transport. Cette situation contraste avec la validation d'un ticket cartonné sur lequel l'usager peut constater, lui-même, sans intermédiaire logiciel ou matériel, que l'horodatage de sa validation a été inscrit sur le ticket et qu'il est lisible.

La seconde forme est une application mobile hébergée directement sur l'ordiphone de l'usager, **mTicket TAN**[5]. Elle permet, entre autres, d'acheter des tickets de transport à l'unité. L'usager consomme un ou plusieurs tickets en utilisant l'application, ce qui génère automatiquement un *Flashcode*[6]. Celui-ci est accessible sur l'ordiphone en suivant quatre liens successifs. L'usager peut également consulter un "*historique de voyage*" qui lui donne accès à "*mes derniers voyages*". La documentation consultée[7] n'a pas permis de savoir si le *Flashcode* était accompagné d'un affichage lisible des informations qui y sont inscrites. Le niveau de détail des informations présentées dans la rubrique "*mes derniers voyages*" n'y figure pas non plus.

---

3 http://www.bicloo.nantesmetropole.fr/Comment-ca-marche/Les-stations/Les-points-d-attache Consulté le 3 décembre 2017.

4 Rapport du Médiateur 2014 VLS France JCDecaux.

5 Sqli, "mTicket TAN, une application mobile innovante et inédite dans le monde du transport urbain", 6 décembre 2012. https://www.strategies-ecommerce.com/actualites/on-parle-de-nous/mticket-tan-application-mobile-transport-urbain-par-sqli-nantes - Consulté le 4 décembre 2017.

6 Un Flashcode est un format de données propriétaire de l'association française du multimédia mobile. Il est représenté par un pictogramme composé de carrés. Il ne faut pas confondre le FlashCode avec le QR code dont le format est public et qui a fait l'objet d'une norme (ISO 18004).

7 TAN, "Guide d'utilisation du service mTicket". https://www.tan.fr/fr/menu/assistance/services/mon-espace/assistance-guide-mticket-21428.kjsp - Consulté le 4 décembre 2017.

Lors d'un contrôle de la validité des titres de transport, le contrôleur procède *via* des outils numériques. Avec la carte *LiberTan*, un usager peut se voir contester la validation de son titre de transport qu'il a pourtant bien effectuée si celle-ci a été mal enregistrée, si l'enregistrement a disparu, ou encore s'il n'est pas accessible. Dans ce cas, il se voit infliger une amende : en l'absence de trace tangible, l'usager ne dispose d'aucun élément pour prouver sa bonne foi. Par ailleurs, d'autres dysfonctionnements lésant l'usager peuvent subvenir, comme la validation non souhaitée de plusieurs titres de transport. Dans le cas du *mTicket TAN*, l'usager dispose d'une trace de sa validation. Il est possible qu'il puisse constater si des titres de transport sont indûment consommés en faisant correspondre ses achats et les traces de ses derniers voyages. Toutefois il peut être victime d'un arrêt du fonctionnement de son ordiphone, par exemple par épuisement de la batterie. La documentation à ce sujet étant succincte, une étude plus poussée devra être menée pour évaluer les risques associés à l'usage du *mTicket TAN*.

Qualification de systèmes inéquitables numériques

Dans le cas du vélo en libre-service ou de la carte de transport, nous constatons que l'inéquité qui permet de qualifier ces dispositifs de SIN trouve sa source dans l'asymétrie quant à l'accès à l'information. En effet la disponibilité de la « trace » résultant de l'utilisation est différente pour l'utilisateur ou la société utilisant les serveurs d'enregistrement des informations.

C'est donc une forme d'inéquité processuelle qui peut être convoquée en ce qui concerne les systèmes inéquitables numériques, les parties n'étant pas placées sur un pied d'égalité. Les systèmes inéquitables numériques décrits plus hauts sont inéquitables précisément parce qu'ils ne placent pas les **personnes en interaction avec lui** (consommateur, usager, société) sur un pied d'égalité, en avantageant certains, le plus souvent les professionnels, au détriment des autres.

Les SIN sont par conséquent susceptibles de générer des difficultés sur le plan juridique, et plus précisément sur le plan de la preuve, lorsque leur utilisation correspond matériellement à l'exécution d'une obligation juridique. En effet, la question de la preuve de l'exécution de cette obligation ou de son inexécution est susceptible de se poser, *a fortiori* si le système est défaillant. En d'autres termes, les difficultés techniques liées à ces systèmes peuvent induire des difficultés probatoires, liées à l'accès à ces preuves et à la possibilité de lutter contre elles.

D'où la question de savoir **comment le droit de la preuve judiciaire appréhende ces systèmes inéquitables numériques.**

La question est d'autant plus intéressante que ces systèmes sont relativement récents et tendent à se développer, alors que le contentieux nourri par ces SIN est un contentieux que l'on pourrait qualifier de quasi-invisible :

En effet, il s'agit principalement de petits litiges, qui n'entraîneront pas nécessairement saisine

d'un tiers pour le régler. En outre, et quand bien même la partie lésée irait contester la fiabilité du SIN, Le contentieux sera réglé le plus souvent par la médiation. Ainsi, les quelques litiges relatifs au raccrochage des Velib à Paris ayant donné lieu à la saisine du Tribunal d'Instance de Courbevoie se sont soldés par une médiation réussie entre l'exploitant JCDecaux et les utilisateurs qui affirmaient avoir correctement raccroché leur vélo à la borne. Le recours à la médiation ne nous permet ainsi pas d'apprécier véritablement la façon dont le droit appréhende les preuves issues des SIN, pas plus que de percevoir la difficulté de les obtenir. Finalement, trop peu de décisions de justice tranchent ce type de litige pour que l'on puisse en tirer des conclusions pertinentes. A notre connaissance, un seul litige a donné lieu à une décision du Tribunal d'instance de Toulouse en faveur des utilisateurs du vélo en libre-service de la ville[8].

La quasi-invisibilité du contentieux lié à l'utilisation des SIN ne doit pourtant pas conduire à considérer que les difficultés n'existent pas, d'autant que ces difficultés sont susceptibles de toucher un public extrêmement large. Notre intention est donc en première lieu de comprendre quelles sont les difficultés créées par les systèmes inéquitables numériques dans le droit de la preuve (I) pour tenter en second lieu de proposer des pistes pour les résoudre ou, mieux, s'en prémunir (II).

I- Les SIN créateurs de difficultés dans le droit de la preuve

En premier lieu, il est nécessaire d'identifier et de comprendre les difficultés engendrées par les systèmes inéquitables numériques au regard du **droit de la preuve**. Ces difficultés sont en réalité de deux ordres. D'une part, les SIN engendrent des obstacles théoriques au droit à la preuve parce qu'ils compliquent l'accès à la preuve (A) et d'autre part, à supposer que l'on puisse y avoir accès, la confiance portée aux systèmes numériques complexifie d'autant leur remise en question par le justiciable, de telle sorte qu'il existerait une distorsion entre la valeur juridique d'une telle preuve et son impact réel sur l'issue du litige (B).

A- La difficulté d'accéder à la preuve numérique issue d'un SIN

Le droit à la preuve peut se définir comme « *le pouvoir d'exiger du juge qu'il accueille l'offre ou la demande de preuve présentant un caractère licite et pertinent* »[9]. Ainsi, lorsqu'une partie ne peut avoir accès par elle-même à un élément de preuve qu'elle sait exister, il devrait lui être possible de solliciter le concours du juge notamment pour obtenir l'ordonnance d'une mesure d'expertise permettant de mettre au jour cet élément de preuve. Il faudrait donc en théorie pouvoir rechercher cette preuve dans le système inéquitable numérique en cas de besoin. Ici des difficultés techniques apparaissent. D'abord, il est difficile de prévoir à l'avance quelle information sera nécessaire pour résoudre un litige et le système numérique n'enregistre pas nécessairement toutes

[8] Cette décision de justice est de surcroît inaccessible sur les bases de données juridiques ; seule la presse locale s'en est fait l'écho…

[9] Sur le droit à la preuve, v. A. BERGEAUD-WETTERWALD, *Le droit à la preuve*, préf. J.-C. SAINT-PAU, LGDJ, 2010, Coll. Bib. Dr. privé, spéc. n° 329.

les informations parmi la multitude qu'il pourrait avoir à enregistrer (date, heure, lieu, numéro de la carte bancaire ayant effectué le paiement, numéro du vélo, numéro du plot auquel a été rattaché le vélo, etc.). Il est donc possible que la preuve n'existe tout simplement pas. Ensuite, pour qu'une information numérique devienne une **preuve numérique**, il est nécessaire qu'elle respecte trois critères que sont l'authenticité, l'intégrité et la traçabilité[10]. En effet, l'**authenticité** garantit l'origine de l'information, l'**intégrité** garantit son contenu et la **traçabilité** garantit la façon dont cette preuve a été copiée[11]. Or, en matière de systèmes inéquitables numériques, et à supposer que les données utiles soient correctement enregistrées, il faudrait pouvoir s'assurer que celles-ci ont été correctement conservées et n'ont pas été modifiées, pour pouvoir garantir leur intégrité. Or, s'il existe des normes de conservation particulières relatives aux données personnelles, il n'est pas certain que l'ensemble des données enregistrées par le système puissent être qualifiées comme telles. En effet, s'il est vrai que les données enregistrées contiennent nécessairement des données personnelles, faute de quoi elles ne seraient pas utilisables, à défaut de pouvoir identifier l'emprunteur du vélo ou l'utilisateur des transports publics, toutes les données utiles à la solution du litige ne sont pas nécessairement des données personnelles – il en va ainsi par exemple de l'heure à laquelle le vélo a été raccroché, par exemple. Il n'est par conséquent par certain que ces données non-personnelles puissent bénéficier de la protection de ces normes de conservation. Par conséquent, il serait utile que des normes similaires s'appliquent à toutes les données enregistrées lors du service (location de vélo, validation d'un ticket de transport) afin d'être certain que le prestataire du service n'a pas pu modifier ces données. Enfin, et de façon plus pragmatique, le prestataire n'a *a priori* aucune obligation de répondre gracieusement aux demandes d'accès aux traces numériques de ses systèmes, puisqu'il ne s'agit pas à proprement parler de « données personnelles ». L'accès aux traces numériques nécessitera alors de mandater un expert, dont le coût serait très vraisemblablement porté par la personne qui sollicite cette expertise. En effet, si l'article 269 du Code de procédure civile prévoit que le juge peut mettre à la charge de l'une ou l'autre des parties ou des deux la consignation de la provision à valoir sur la rémunération de l'expert, le demandeur à la mesure d'expertise est, la plupart du temps, désigné pour consigner cette provision. En effet, l'article 271 du Code de procédure civile prévoit, qu'en principe, à défaut de consignation dans le délai imparti, la désignation de l'expert est caduque : par conséquent, pour éviter que le défendeur à la demande d'expertise ne paralyse la procédure en ne payant pas la consignation, c'est sur la tête du demandeur que reposera dans la grande majorité des cas de cette consignation, ce qui vient entraver, au moins matériellement l'accès aux preuves issues des systèmes inéquitables

[10] En ce sens, v. S. MIGAYRON, « Critères d'appréciation techniques, vraies et fausses preuves numériques », intervention à l'occasion du colloque du 13 avril 2010 à la première chambre de la Cour d'appel de Paris, consacré à « La preuve numérique à l'épreuve du litige », accessible en ligne à l'adresse http://www.lagbd.org/images/3/3a/Colloque_CNEJITA_13_Avril_2010.pdf [consulté le 8 décembre 2017], p. 19 et s.

[11] *Ibid*, spéc. p. 22 et s.

numériques, d'autant qu'il s'agit là de montants élevés puisque ces expertises sont coûteuses et que le montant de la consignation doit être aussi proche que possible de la rémunération définitive prévisible[12].

B- La difficulté de contester la preuve numérique issue des SIN

A supposer qu'une preuve soit accessible et puisse être utilisée, il faut s'interroger sur sa valeur et sur la possibilité dont pourrait disposer les parties de contester cette preuve. Or, il existe une véritable distorsion entre la valeur théorique d'une preuve et son impact réel sur l'issue d'un litige.

En effet, théoriquement, cette trace numérique correspond à un code, qui est nécessairement intelligible soit directement (si, par exemple, la trace indique clairement « retrait le 12/1/10/2016 à 16h03 »), soit indirectement parce qu'elle a été chiffrée – mais dans cette hypothèse il est toujours possible de la dé-chiffrer grâce à la clé numérique adéquate. Par conséquent, cette trace numérique peut être qualifiée d'écrit au sens de l'article 1365 du Code civil qui dispose que « *l'écrit consiste en une suite de lettres, de caractères, de chiffres ou de tous autres signes ou symboles dotés d'une signification intelligible, quel que soit leur support* ». Toutefois, au regard de la classification des modes de preuve proposée par le Code civil, il ne s'agit à l'évidence ni d'un acte authentique ni même d'un acte sous seing privé. Cet écrit devrait alors être analysé tout au plus comme un « indice » susceptible de constituer une présomption au sens de l'article 1382 du Code civil. La valeur probante de ces présomptions est alors laissée à la libre appréciation du juge et il est en principe possible de les renverser.

Pourtant, et on touche là aux confins de la psychologie, la confiance dans le numérique est extrêmement solide, à tel point qu'elle est d'ailleurs un véritable objectif politique affiché dans les lois « pour la confiance dans l'économie numérique »[13], ou encore « pour une république numérique »[14] de ces dernières années. Cette confiance est telle que la preuve numérique sera vraisemblablement jugée plus fiable qu'un témoignage et qu'il sera difficile de la contester. En réalité, pour être fiable, une trace numérique doit être signée numériquement (c'est-à-dire chiffrée) pour garantir qu'elle n'est pas volontairement modifiée, et rien ne garantit à l'heure actuelle que les systèmes inéquitables numériques signent numériquement leurs traces numériques. En outre, le système n'est, quoi qu'il arrive, pas à l'abri d'un bug qu'il pourra être extrêmement difficile de prouver.

Nous avons démontré que les systèmes inéquitables numériques sont loin d'être anodins sur le plan de la preuve. Il est donc nécessaire de chercher à proposer des solutions pour résoudre ces difficultés, ou mieux, les faire disparaître. C'est l'objet de notre seconde partie.

---

12 Art. 269 *C. proc. civ.*

13 Loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

14 Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

II- Le droit de la preuve, source limitée de solutions aux difficultés créées par les SIN

En second lieu, puisque les systèmes inéquitables numériques se développent et sont sources de difficultés, il nous appartient de chercher des solutions pour minimiser ces difficultés. Plusieurs pistes peuvent alors être proposées, qui sont plus ou moins pertinentes. L'hypothèse est la suivante : un voyageur en transport en commun doit valider son titre de transport numérique sur une borne prévue à cet effet. Toutefois, un litige survient : il affirme avoir validé correctement son titre de transport alors que le système semble *a priori* n'en avoir gardé aucune trace. Toute la difficulté réside dans la nécessité de prouver l'utilisation conforme du système et donc de prouver qu'il y a eu ou non un dysfonctionnement du système numérique. En théorie, pour prouver l'existence ou l'inexistence d'un dysfonctionnement, une expertise est possible. Mais cette solution intervenant *a posteriori* n'est que peu pertinente (A) car difficile à mettre en œuvre. Il serait par conséquent plus opportun de chercher à prévenir ces difficultés (B).

A- La résolution *a posteriori* des difficultés peu opportune

Résoudre les difficultés probatoires posées par les systèmes inéquitables numériques impliquerait de pouvoir recourir à une expertise capable de démontrer qu'un bug existe et d'en préciser la portée.

Il est alors envisageable de demander au juge d'avoir copie des traces d'usages ou de mandater un expert afin que ce dernier accède aux traces d'usage.

Or un prestataire ne peut garder trace de tous les traitements d'information effectivement réalisés car la quantité d'informations serait pléthorique. Il doit donc forcément effectuer une sélection en décidant la nature des traces conservées et la durée de leur conservation. Il donc possible qu'un bug ne soit pas perceptible par l'analyse des traces effectivement conservées par le prestataire. De plus, les traces d'usage sont elles-mêmes sujet à caution. Techniquement, pour déterminer dans quelle mesure elles reflètent l'usage véritable, il est nécessaire que l'expert évalue la sûreté de la chaîne d'horodatage et de signature numérique et qu'il détermine si des incidents sont survenus au moment des faits et qui auraient pu compromettre le fonctionnement de cette chaîne (une coupure de courant par exemple).

Par ailleurs, à supposer qu'elles puissent théoriquement aboutir, ces expertises sont extrêmement délicates, longues et complexes, ce qui les rend onéreuses et par conséquent disproportionnées au regard du montant des litiges, eu égard à la matière.

Puisqu'une résolution *a posteriori* des difficultés paraît peu opportune, il serait bien plus pertinent de réfléchir à des solutions permettant d'éviter sinon de contourner le problème.

B- La nécessaire prévention des difficultés

Pour prévenir ces difficultés probatoires liées à l'usage des systèmes inéquitables numériques, il est envisageable de renverser la charge de la preuve : en principe, si l'usage du système inéquitable matérialise l'exécution d'une obligation civile, ce devrait être à l'usager débiteur de cette obligation de prouver qu'il s'est correctement exécuté. Or, puisque la preuve est détenue par le prestataire de services, il serait sans doute envisageable de renverser la charge de la preuve pour qu'elle pèse sur celui-ci. Afin de compenser l'inégalité dans l'accès à la preuve, cette preuve devrait alors être apportée par celui qui la détient matériellement s'il veut emporter la conviction.

Cette solution ne fait toutefois que déplacer le problème : le prestataire démontrera à l'aide des traces qu'il a à sa disposition que l'utilisateur n'a pas correctement exécuté son obligation ce qui n'empêchera pas l'usager de contester ce fait en soutenant qu'il y a eu un bug dans le système, et donc en demandant une expertise.

La solution la plus pertinente serait donc de permettre à l'usager de détenir *a priori* des éléments pour prouver sa bonne foi. En d'autres termes, il faut rétablir la symétrie de l'accès à la preuve. Cette solution est d'ailleurs déjà mise en œuvre dans le cadre de certains systèmes identifiés plus haut comme pouvant être des systèmes inéquitables numériques. Ainsi, à Paris, il est possible de retirer un ticket aux bornes Vélib pour attester que le vélo a été correctement reposé, étant précisé que cette solution n'est pas infaillible non plus puisque des problèmes de panne d'imprimante, d'absence de papier ou d'encre peuvent subvenir.

Une autre possibilité équivalente mais conférant au domaine du numérique serait d'envisager l'envoi d'un message électronique ou d'un SMS ou encore la production d'un fichier signé téléchargeable (par exemple en s'inspirant des preuves d'envoi de Lettres recommandées en ligne par La Poste) que l'utilisateur pourrait conserver. Cette suggestion avait d'ailleurs été émise par le médiateur de JCDecaux mais il n'a pas été possible de savoir si elle a effectivement été mise en œuvre[15]. Toutefois, si elle était réalisée, cette procédure devrait être étudiée avec précision afin d'éviter certains écueils. Ainsi, l'envoi d'un simple message électronique serait insuffisant car il n'existe pas de garantie de délivrance des courriers électroniques : un message électronique peut ne jamais parvenir à son destinataire ou ne lui parvenir qu'après des semaines de délai.

Cette approche reste cependant prometteuse et devrait être approfondie car elle a pour conséquence d'éliminer *in fine* les SIN.

[15] Rapport du Médiateur 2014 VLS France JCDecaux, Recommandation n°4, page 43.

Conclusion

Cette première étude consacrée aux Systèmes Inéquitables Numériques permet de mettre en évidence que, malgré la confiance qu'on voudrait leur accorder aveuglément, il est nécessaire aujourd'hui d'admettre que les systèmes numériques dysfonctionnent. Ce n'est en effet qu'à cette condition qu'il est possible d'affiner l'analyse des difficultés que les systèmes numériques sont susceptibles de poser. Cette première observation se heurte toutefois aujourd'hui à l'idée reçue selon laquelle les systèmes numériques sont quasiment infaillibles. Il serait donc opportun de pouvoir effectuer des mesures quantitatives de ces dysfonctionnements afin de pouvoir mieux en appréhender l'étendue.

Ces dysfonctionnements peuvent en effet être à l'origine de difficultés probatoires, accentuées dans l'hypothèse des systèmes inéquitables numériques, qu'il convient donc de savoir identifier pour mieux tenter de résoudre les difficultés. Ainsi, il a été proposé de retenir le critère de l'asymétrie de l'accès à l'information pour qualifier le système de système *inéquitable* numérique. Cette étape d'identification des systèmes inéquitables numériques est indispensable puisqu'elle est le préalable nécessaire à la résolution des difficultés, résolution qui passe par le rétablissement de la symétrie dans l'accès à l'information numérique susceptible de devenir preuve numérique. Il s'agit donc, à terme, d'identifier les SIN pour mieux les éliminer…

Ces réflexions doivent être diffusées auprès des acteurs privés et public faisant usage de systèmes numériques dans leurs échanges avec les personnes afin de le inciter à adopter la démarche ci-dessus car le numérique ne devrait pas induire de risques supplémentaires pour les usagers.

De l'annulation d'élections par Internet par le moyen des insuffisances du système de vote

Chantal Enguehard (LS2N, UMR CNRS 6004, Université de Nantes)

Tatiana Shulga-Morskaya (CERCCLE-EA 7436, Université de Bordeaux)

« *La circonstance que des opérations électorales se déroulent par le truchement de machines à voter ne prive pas l'électeur de son droit de saisir le juge de l'élection de conclusions à fins d'annulation du scrutin, lorsque les irrégularités sont établies. Mais si le juge accueille les griefs relatifs au fonctionnement des machines à voter dans le cadre traditionnel de sa jurisprudence électorale, en assimilant la machine à voter à une urne, il n'a, jusqu'à présent, pas complètement adapté son office à la technicité particulière de ce procédé de vote et, pas conséquent, a échoué à en assurer un réel contrôle* »[1]. Ces propos du Maître Philippe Bluteau peuvent être appliqués de plein droit au vote par Internet.

Le vote par Internet est une des modalités du vote électronique. En France, comme le note Gilles Guglielmi, le vote électronique prend aujourd'hui essentiellement deux formes : celle de machines à voter ou celle du vote à distance par Internet[2]. Ce dernier est défini dans les textes officiels par les termes « *vote par correspondance électronique* » ou « *vote électronique à distance* ». L'électeur peut voter en utilisant un ordinateur connecté à internet depuis son domicile, son travail, etc.

Comme tout vote, le vote par Internet n'est pas à l'abri des irrégularités qui peuvent donner lieu à l'annulation du scrutin. En effet, des manquements à un certain nombre de principes fondamentaux du droit électoral peuvent avoir pour conséquence que la volonté réelle des électeurs « *ne peut pas être connue de manière certaine, et donc qu'il est impossible de connaître avec certitude le choix majoritaire des électeurs* »[3]. Dans ce cas, l'élection est annulée par le juge. Ces principes fondamentaux sont l'égalité, la liberté et le caractère secret du vote[4]. Le respect de ces principes vise à assurer la sincérité du scrutin. En ce sens, l'observation des élections est cruciale puisqu'elle a précisément pour but de s'assurer de la sincérité du scrutin. C'est l'observation qui

1 Gilles J. Guglielmi, Olivier Ihl. *Le vote électronique*. Issy-les-Moulineaux: LGDJ-Lextenso éd. 2015. p.207.

2 Ibid. p.121.

3 Richard Ghevontian. La notion de sincérité du scrutin. *Cahiers du Conseil constitutionnel n° 13 (Dossier : La sincérité du scrutin)* [en ligne]. 2003. [réf. 02/06/2017]. v.: http://www.conseil-constitutionnel.fr/conseil-constitutionnel/francais/nouveaux-cahiers-du-conseil/cahier-n-13/la-notion-de-sincerite-du-scrutin.52035.html.

4 Ibid.

permet de révéler au juge de l'élection des manquements aux principes mentionnés afin d'obtenir l'annulation de l'élection.

Lors d'une élection organisée à l'aide de bulletins en papier, la procédure d'identification de l'électeur, du vote et du dépouillement se déroulent en présence de personnes autorisées à contrôler les opérations électorales et d'électeurs. Dans la mesure du possible et la limite de leurs sens (vue, ouïe, etc.), ces personnes peuvent percevoir des irrégularités sans que des connaissances techniques spécialisées soient nécessaires. En revanche, avec le vote par Internet, la situation est toute autre. Une partie de la procédure (l'identification de l'électeur, l'expression de l'intention de vote et son envoi) s'effectue dans un environnement non protégé, en l'absence d'observateur. L'électeur vote depuis son domicile ou son lieu de travail (il n'y a donc pas d'assurance que l'électeur vote seul, en toute confidentialité, sans pressions). Le reste de la procédure (acheminement de l'intention de vote et décompte des voix) est réalisé par des dispositifs numériques intermédiaires puis de(s) serveur(s) de vote qui échappent à la vigilance des observateurs. Au mieux, ils peuvent observer une représentation du déroulement du scrutin sur un écran mais il n'y a pas de garantie que ce qu'ils voient correspond aux traitements des informations effectivement réalisés par les dispositifs numériques. En ce sens, la procédure d'observation électorale « *traditionnelle* » élaborée pour des élections organisées à l'aide de bulletins en papier, devient partiellement dénuée de sens en ce qui concerne les procédures dématérialisées du vote. En outre, les traitements appliqués aux intentions de vote étant informatisés[5], il est nécessaire (mais non suffisant) de disposer de connaissances en informatique, ce qui est loin d'être le cas de tous les observateurs. De ces obstacles, il résulte qu'il est difficile de contester en justice un vote par Internet en vue de l'annulation de ses résultats. Non seulement les requérants ne sont pas en capacité de détecter toutes les irrégularités et de les prouver devant le juge, mais, en outre, ce dernier, en essayant d'appliquer le raisonnement « *traditionnel* » à ce nouveau procédé, rencontre très vite la limite d'une telle assimilation.

La contestation des résultats du vote par Internet se complique encore du fait de la grande diversité des textes applicables et des juges compétents alors que ce procédé de vote reste assez peu utilisé en France. En effet, le vote par Internet est limité aux élections professionnelles et aux élections de conseillers consulaires. En 2012, ce type de vote avait été autorisé lors des élections des députés des Français établis hors de France mais il a été suspendu, par le gouvernement, pour les élections législatives de 2017. Cette décision a été fondée sur la recommandation de l'Agence

5 Les traitements informatisés appliqués aux intentions de vote sont multiples. D'abord l'électeur exprime son intention de vote en appuyant sur une touche, ou sur l'écran, pour signifier pour quel candidat il a l'intention de voter. Ce geste est converti en une impulsion électrique. Cette impulsion électrique est ensuite convertie en un codage numérique. Plusieurs transformations numériques sont ensuite appliquées sur ces données qui sont finalement agrégées.

nationale de la sécurité des systèmes d'information (Anssi), de ne pas recourir à cette modalité de vote « *en raison du contexte actuel, caractérisé par un niveau de menace extrêmement élevé de cyberattaques* »[6]. Ce moyen de vote reste cependant toujours présent dans l'article 330-13 du Code électoral, il n'est donc pas exclu qu'il soit rétabli et que la question de la contestation de ses résultats devant le juge constitutionnel, compétent pour la connaître[7], se présente à nouveau. En ce qui concerne les élections des conseillers consulaires (ainsi que les élections des représentants du personnel au sein des ministères[8]), une telle question se pose devant le juge administratif[9] alors que pour les élections professionnelles, c'est le juge judiciaire[10] qu'il faut saisir. Par ailleurs, le vote par Internet a parfois été utilisé lors de primaires de partis politiques : avant 2016 par l'UMP, le Modem et EELV ; en 2016-2017 par EELV, et (mais uniquement pour les français de l'Etranger) lors des primaires de la Belle Alliance Populaire[11] et de celles de la droite et du centre. Faute de textes officiels applicables, nous ne traitons pas le cas des élections internes aux partis politiques dans cet article.

Il est intéressant de noter qu'en dépit du caractère opaque du dispositif, et donc de la difficulté à surveiller et à contester les résultats, on peut observer de nouvelles mises en œuvre du vote par Internet[12]. Malgré les délibérations de la CNIL restant assez réservée quant à l'utilisation de ce moyen de vote, en dépit de l'impossible sécurisation du Web et de plusieurs difficultés liées au caractère nouveau du dispositif (bugs, problèmes organisationnels, possibilités de fraude massive[13]), l'usage du vote par Internet ne semble pas freinée en France. En outre, le programme politique d'Emmanuel Macron prévoit la "*généralisation du vote électronique d'ici 2022*"[14]. Comme la feuille de route adressée au ministère de l'Intérieur en septembre 2017[15] annonce l'"*interdiction des machines à voter*" qui étaient utilisées pour des élections politiques dans quelques dizaines de

6 France Diplomatie, Vote électronique, 2017, v.: http://www.diplomatie.gouv.fr/fr/services-aux-citoyens/droit-de-vote-et-elections-a-l-etranger/article/vote-electronique.

7 Article 33 de l'ordonnance n° 58-1067 du 7 novembre 1958 portant loi organique sur le Conseil constitutionnel.

8 V. par ex., CE, 1 juin 2016, n° 382233. v.: www.dalloz.fr.

9 Plus précisément, le Conseil d'État. L'article 23 du décret n° 2014-290 du 4 mars 2014 portant dispositions électorales relatives à la représentation des Français établis hors de France.

10 Le tribunal d'instance a la compétence exclusive en matière de la régularité des élections professionnelles en vertu des articles R2314-27, R2324-23, R2122-93 du Code du travail.

11 http://www.lesprimairescitoyennes.fr/francais-de-letranger/ (page consultée le 28 août 2017)

12 En l'absence d'un recensement global des élections professionnelles, il est toutefois difficile de déterminer si le vote par Internet gagne du terrain.

13 Enguehard, C. Internet Voting: situation, questions and trends. Politics and Government in the Information Age, Springer 2013.

14 Le Programme d'Emmanuel Macron. Vie politique et vie publique. 2016. v. : https://en-marche.fr/emmanuel-macron/le-programme/vie-politique-et-vie-publique.

15 Protéger, garantir et servir. La feuille de route du ministère de l'Intérieur. Septembre 2017. v. : https://media.interieur.gouv.fr/feuille-de-route/Feuille-de-route-mi.pdf

mairies, il faut s'attendre à une augmentation des usages du vote par Internet même si les élections concernées (politiques / professionnelles) ne sont pas connues. La possibilité d'un nombre croissant de contentieux électoraux en matière du vote par Internet peut donc être envisagée.

Cet article est conçu dans cette perspective, comme une tentative de rassembler les expériences et les jurisprudences éparses de contestation des résultats de votes par Internet, afin de les synthétiser et d'ébaucher un cadre possible de surveillance, de détection et de contestation d'irrégularités survenues lors de l'utilisation de ce dispositif devant le juge de l'élection. Etant donné le caractère opaque par nature du vote par Internet, la question principale est de savoir comment observer puis prouver les éventuelles irrégularités d'un scrutin par Internet.

En effet, en ce qui concerne les systèmes de vote par internet existants, seul un nombre limité d'irrégularités techniques peuventt être prouvées et par conséquent, entraîner l'annulation de l'élection en cause. Ainsi, il semble opportun d'analyser d'autres moyens d'annulation, non liés aux insuffisances ou malfonctionnements techniques du système de vote en cause mais relatifs à la non observance d'autres règles encadrant l'organisation du scrutin. En premier lieu, il convient d'établir les difficultés, pour un juge, de traiter une contestation d'un vote par Internet et des conséquences de ces difficultés, à savoir le caractère *quasi* inattaquable de ce type de vote (I). En second lieu, il est opportun de considérer les possibilités d'obtenir l'annulation de l'élection restant à disposition des requérants (II).

I. Le vote électronique (quasi) inattaquable

L'analyse de la jurisprudence relative au vote par Internet mène à un constat plutôt amer : il est difficile de faire annuler une élections par moyens liés au fonctionnement technique du système de vote. La plupart de tels moyens ont été réfutés par le juge (A) ce qui démontre la difficulté de conjuguer le fonctionnement « traditionnel » du contrôle des élections et la difficulté de prouver techniquement les irrégularités du vote en ligne (B).

A. Les moyens de preuve réfutés par le juge

Aux yeux du juge constitutionnel, l'élection doit être annulée si sa sincérité est atteinte, autrement dit, s'il n'est pas possible de révéler la volonté réelle des électeurs. Cependant, l'irrégularité donnant lieu à cette atteinte doit influencer les résultats du scrutin, eu égard l'écart de voix. Tout d'abord, il faut prouver que l'irrégularité contestée a influencé les résultats du décompte des voix, puis, que les voix influencées ont été assez ou plus nombreuses en comparaison avec

l'écart entre les candidats pour qu'il soit possible de prouver que l'élection en cause a été faussée. Ainsi, face à des contestations des résultats des élections des députés des Français de l'étranger en 2012, le Conseil Constitutionnel a considéré que les requérants n'ont pas pu prouver qu'un nombre significatif d'électeurs n'étaient pas parvenus à exprimer leur suffrage. Ni les insuffisances alléguées du dispositif de vote[16] ou dans l'organisation du scrutin[17], ni le fait que certaines recommandations de la Cnil[18] n'ont pas été mises en œuvre dans la définition des modalités de vote[19], n'ont été prises en compte dès lors qu'il n'y avait pas de preuve que ces circonstances aient affecté les résultats du scrutin.

Les élections en ligne des conseillers consulaires ont donné lieu à peu de contentieux mais il est possible de faire l'hypothèse que le juge administratif suivrait le même raisonnement que le Conseil constitutionnel. Ainsi, une erreur commise par le bureau de vote et ayant permis à un électeur de voter deux fois (à l'urne et en ligne), n'a pas constitué une atteinte à la sincérité du scrutin[20]. Il en a été de même quant à l'invitation d'une candidate adressée à dix-huit électeurs à lui envoyer leurs codes de vote « *dès lors qu'il n'est pas établi que les deux votes émis par la voie électronique par les destinataires du courriel litigieux aient été effectués sous une fausse identité* »[21].

L'approche du juge judiciaire est sensiblement identique : « *Il n'y a pas lieu d'annuler un scrutin dès qu'une irrégularité est relevée* »[22]. Dans l'arrêt du 13 janvier 2010, la Cour de cassation s'explique sur les causes d'annulation : « *à moins qu'elles soient directement contraires aux principes généraux du droit électoral, les irrégularités commises dans l'organisation et le déroulement du scrutin ne peuvent constituer une cause d'annulation que si elles ont exercé une influence sur le résultat des élections ou depuis l'entrée en vigueur de la loi n˚2008-789 du 20 août*

16 Cons. const., 2012- 4597/4626 AN, 15 février 2013, *A.N., Français établis hors de France (4ème circ.)*, v.: www.conseil-constitutionnel.fr/decision/2013/20124597_4626an.htm, cons. 6.

17 Cons. const., 2012-4554 AN, 15 février 2013, *A.N., Français établis hors de France (7ème circ.)*, v.: www.conseil-constitutionnel.fr/decision/2013/20124554an.htm, cons. 2 ; Cons. const., 2012-4627 AN, 15 février 2013, *A.N., Français établis hors de France (2ème circ.)*, v.: www.conseil-constitutionnel.fr/decision/2013/20124627an.htm, cons. 1 et 2.

18 V. Cnil, délibération n° 2012-083 du 15 mars 2012, v.: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000025826116 ; délibération n° 2010-371 du 21 octobre 2010, v.: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000023124205.

19 Cons. const., 2012- 4597/4626 AN, 15 février 2013, préc., cons. 3.

20 Conseil d'État, 17 février 2015, 380893. V. : https://www.legifrance.gouv.fr/affichJuriAdmin.do;jsessionid=5FF9F46CA63B047DC1D81E986AE97B1B.tpdila23v_2?oldAction=rechJuriAdmin&idTexte=CETATEXT000030262908&fastReqId=421180985&fastPos=5.

21 Conseil d'État, 10 décembre 2014, 380933. V.: https://www.legifrance.gouv.fr/affichJuriAdmin.do;jsessionid=5FF9F46CA63B047DC1D81E986AE97B1B.tpdila23v_2?oldAction=rechJuriAdmin&idTexte=CETATEXT000029882514&fastReqId=623311466&fastPos=12.

22 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou. *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise*. Paris: Dalloz. 2015. p.1247.

2008 si, s'agissant du premier tour, elles ont été déterminantes de la qualité représentative des organisations syndicales dans l'entreprise, ou du droit pour un candidat d'être désigné délégué syndical »[23]. En ce qui concerne le fonctionnement du système de vote par Internet, c'est donc sur le fondement de ce raisonnement que le juge prend la décision d'annulation. Cependant, l'approche du juge judiciaire est beaucoup plus attentive en ce qui concerne l'aspect technique de l'affaire. Ainsi, la Cour de cassation a-t-elle invalidé l'arrêt du tribunal d'instance soutenant que les demandeurs n'ont pas démontré ni une violation des principes généraux du droit électoral ni des irrégularités qui auraient influencé le résultat du scrutin alors qu'il avait été établi que l'expert indépendant n'avait pas eu accès à l'intégralité du code source du système de vote. En effet, en vertu de l'article R.2314-12 du code du travail, préalablement à sa mise en place ou à toute modification substantielle[24] de sa conception, le système de vote électronique est soumis à une expertise indépendante qui statue sur sa conformité aux exigences des articles R.2314-8 à R.2314-10 du même code. L'expert, n'ayant pas accès à tout le code source, était dans l'impossibilité de statuer que le système de vote répondait aux exigences du code du travail, contrairement à l'arrêt cassé[25]. Dans une récente affaire de 2016, le juge de cassation a retenu « *qu'en s'abstenant de rechercher si les difficultés techniques rencontrées pour accéder au système de vote électronique ne portaient pas atteinte à la sincérité du scrutin, le tribunal d'instance a privé sa décision de base légale au regard des articles L. 2324-19, L. 2324-21, R. 2324-4 à R. 2324-17 du code du travail, ensemble les principes généraux du droit électoral* »[26]. Ainsi, il semble que le juge judiciaire soit le plus ouvert aux moyens techniques de preuve pourvu que les irrégularités soient de nature à fausser le scrutin.

En tout état de cause, devant tout juge, il n'est pas simple d'obtenir l'annulation de l'élection en s'appuyant exclusivement sur les moyens techniques.

23 Cass., Soc. 13 janvier 2010, n˚09-60.203. V.: https://www.courdecassation.fr/publications_26/arrets_publies_2986/chambre_sociale_3168/2010_3326/janvier_3327/127_13_14897.html.

24 La question se pose de savoir quelle modification peut être qualifiée de substantielle. Il semble que la qualification soit effectuée par le juge sur le fondement de l'expertise. V. par ex. Cass., Soc. 14 janvier 2014, n˚12-29.253. V. : https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000028483383. V. d'ailleurs Conseil d'État, 11 mars 2015, 368748. V. : https://www.legifrance.gouv.fr/affichJuriAdmin.do?idTexte=CETATEXT000030445583 où le Conseil d'Etat a jugé que l'expertise indépendante devait être effectuée préalablement à chaque scrutin.

25 Ibid.

26 Cass., Soc. 10 mars 2016, n˚15-19.544. V.: https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000032198951.

B. La difficulté de prouver les irrégularités par moyens techniques

« *Le contentieux électoral est un contentieux objectif, il ne s'agit pas de sanctionner une faute éventuelle, mais de s'assurer que le résultat correspond à la volonté des électeurs* »[27]. Ces propos illustrent bien l'approche de tout juge qui n'accepte pas d'argument hypothétique. En ce qui concerne les contentieux liés à l'utilisation des systèmes de vote électronique, le respect de ce principe peut poser problème. En effet, la procédure d'expression des votes[28], de leur acheminement puis du décompte des voix est opaque, autrement dit, il n'est pas possible de mettre en place des observations (numériques ou humaines) qui soient capables de percevoir toutes les atteintes à la sincérité de l'élection, même si ces atteintes sont massives. La question est donc de savoir comment démontrer l'existence d'atteintes à la sincérité de l'élection. En effet, le juge n'accepte pas de moyens de preuve liés à des insuffisances techniques du logiciel de vote[29] en tant que telles dès lors que l'expertise indépendante a bien été effectuée. Notamment, il s'avère peu utile de soulever indépendamment un moyen relatif aux insuffisances techniques du logiciel de vote expertisé. Par exemple, le juge constitutionnel a écarté des griefs soulevant la possibilité de voter par Internet en faveur d'un candidat ne figurant pas sur la liste, la non observance de certaines recommandations de la CNIL dans la définition des modalités de vote, ou encore la sécurisation insuffisante du dispositif de vote contre toute utilisation frauduleuse dès lors que l'expertise indépendante du système de vote électronique avait été effectuée et qu'il n'était pas établi que des faits de fraude aient été commis à l'occasion de l'élection[30].

Cette position semble ignorer l'état de l'art informatique sur ce sujet, à savoir le caractère forcément incomplet d'une expertise technique portant sur un système aussi complexe qu'un vote par internet[31] et l'absence de « *mise en œuvre par le fabricant de processus normalisés de développement logiciel, ni de certification de la qualité de ce dernier selon des normes industrielles* »[32]. Ainsi, les recommandations de la CNIL ne portent que sur des « *garanties*

27 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise*, *préc.*, p.1247

28 Les choix effectués par les électeurs sont dématérialisés, il sont réduits à quelques bits d'information qui sont transformés à plusieurs reprises mais ces transformations successives ne peuvent être observées de bout en bout du fait du caractère secret du vote.

29 Voici deux exemples d'insuffisances techniques à la fois inobservables et susceptibles d'affecter la sincérité du scrutin : des suffrages sont attribués à un autre candidat que celui désigné lors de l'expression de l'intention de vote, des votes sont annulés tandis que d'autres votes sont dupliqués.

30 Conseil Constitutionnel. *Décision n° 2012-4597/4626 AN du 15 février 2013. A.N., Français établis hors de France (4ème circ.)* [réf.29/03/2017]. v.: www.conseil-constitutionnel.fr/decision/2013/20124597_4626an.htm.

31 Enguehard, C. Vote par internet : failles techniques et recul démocratique. Jus Politicum, N°2, Droit, politique et justice constitutionnelle, mars 2009.

32 F. Pellegrini, « Chaînes de confiance et périmètre de certification en matière de ``vote électronique" », in G. J. Guglielmi, et O. Ihl, eds., "Le vote électronique", LGDJ, Lextenso Éditions, 2014, pp. 239-276, ISBN 978-2-275-04466-8.

minimales » limitées au respect du secret du vote mais ne concernent pas tous les aspects de la sincérité du scrutin. En effet, la CNIL a pour mission de faire respecter la loi Informatique et Libertés et est donc limitée au respect du secret du vote. Des dispositions organisant des tests destinés à vérifier que les intentions de vote des électeurs vont bien aux candidats qu'ils ont choisis peuvent être prévues lors du protocole conclu entre l'employeur et les organisations signataires, toutefois des tests sont insuffisants pour garantir le respect de la sincérité de l'élection et une expertise n'est pas en mesure de détecter tous les bugs potentiels. L'article R2314-9 du code du travail précise que « *Le système retenu assure […] la sécurité de l'adressage des moyens d'authentification, de l'émargement, de l'enregistrement et du dépouillement des votes* » et passe ainsi sous silence l'étape pourtant cruciale de l'expression des intentions de votes. Cette étape se déroule sur l'ordinateur utilisé par l'électeur, or ce matériel susceptible d'héberger des virus échappe à l'expertise du système de vote.

Il semble que la position du juge administratif soit identique. Faute de jurisprudence relative à ce sujet, cette position peut être déduite d'après les jugements portant sur les machines à voter. En rejetant le moyen relatif à ce que ces dernières ne satisfont pas aux exigences de l'article L.57-1 du Code électoral, le juge considère en effet « *qu'à la supposer établie, une telle circonstance ne permet pas, à elle seule, de caractériser une atteinte grave à la liberté fondamentale que constitue l'exercice du droit de suffrage* »[33]. Il faut donc non seulement évoquer les insuffisances du logiciel mais aussi démontrer que ces dernières ont donné lieu à des altérations du résultat du scrutin en question. En ce qui concerne le juge judiciaire, il était déjà montré qu'il ne semblait pas non plus opposer une fin de non recevoir au moyen relatif à la légalité d'un logiciel de vote par Internet. Pourtant, il est indispensable de prouver que les insuffisances ont faussé le résultat du scrutin. La question essentielle qui se pose à cet égard est le nombre d'électeurs concernés par le malfonctionnement du logiciel par rapport à l'écart de voix entre candidats. En effet, tout juge réfute systématiquement les moyens soulevant un nombre de votes litigieux inférieur à cet écart.

Un autre angle serait de soulever l'absence de respect, par le système de vote par Internet ou par la procédure de l'expertise indépendante ou par les contrôles dont le système de vote fait l'objet, des exigences techniques prévues par la loi. La difficulté principale dans ce cas est d'obtenir l'ensemble des informations et des documents sur le dispositif de vote en cause. Certaines obligations de transparence sont fixées dans les textes pour les élections professionnelles. En ce qui concerne la représentativité syndicale, l'article R2122-54 du code du travail dispose que le rapport de l'**expertise indépendante** du système de vote est communiqué par le ministère du travail aux

33 TA Versailles, ord., 17 avril 2007, n˚0703990. V. : www.dalloz.fr.

membres du bureau du vote ainsi qu'aux délégués des syndicats candidats. Quant aux élections mises en place au sein de l'entreprise, l'article R2314- 8[34] du même code fixe l'obligation de l'employeur de mettre le **cahier des charges** du système de vote à la disposition des salariés ; les articles R2314-14 et R2314-15[35] - d'informer les syndicats de l'accomplissement des formalités déclaratives auprès de la CNIL, de faire parvenir aux salariés une notice d'information détaillée sur le déroulement des opérations électorales, ainsi que d'assurer une formation sur le système de vote aux représentants du personnel, aux délégués syndicaux[36] et aux membres du bureau de vote. Cependant, le texte du code est muet en ce qui concerne l'obligation explicite de l'employeur de transmettre le rapport de l'expertise indépendante aux personnes mentionnées aux articles R2314-15 et R2324-11. Certains auteurs déduisent du texte que l'employeur doit délivrer aux salariés une information complète sur les modalités et le fonctionnement du système de vote électronique[37]. Cependant, à notre connaissance, il n'y a pas encore de jurisprudence relative à cette question. Il faut ajouter qu'il n'y a aucune garantie ou exigence que les personnes destinataires d'informations ou de rapport d'expertises soient en capacité technique de les apprécier. Par conséquent, il serait nécessaire qu'elles fassent elles-même appel à un expert pour en faire une lecture critique.

Les rapports d'expertise des autres élections dont il s'agit ici (élections consulaires par exemple) ne sont pas publics. Cette absence de communication est justifiée par la protection du secret industriel et commercial ou pour "*raisons de sécurité*" comme nous en avons récemment fait l'expérience avec le refus de communication du "*rapport d'analyse technique du résultat du dépouillement*"[38] des élections consulaires de 2014 par le Ministère des Affaires Etrangères. Or, l'absence de publicité des rapports d'expertise est préjudiciable d'un point de vue de la transparence des élections et ce risque est identifié depuis longtemps : « L*'introduction de technologies au sein du processus démocratique peut diminuer la transparence et risque de donner la priorité aux intérêts privés et commerciaux des producteurs des logiciels, devant l'intérêt public de tenir des élections sincères.* »[39]

34 Pour les élections des délégués du personnel. La même obligation est prévue pour les élections du comité d'entreprise à l'article R2324-4.

35 Pour les élections des délégués du personnel. Les mêmes obligations sont prévues pour les élections du comité d'entreprise aux articles R2324-10 et R2324-11.

36 Sauf les élections du comité d'entreprise.

37 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise*, *préc.*, p.835.

38 Ce rapport est mentionné par le référé n° S2016-3241 de la Cour des Comptes (note de bas de page n°6, page 6).

39 « the introduction of technology into the democratic process can reduce transparency, and risks private commercial interests being given priority over public democratic interests. » Margaret McGaley, Joe McCarthy, Transparency and e-Voting Democratic vs. commercial interests, *Electronic Voting 2006, Workshop "Electronic Voting in Europe – Technology, Law, Politics and Society", Lecture Notes in Informatics*, Robert Krimmer (Ed.), pp.153-163, Bregenz, Austria, July, 7th-9th, 2004.

En ce qui concerne le contrôle du système de vote par Internet, il est très important que les membres de bureau de vote aient des compétences en informatique, ne serait-ce que pour comprendre ses limites et son incomplétude. En effet, le contrôle (nécessairement partiel) du vote électronique est une procédure en grande partie technique. Par exemple, le code du travail prévoit pour les élections dans l'entreprise que les membres de bureau de vote soient accompagnés par la cellule d'assistance technique qui en réalité effectue des contrôles du système : elle procède à des tests du système, vérifie que l'« *urne électronique est vide* », scellée et chiffrée, scelle le système avant et après le vote[40]. Ainsi, la surveillance de l'élection par le bureau de vote devient pour partie vide de sens[41] si ses membres confient entièrement ces opérations à la cellule qui est d'ailleurs mise en place par l'employeur[42]. Même si les les membres du bureau de vote observent une procédure du contrôle sur l'écran d'un ordinateur dédié, il faut bien comprendre que ce qui est affiché sur l'écran peut ne pas correspondre à la réalité des traitements. Ainsi, effectuer un contrôle suppose au moins de surveiller le déploiement du système de vote et de comprendre son fonctionnement. Or, pour les élections organisées par le ministère du travail, ce sont les membres du bureau du vote qui effectuent des contrôles du système, dont les résultats sont ensuite transmis au comité technique et aux délégués syndicaux[43]. Cependant, le bureau du vote est composé des trois magistrats de l'ordre judiciaire et deux de l'ordre administratif[44] dont il n'est pas exigé qu'ils aient des compétences en informatique et se fassent assister par le comité technique. Ce dernier comprend un expert indépendant et deux membres nommés par le ministère du travail[45]. Il est donc possible de constater encore une fois que la possibilité de contrôler l'opération du vote est partielle et, de nouveau, qu'il n'y a pas de garantie qu'un dysfonctionnement affectant la sincérité de l'élection serait détecté.

Le même problème se présente lors du contrôle d'autres élections. Pour les élections des conseillers consulaires, le bureau du vote est composé d'un membre du Conseil d'État, de représentants du ministère des affaires étrangères (dont son directeur des systèmes d'information) ainsi que de trois personnalités qualifiées[46] désignées par l'Assemblée des Français de l'Étranger[47]. De même, les partis et les associations représentatives au niveau national des Français établis hors

40 Articles R2314-18 et R2324-14 du code du travail.

41 V. en ce sens François Pellegrini. Chaînes de confiance et périmètre de certification : le cas des systèmes de « vote électronique » dans Gilles J. Guglielmi, Olivier Ihl, *Le vote électronique, préc.*, p.270-271.

42 Article R2314-13 et R2324-9 du code du travail.

43 Articles R2122-60, R2122-67, R2122-68 du même code.

44 Article R2122-57 du même code.

45 Article R2122-58 du même code.

46 Par exemple, un cadre d'une société commercialisant des systèmes de vote électronique pourrait être désigné comme une personnalité qualifiée.

47 Article 16 du décret n° 2014-290 du 4 mars 2014 portant dispositions électorales relatives à la représentation des Français établis hors de France.

de France qui sont candidats dans au moins trois circonscriptions peuvent désigner un délégué pour contrôler les opérations de vote[48]. Pour ces élections, le bureau du vote peut saisir les autorités et les prestataires de toute question liée au déroulement de l'élection. En vertu de l'article 176-3-4 du code électoral, il se fait également assister par les experts délégués par les responsables du traitement avec le risque que ces experts se substituent *in fine* au bureau de vote. En ce sens, Laurent Touvet et Yves-Marie Doublet notent que « *[l]e contrôle exercé par les assesseurs se réduit ainsi à vérifier que personne ne pénètre dans le local qui contient les machines informatiques ; il ne porte ni sur l'intégralité ni sur la confidentialité du flux des données* »[49].

Face aux difficultés évoquées ci-dessus, on peut se demander si le vote par Internet est vraiment attaquable en justice. Malgré la complexité de la tâche, il est possible néanmoins de distinguer certains moyens de preuve acceptables par le juge.

II. Les moyens d'obtenir l'annulation du scrutin restant à la disposition des requérants

Certes, il est toujours possible d'attaquer le système de vote par Internet par les moyens liés aux exigences techniques et au fonctionnement du système de vote. La difficulté est ce que, tout d'abord, seule une partie des dysfonctionnements ou des vices du système peuvent être perçus et que ceux qui peuvent être perçus ne peuvent pas être toujours prouvés. Ainsi, un malfonctionnement d'un système de vote électronique énonçant des résultats insincères pourrait passer inaperçu. Seule une partie réduite des insuffisances techniques est susceptible de servir de preuve entraînant l'annulation de l'élection par le juge (A). Faute de tels moyens de preuve, il est conseillé aux demandeurs de se tourner vers d'autres moyens, non liés aux insuffisances techniques du logiciel du vote (B).

A. Les preuves possibles des insuffisances du logiciel de vote

En ce qui concerne les élections professionnelles, la solution la plus simple est de comparer le système de vote utilisé dans l'entreprise avec le cadre des « *précautions suffisantes* » pour garantir la confidentialité des données élaboré par le juge dans l'arrêt du 21 septembre 2016. Les mesures suivantes constituent ce cadre : « *les codes et les identifiants [sont] personnels, obtenus de manière aléatoire et à usage unique, que [le prestataire] avait mis en place une phase postérieure de*

48 Article 17 du même décret.

49 Laurent Touvet, Yves-Marie Doublet. *Droit des élections*. Paris: Economica. 2014. p.551.

validation du vote par l'électeur lui-même, que les documents internes à [l'entreprise] mettent en évidence une restriction et une sécurisation non seulement de la messagerie avec des adresses électroniques uniques et des mots de passe strictement personnels à chaque salarié, mais aussi des accès à la messagerie professionnelle par des administrateurs réseau eux-mêmes avec la traçabilité des interventions et des engagements de confidentialité, que les codes ont été envoyés par [le prestataire], que le vote se faisait exclusivement sur les serveurs de cette société dédiés à cette élection et sécurisés contre les intrusions, que le système informatique de [l'entreprise] n'était pas impliqué dans le processus de vote, que le vote en lui-même faisait l'objet de trois chiffrements successifs sécurisant ainsi l'échange entre le terminal de l'utilisateur et la plate-forme [du prestataire], de sorte que la direction ne pouvait avoir connaissance du vote crypté immédiatement stocké dans l'urne dédiée, qu'il y avait deux flux, l'un pour le vote et l'autre pour l'émargement, de sorte que pendant les opérations électorales les administrateurs (assesseurs et organisateurs) avaient accès au second et non au premier, le décryptage des votes ne pouvant se faire qu'à la clôture du scrutin avec l'introduction de deux clés d'accès simultanément »[50]. Certes, ce cadre n'est pas à l'abri de critiques : par exemple, la mise en place de « *la phase postérieure de validation du vote par l'électeur lui-même* » ne garantit aucunement que l'information effectivement envoyée vers le serveur de vote porte bien l'intention de vote de l'électeur[51]. Cependant, il peut servir de point de départ pour contester les élections.

Si le système de vote en question satisfait au cadre cité mais que des irrégularités sont constatées, il convient de comparer le cahier de charges élaboré conformément aux articles R2314-8 et R2324-4 du code du travail qui est mis normalement à la disposition des salariés avec le rapport de l'expertise indépendante. Si l'employeur refuse de transmettre ce rapport d'expertise, il est possible de l'obtenir en s'adressant au juge car, même si la charge de la preuve incombe habituellement au demandeur, il existe une certaine inégalité des parties du fait que seul l'employeur détient toutes les informations et les documents relatifs au système de vote[52]. Il convient en outre de comparer le rapport de l'expertise avec les exigences des articles R2314-9 et R2324-5 qui prévoient que le système de vote doit assurer « *la confidentialité des données transmises, notamment de celles des fichiers constitués pour établir les listes électorales des collèges électoraux, ainsi que la*

50 Cass., Soc. 21 septembre 2016, n˚15.60-216. V.: https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000033145504.

51 Enguehard, C. Le vote en ligne : expertise du système de vote électronique et chiffrement du bulletin de vote. Colloque "La démocratie : du crépuscule à l'aube ?", CREDOF, Nanterre, 13-14 juin 2013.

52 Par exemple, c'est déjà acquis concernant les listes électorales. Cass., Soc. 13 novembre 2008. V. : https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000019772436. V. aussi Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise*, *préc.*, p.1216, 1234.

sécurité de l'adressage des moyens d'authentification, de l'émargement, de l'enregistrement et du dépouillement des votes ». Si le système de vote est conforme à ces exigences selon le rapport de l'expertise, il est possible néanmoins de vérifier si la mise en œuvre de ces exigences est efficace. En ce sens, il pourrait être utile de recourir à un autre expert indépendant. En effet, une mise en œuvre insuffisante de ces exigences au niveau technique, pourrait donner lieu à l'annulation de l'élection. Il semble cependant qu'il serait très difficile d'obtenir l'intervention d'un second expert. Un autre écueil non négligeable est le coût élevé d'une telle expertise.

Par ailleurs, ces spécifications ne sont que partielles : elles ne visent qu'à sauvegarder le secret du vote et ainsi n'assurent que partiellement le respect des exigences constitutionnelles au droit du suffrage qui, en vertu de l'article 3 de la Constitution de 1958, est libre[53], universel, égal et secret. De plus, les articles en question n'ont qu'une valeur réglementaire. Au niveau de la loi, seule l'exigence du vote secret existe ce qui est effectivement prévue par les articles L2314-21 et L2324-19 du code électoral. Il ne semble pas opportun de soulever l'inconstitutionnalité de ces articles, le fait qu'ils n'évoquent pas toutes les exigences de l'article 3 de la Constitution n'étant pas en lui-même contraire à la Constitution. Ainsi donc, même en sachant le caractère insuffisant des exigences au système de vote par Internet prévues par le code du travail, il n'est pas possible de les contester en justice. Face à des irrégularités liées au non respect des exigences absentes dans les articles sus-mentionnés, il est en revanche toujours possible d'évoquer directement la Constitution.

En ce qui concerne la représentation des Français établis hors de France, les dispositions législatives[54] instaurent les exigences de secret du vote et de sincérité du scrutin au système de vote utilisé. Il a été montré plus haut que cette dernière est une notion comprenant l'égalité, la liberté et le secret du vote. Le Conseil d'Etat, saisi par un recours pour excès de pouvoir visant l'annulation de l'arrêté du 27 avril 2012 relatif au traitement automatisé de données à caractère personnel prévu à l'article R. 176-3 du code électoral, a précisé le cadre de leur respect par le système de vote en cause. Selon lui, « *le recours à un système de vote électronique est subordonné à la réalisation d'une expertise indépendante, lors de sa conception initiale, à chaque fois qu'il est procédé à une modification substantielle ainsi que préalablement à chaque scrutin ; qu'en outre, [...] l'identifiant et l'authentifiant sont transmis à l'électeur par des modes d'acheminement différents, que l'authentifiant est renouvelé en cas de second tour de scrutin et qu'en cas de perte, seul l'identifiant peut être récupéré par l'électeur ; qu'enfin, [...] le respect du secret du vote, de la sincérité du*

53 « Même si cet article ne fait pas référence de manière expresse à la liberté de suffrage, celle-ci est implicite: toute violation de cette liberté serait, en effet, directement contraire au principe démocratique réaffirmé à l'article premier de la Constitution. » RICHARD GHEVONTIAN. La notion de sincérité du scrutin, *préc.*

54 Article 22 de la loi n°2013-659 du 22 juillet 2013 et article L330-13 du code électoral.

scrutin et de l'accessibilité au suffrage doivent être garantis au stade de la mise en oeuvre du traitement [automatisé de données à caractère personnel] »[55]. Autrement dit, le système de vote en cause doit assurer la sincérité du scrutin, sauf qu'il est difficile s'en assurer faute de la publicité des rapports de l'expertise, ainsi qu'étant donné son caractère inévitablement partiel (v. plus haut sur les limites de l'expertise).

Force est de constater qu'en tant que telles, les insuffisances techniques du système de vote sont difficilement contestables afin d'obtenir l'annulation de l'élection. Il convient donc de les assortir par d'autres moyens, non liés à l'aspect technique de l'organisation de l'élection.

B. Les moyens non techniques d'obtenir l'annulation du scrutin

Certaines irrégularités des élections professionnelles donnent lieu à l'annulation sans avoir à chercher si elles ont influencé le résultat du scrutin. Avant l'entrée en vigueur de la loi Travail de 2016, tel était le cas de l'absence d'un accord d'entreprise ou de groupe autorisant le recours au vote par Internet. Maintenant, en l'absence d'un tel accord, l'employeur peut décider unilatéralement. En cas d'échec des négociations sur le protocole préélectoral, les modalités de la mise en œuvre d'un tel procédé seront fixées par le tribunal d'instance (dans le cas où il y a un accord d'entreprise, les modalités seront fixées dans les conditions prévues par l'accord d'entreprise[56]). Il y a donc tout intérêt, pour les syndicaux représentatifs, à participer à la négociation de l'accord d'entreprise et du protocole préélectoral[57], à demander tous les documents et les informations liés au système de vote y compris le rapport de l'expertise ainsi qu'à définir la composition du bureau de vote afin qu'il y ait des personnes ayant des compétences en informatique.

Plus généralement, selon Marie-Laure Morin, Laurence Pécaut-Rivolier et Yves Struillou, font partie des irrégularités donnant lieu à l'annulation la violation des règles générales d'ordre public en matière électorale (comme la présence d'un non électeur dans le bureau de vote ou la composition de ce dernier par son seul président) ou des principes généraux du droit électoral[58] (la violation du secret du vote ou de la confidentialité des données transmises[59], l'absence des représentants des

55 CE, 27 juillet 2015, N° 360813. V.: www.dalloz.fr.

56 L'art. L2324-21 du code du travail. V. aussi Cass., Soc. 4 juin 2014, n˚13-18.914. V. : https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000029057354.

57 L'employeur ne peut pas refuser de tenir une réunion en vue de la négociation du protocole préélectoral suite aux demandes d'une organisation syndicale. Un tel refus constitue une cause d'annulation de l'élection. Cass., Soc. 9 novembre 2011, n° 11-60.029 11-60.030 11-60.031 11-60.032. V. : www.dalloz.fr.

58 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise*, *préc.*, p.842.

59 Cass., Soc. 14 décembre 2015, n˚15-16.491. v. : https://www.legifrance.gouv.fr/affichJuriJudi.do?idTexte=JURITEXT000031656591; Cass., Soc. 27 février 2013, n˚12-16.789. v. : www.dalloz.fr.

syndicats pour surveiller le scrutin, le non respect de l'obligation de clôturer publiquement le scrutin sous le contrôle des électeurs et des délégués), puis l'existence d'irrégularités particulièrement graves (le vote d'un très grand nombre d'électeurs a été affecté, il y avait des violences, des abus manifestes de propagande[60], etc.) ou d'irrégularités formelles de nature à jeter un doute sur la sincérité et la transparence du scrutin (l'absence de mention de l'heure d'ouverture et de clôture du scrutin)[61]. D'un autre côté, les auteurs distinguent les irrégularités de nature à fausser les résultats du scrutin qui peuvent donner lieu à l'annulation du scrutin si elles ont influencé ses résultats. Par exemple, l'arrivée trop tardive du matériel de vote chez les électeurs[62] (dans le cas du vote par Internet, l'on peut songer à l'arrivée tardive des codes de vote) dans le cas où l'irrégularité a influencé un tel nombre d'électeurs que la sincérité du scrutin peut être mise en doute. Toutefois, nous pouvons soulever les difficultés pratiques pour rassembler les témoignages des électeurs affectés. En ce qui concerne la représentativité syndicale, les auteurs distinguent encore deux hypothèses d'annulation du scrutin : « *l'irrégularité qui affecte le calcul de l'audience électorale ne peut entraîner l'annulation des élections que si, s'agissant du premier tour, elle a été déterminante pour l'accès du syndicat à la qualité représentative ou la possibilité pour un candidat d'être désigné délégué syndical. Concrètement, il en est ainsi si l'irrégularité est susceptible d'avoir influé sur le respect du seuil de 10% revendiqué par un syndicat ou un candidat.* »[63].

Pour les élections des députés des Français établis hors de France, « *l'élément qui détermine le Conseil [constitutionnel] à annuler ou non une élection est l'écart des voix entre le candidat battu et l'élu. Quel que soit l'objet des irrégularités – pressions, tracts mensongers ou diffamatoires, erreur de recensement des votes, procurations mal établies..., quelle que soit même leur importance, elles n'entraînent pas l'annulation de l'élection si un écart de voix important sépare le candidat élu de son adversaire ou, plus généralement, si ces irrégularités sont jugées par le Conseil n'avoir pas eu d'influence déterminante sur le résultat de l'élection. En revanche, la combinaison d'irrégularités graves et d'un faible écart de voix entre le candidat élu et son concurrent entraîne l'annulation de l'élection* »[64]. Dans l'impossibilité de se prévaloir d'un tel écart, il est possible de

60 Cass., Soc. 14 janvier 2014, n˚12-29.253, préc.

61 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise, préc.*, p.1249.

62 Cass., Soc. 21 octobre 1985, n˚85-60.221. v.: www.dalloz.fr. A comparer avec Conseil constitutionnel. *Décision n°2012-4554 AN, 15 février 2013, A.N., Français établis hors de France (7ème circ.)*. [réf.31/03/2017]. v.: www.conseil-constitutionnel.fr/decision/2013/20124554an.htm. où le juge ne s'est pas prononcé sur la reception tardive de l'authentifiant par certain nombre d'électeur en raison de la requête tardive et donc irrecevable.

63 Marie-Laure Morin, Laurence Pécaut-Rivolier, Yves Struillou, *Le guide des élections professionnelles et des désignations de représentants syndicaux dans l'entreprise, préc.*, p.1250.

64 Dominique Rousseau, Pierre-Yves Gahdoun, Julien Bonnet. *Droit du contentieux constitutionnel*. 9e édition. Issy-les-Moulineaux: LGDJ-Lextenso éditions. 2016. p.474-475.

demander l'annulation sur le moyen d'inéligibilité du candidat élu ou sur la violation des règles du financement de la campagne. Tel était le cas lors des élections de 2012 où le Conseil a invalidé l'élection dans la 1ère circonscription (Etats-Unis et Canada)[65] en prononçant l'inéligibilité d'une candidate élue pour cause du rejet de son compte de campagne par la Commission nationale des comptes de campagne et des financements politiques, « *sans qu'il soit besoin d'examiner les griefs de la requête* » alors que le requérant alléguait également qu'un très grand nombre d'électeurs n'ont pas pu accéder au système de vote par Internet[66].

Enfin, en ce qui concerne le juge administratif, faute de jurisprudence permettant de retracer, avec un degré suffisant de certitude, sa perception des possibilités d'annulation du vote par Internet, on ne peut que réfléchir par analogie. Etant donné qu'il se fonde sur le rapport de l'expertise quant à l'évaluation de la comptabilité de la solution logicielle utilisée avec les exigences techniques, l'absence d'expertise avant chaque scrutin semble pouvoir servir de fondement d'annulation des résultats du scrutin[67]. Puis, en comparant le vote par internet avec le vote par le biais des machines à voter, il est possible de voir que l'approche du juge est assez formelle. Même s'il utilise ses pouvoirs d'instruction pour demander aux autorités publiques les documents relatifs aux machines à voter en cause, il les analyse non d'un point de vue technique, par le biais d'une expertise supplémentaire par exemple, mais pour s'assurer que la procédure de l'agrément[68] et de l'information des intéressés a été suffisamment bien organisée. Dans l'arrêt du 4 novembre 2009, il retient que « *le dossier de demande d'agrément présenté comportait notamment un dossier communes, à destination des mairies, des manuels d'installation et d'utilisation, une notice de maintenance, et une documentation matériel, documents tous rédigés en français ; qu'il en ressort également qu'ont été organisés plusieurs visites ou audits des sites de conception et de production des machines à voter et que le bureau Veritas a disposé de tous les éléments, dispositifs et documents nécessaires à son contrôle, en vue de vérifier la conformité de la machine à voter en cause à l'ensemble des exigences*

65 Conseil constitutionnel. *Décision n°2012-4551 AN, 15 février 2013, A.N., Français établis hors de France (1ère circ.)*. [réf.08/06/2017]. v.: www.conseil-constitutionnel.fr/decision/2013/20124551an.htm.

66 Pierre Ciric. How I Convinced the French Supreme Court to Cancel the First French Legislative Elections in the United States. *South Carolina Journal of International Law and Business*, 2014, 10 (2), p.270-297. p.288.

67 V. *mutatis mutandis* CE, 11 mars 2015, 368748. V. : https://www.legifrance.gouv.fr/affichJuriAdmin.do?idTexte=CETATEXT000030445583.

68 Il faut toutefois noter que, dans les faits, des "*exigences*" du règlement technique ne sont pas respectées par des machines à voter qui ont pourtant été agréées par le ministère de l'intérieur, ce qui a attiré défavorablement l'attention de la mission d'observation des élections présidentielles 2007 de l'Organisation pour la Sécurité et la Coopération en Europe (OSCE) (voir le Rapport de la Mission d'évaluation électorale de l'élection présidentielle France 22 avril et 6 mai 2007). De plus, le juge se désintéresse du fait qu'il s'agit d'un agrément d'un modèle de machine à voter, et non de chaque exemplaire de machine à voter. Or, il n'existe aucune possibilité de vérifier que les machines effectivement utilisées dans les bureaux de vote sont identiques à la ou les machines qui ont été examinées par les experts du bureau Veritas.

du règlement technique »[69]. S'il constate la discordance entre le nombre de suffrages comptabilisés et le nombre d'émargements, dès qu'il est possible de rectifier les résultats, il rectifie après s'être assuré par le biais de l'instruction, que la discordance relève d'erreurs de manipulation et non provoquée par l'utilisation des machines à voter « *dont le modèle a été agréé par arrêté du ministre de l'intérieur* »[70].

Plus généralement, son attitude envers les irrégularités ne diverge pas beaucoup de celle du juge constitutionnel : il faut qu'il y ait un écart des voix suffisant pour mettre en doute la sincérité du scrutin pour que le juge administratif accepte les moyens relatifs aux irrégularités en vue d'annuler l'élection en cause. Il existe néanmoins des cas permettant annuler l'élection en toute hypothèse, soit parce qu'il est impossible de définir les véritables résultats du scrutin[71], soit pour cause d'inéligibilité[72] ou d'incompatibilité[73] du candidat élu.

69 Conseil d'État, 4 novembre 2009, n° 306563. V.: www.dalloz.fr.

70 Conseil d'État, 1 décembre 2010, n° 337945. V. : www.dalloz.fr.

71 Conseil d'État, 2 septembre 1983, n° 51182, 51853. V. : www.dalloz.fr. « Lorsqu'une irrégularité de nature à altérer les résultats du scrutin a été commise mais que le nombre de suffrages recueillis par chacune des deux listes en présence ne peut être déterminé avec certitude au vu du dossier soumis à la juridiction administrative, il y a lieu pour celle-ci d'annuler les opérations électorales ».

72 Conseil d'État, 16 décembre 1983, n° 52117. V. : www.dalloz.fr.

73 Conseil d'État, 3 mars 1984, n° 53067. V. : www.dalloz.fr.

Fiabilité et sincérité des systèmes blockchain

Nicolas Herbaut[1] and François-Vivien Guiot[2]

[1]Université Paris 1 Panthéon Sorbonne
[2]Université Toulouse 1 Capitole

Introduction

Le terme de "Système" est un terme qui a priori fait sens aussi bien pour les juristes que pour les "vrais" scientifiques. Il devrait donc constituer un bon vecteur de convergences. Il n'est cependant pas certain que la notion fasse tout à fait sens de la même façon pour les uns et pour les autres. Ainsi, si dans le domaine du numérique on devrait accepter de retenir le système comme un " ensemble d'éléments qui dépendent réciproquement les uns des autres de manière à former un tout organisé " (Lalande, 1926 : 1097), au sein de la doctrine juridique on s'arrêtera également et peut-être davantage sur une seconde acception qui voit dans le système un " ensemble d'idées logiquement solidaires et tendant à offrir une vue cohérente d'un objet ou d'un champ d'étude " (id.). D'un côté le système est l'objet de la pensée, tandis que de l'autre il est une qualité de la pensée elle-même – pensée qui vise alors à la systématisation.[1]

Quoiqu'il en soit de ces nuances, il semble que la notion de système a soulevé lors de nos travaux préparatoires une même préoccupation chez les différents intervenants au premier atelier des convergences du droit et du numérique, qu'ils soient juristes ou scientifiques : celle de la confiance. Peut-on avoir confiance dans le système ? Que l'on appréhende le système construit par les algorithmes, en tant que tel, ou dans sa relation avec une pratique qu'il entend "disrupter", la mécanisation, l'automatisation, ou – plus globalement – la logicisation sont ou restent sources de méfiance.

C'est ainsi à travers la double problématique de la fiabilité et de la sincérité que nous avons décidé de converger à propos d'un objet d'intérêt commun, les systèmes Blockchain. Il s'agissait pour le juriste d'adopter une approche critique suivie de longue date par la recherche en informatique visant à démythifier le caractère infaillible comme l'objectivité des systèmes informatisés.

Il existe un terme qui exprime très bien cela : celui de " système inéquitable numérique ". Le"SIN" a pour objet de désigner un système qui pêche par ce qu'il fait peser le coût de certains dysfonctionnements sur l'utilisateur – lequel sera évidemment la partie juridiquement et économiquement faible en cas de litige[2]

Dans certains cas, c'est au législateur qu'il est revenu d'intervenir pour rétablir l'équité face au système. En vertu de la loi, c'est ainsi à la banque d'assumer l'essentiel des risques liés à l'utilisation de la carte bancaire (même en absence d'opposition, la responsabilité de l'utilisateur est limitée à 150 euros en vertu de l'art. L 133-19 du Code mon. et fin.).

[1]A dire vrai, il y a dans cette présentation des approches respectives de la science et du droit une forme de réduction puisque dans le champ juridique les deux acceptions de la notion de système sont connus. La question de savoir si c'est l'objet juridique en lui-même qui est organisé, ou si cette organisation est l'œuvre de son observateur est ainsi une question bien connue de la théorie du droit (Grzegorczyk 2002).

[2]Il est possible de prendre un exemple topique avec le cas du défaut de validation d'un titre de transport : l'utilisateur ne peut pas prouver qu'il a bien entendu le composteur biper à la présentation de sa carte numérique.

Confiance dans les systèmes blockchain

Compte tenu du rôle de plus en plus prépondérant des solutions Blockchain dans l'organisation de la vie économique et sociale, de leurs potentialités et de leurs risques[3], nous avons souhaité investir notre travail dans la question de la fiabilité et de la sincérité de ce type de système.

Présentation de la blockchain

La blockchain est perçue comme un composant d'une architecture logicielle permettant d'atteindre un consensus distribué pour les données transactionnelles sans recourir à un tiers de confiance (Xu et al. 2016). Ces systèmes consistent en une base de données permettant la lecture et l'ajout d'information sous forme d'une liste d'enregistrements chaînés appelés les blocks. Par construction, les systèmes de blockchain ne peuvent être falsifiés ou modifiés puisque chaque block contient un marqueur temporel couplé à un lien vers le block précédent (Bozic, Pujolle, and Secci 2016). La blockchain offre donc supposément l'assurance que les données ne peuvent être modifiées rétroactivement une fois enregistrées.

Un tel consensus décentralisé peut être atteint à l'aide d'algorithmes tels que proof-of-work (Nakamoto 2008) (preuve de travail), proof-of-stake (Kiayias et al. 2017) (preuve d'enjeu) ou d'algorithmes dits tolérants aux fautes Byzantines (BFT) (Veronese et al. 2013). Les blockchain peuvent être utilisées dans une grande variété de cas d'utilisation tels que les transactions financières comme Bitcoin, les dossiers médicaux ou encore le contrôle des réseaux (Herbaut and Négru 2017).

Les implémentations les plus répandues des blockchain, telle celle utilisée pour le réseau Bitcoin, ont démontré leur efficacité dans la gestion de transactions très simples, comme l'échange de devises. Néanmoins, le manque de moyen technique permettant une programmabilité extensible à d'autres cas d'utilisations plus complexes ont conduit au développement d'une nouvelle génération de blockchain. Celle-ci étend la sémantique des transactions au travers l'utilisation de "Smart Contracts" (Szabo 1997).

Parmi les différents essais lancés à grande échelle sur les marchés, l'épisode de *The DAO* (Jentzsch 2016) a constitué notre point d'entrée dans notre réflexion sur la problématique de la confiance des systèmes Blockchain.

The DAO ou l'échec de la confiance

The Dao a été fondé le 30 avril 2016, en tant qu'organisation autonome décentralisée et formait un fond d'investissement piloté directement par ses actionnaires, en fournissant un nouveau business model décentralisé appliqué à l'organisation des entreprises commerciales et organismes sans but lucratif.

Au démarrage de l'activité, le système a été techniquement déployé sur la blockchain Ethereum et était dépourvu de structure managériale ou de conseil d'administration. Son code fut publié sous licence open-source. Issu d'un financement participatif, il a détenu le record de capitalisation pour ce type de création. En Juin 2016, une vulnérabilité dans son code a permis à des utilisateurs malveillants de détourner un tiers de la valeur du fond (Atzei, Bartoletti, and Cimoli 2017).

En Juillet 2016, la communauté Ethereum a décidé de revenir en arrière sur ces transactions délictueuses en les supprimant de ses registres. Dans la controverse, une partie de la communauté, opposée à ce "hard fork" continue de maintenir la blockchain Ethereum dans son état non modifié sous le nom d'Ethereum Classic. Les deux communautés s'opposent de façon quasi philosophique sur ce retour en arrière.

- La communauté pro-fork **Ethereum** soutient que l'intentionnalité du smart contrat défectueux n'avait pas été respectée, et que les dé-

[3] dernièrement soulignés par Christine Lagarde en sa qualité de directrice du FMI : https://blogs.imf.org/2018/04/16/

tenteurs de theDao n'avaient pas exprimé leur accord sur les transactions frauduleuses.
- La communauté anti-fork **Etherum Classic** soutient que le code contenu dans les Smarts Contracts représente le plus haut niveau de vérité, et qu'il fait loi (selon le principe "code is law" : (Lessig 1999)). Ces représentants arguent également que la propriété d'immutabilité de la blockchain, sur laquelle repose la confiance des utilisateurs doit être garantie.

Dans la suite de cet article, nous proposons un modèle pragmatique qui est un compromis entre les deux approches. D'une part, puisqu'il est difficile aujourd'hui de prouver mathématiquement qu'un smart contrat ne peut être exploité par un utilisateur malveillant à des fins contraires à l'intentionnalité de l'auteur du contrat, nous proposons d'exprimer un ensemble d'invariants très simple, permettant de s'assurer que les cas limites ne peuvent entraîner d'exploitation abusive. D'autre part, nous proposons qu'une fois que les invariants ont été vérifiés, les transactions soient inscrites dans la blockchain de manière immuable.

Code is law vs code by law

Il s'agit d'envisager la possibilité de concevoir des Smart Contracts " non directement et automatiquement opposables ", en ajoutant au système un mécanisme d'appel ou d'arbitrage au stade de l'exécution du contrat. Au-delà d'une réflexion sur les valeurs du cyberespace, à laquelle nous invite le constitutionnaliste américain Laurence Lessig depuis le début des années 2000, il y a un enjeu très pragmatique : comme ce fut le cas pour toute innovation du monde économique, depuis la monnaie scripturale à Airbnb, c'est sur la confiance que repose la capacité des Smart Contracts à se développer.

En pratique, il existe deux voies essentielles et complémentaires afin d'assurer une bonne diffusion de cette nouvelle technologie : il faut tout d'abord travailler sur l'architecture du système, c'est-à-dire sur le code lui-même afin de sécuriser son fonctionnement et l'automatisation de ses fonctions. Il faut ensuite réfléchir à la manière dont le droit peut ou doit garantir les parties contractantes, non plus uniquement de la bonne exécution des engagements réciproques, mais de l'existence des conditions propres à garantir une bonne auto-exécution du contrat. Cela revient à adopter deux approches : une première, "code is law", qui tire les conséquences du fait que pour tout système numérique, c'est d'abord sa conception qui régule son fonctionnement ; une seconde, "code by law", qui s'interroge sur la capacité du droit à investir ce système afin d'y exercer son pouvoir normateur.

Code is Law

Les Smart Contracts sont généralement écrits à l'aide de langages de programmation non spécialisés (Dannen 2017) (Androulaki et al. 2018) afin de permettre au développeur les réalisant une productivité similaire au développement d'applications traditionnelles. On peut concevoir les Smart Contracts comme l'exécution de code déterministe ayant comme entrée un état donné de la blockchain et produisant des sorties également inscrites dans la blockchain. Entrées et sorties peuvent être considérées comme un ensemble de valeurs rattachées à un compte utilisateur, les contrats permettant de transférer des valeurs d'un compte à un autre.

Face aux difficultés de sécuriser les Smart Contracts par la production de preuves (Bhargavan et al. 2016) (Hirai 2017) ou par l'emploi de chasseur de prime pour détecter les exploitations possibles des anomalies des Smart Contracts (Breidenbach et al. 2018), nous proposons une approche complémentaire basée sur l'encapsulation des Smart Contracts dans des containers d'exécution arbitraux (CEA), présenté Figure 1 possédant différentes propriété que nous allons décrire.

Containers d'Exécution Arbitraux

Premièrement, l'exécution du Smart Contract dans le CEA reste similaire aux smart contacts classiques. Ils accèdent aux même données, et leurs résultats sont

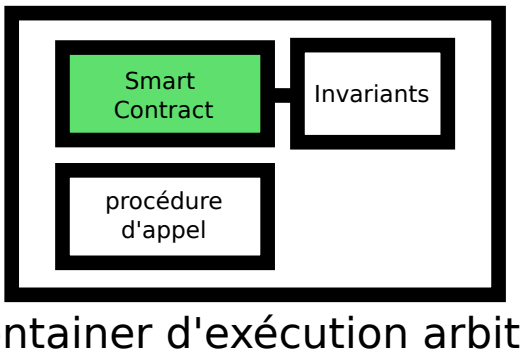


Figure 1: Encapsulation du Smart Contract dans son container d'exécution

également stockés dans la blockchain. Tout code exécuté dans le container avec les mêmes conditions initiales doit aboutir au même état de sortie. La principale différence entre le contrat exécuté directement dans la blockchain et le contrat rattaché à un CEA est que le CEA ajoute une série d'invariants concernant les entrées et sorties des contrats. Ainsi, le CEA consiste en un environnement d'exécution de plus haut niveau que le contrat initial, permettant de vérifier si certains invariants exprimables en fonction des paramètres d'entrée et de sortie sont vérifiés lors de l'exécution du contrat, comme montré sur la Figure~2.

Dans le cas nominal, tous les invariants sont satisfaits et le résultat de l'exécution du CEA est directement publié sur la blockchain sans attendre. Dans ce cas, l'exécution du Smart Contract originel est directement publiée sur la blockchain et les transactions afférentes à la consommation des ressources des différentes parties prenantes sont définitivement validées. En cas de violation des invariants du CEA, le résultat associé, appelé *résultat transactionnel* est publié sur la blockchain, mais il est décrété *non opposable* (celui-ci n'est pas encore définitif, et les ressources générées par ce contrat ne sont pas utilisables dans d'autres contrats). Le *résultat transactionnel* est associé à un nouveau contrat ad-hoc dit *contrat d'appel* permettant aux parties prenantes de les contester dans le cadre d'une *procédure d'arbitrage*.

Procédure d'Arbitrage

La procédure d'Arbitrage permet aux parties prenantes de contester les résultats de l'exécution d'un contrat sans altérer les propriétés d'immutabilité de la blockchain. En effet, les résultats transactionnels, avant d'être déclarés *opposables* peuvent faire l'objet d'un appel permettant d'aboutir à 3 résultats différents, illustré Figure 3

- Si aucune partie prenante ne souhaite faire appel de l'exécution initiale du contrat, le mécanisme d'arbitrage contourne la violation des invariants. Les conditions initiales du contrat sont validées et les résultats sont inscrits directement sur la blockchain et sont décrétés opposables.
- Si une partie prenante fait appel à un arbitrage, deux cas de figure peuvent intervenir: l'arbitrage peut invalider l'exécution du contrat, jugeant que la violation des invariants est contraire à l'esprit initial du contrat. Dans ce cas, le résultat transactionnel est déclaré non écrit, et les ressources dépensées par les parties prenantes dans le cadre de l'exécution du contrat initial sont restituées conformément à la procédure d'arbitrage.1 Dans le cas contraire, l'arbitrage peut être rendu en conformité avec l'exécution initiale. Dans ce cas, le résultat transactionnel devient un résultat opposable et est écrit en tant que tel sur la blockchain[4].

[4]L'arbitrage pouvant précisé par exemple l'annulation pure

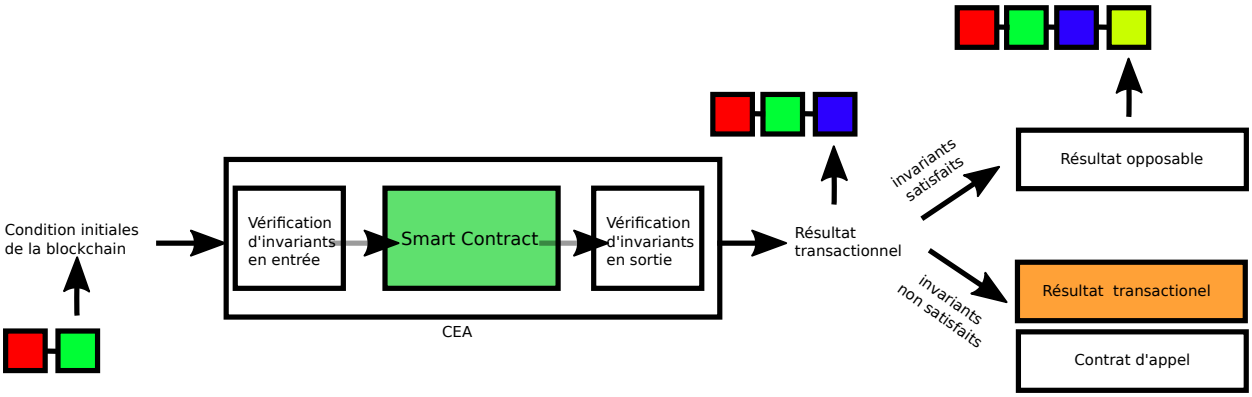


Figure 2: Exécution d'un Smart Sontract encapsulé permettant la vérification des invariants

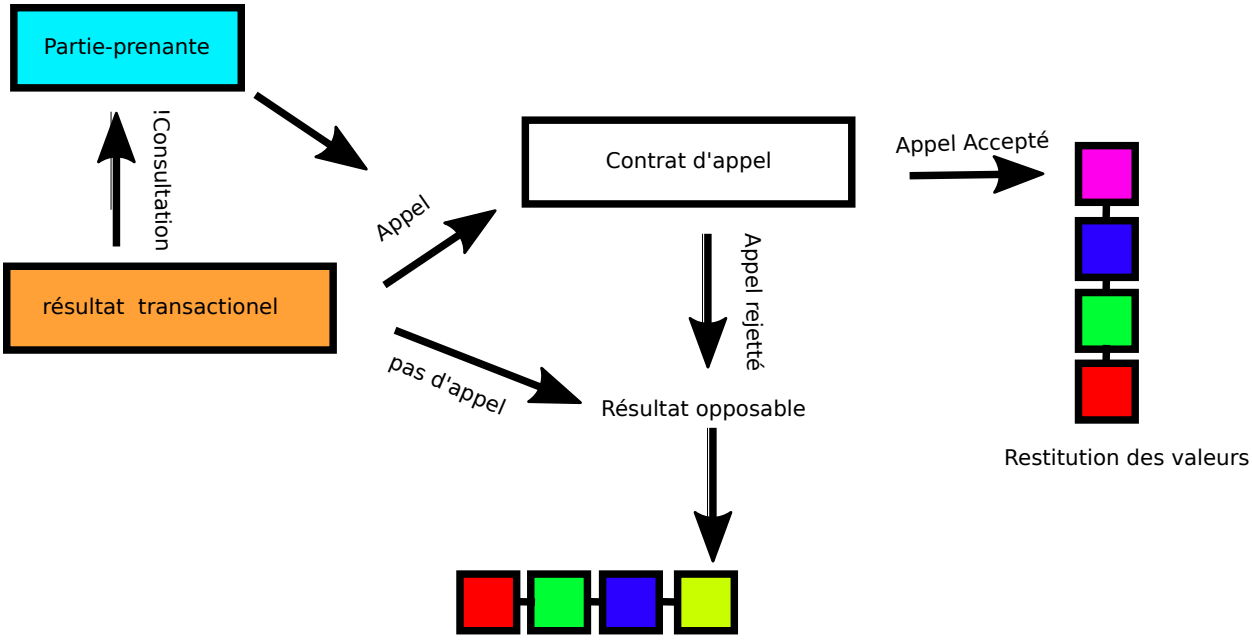


Figure 3: Exécution de la procédure d'arbitrage

Le CEA joue le rôle dépôt fiduciaire, qui ne procède au déblocage des valeurs que lorsque les résultats du smart contrats deviennent opposables.

Exemple du TheMcDAO

Afin d'illustrer l'utilisation des CEA et de la procédure d'arbitrage, nous proposons un exemple basé sur une solution de livraison de repas à domicile basée sur la blockchain: theMcDAO.

Une entreprise autonome theMcDAO propose la livraison de repas à domicile, directement basé sur la blockchain. Le service est basé sur la collaboration de 4 acteurs différents qui vont entrer en jeu pour livrer un hamburger chaud à un client, au meilleur prix. Le premier type d'acteur est le client. Celui-ci spécifie le menu de son choix, basé sur la carte d'une franchise de restauration rapide. Il spécifie également le délai de livraison attendu dans une application mobile et procède au paiement dans une devise définie par la blockchain. Les valeurs sont stockées dans le CEA du contrat de livraison de repas, qui précise la répartition de la valeur ajoutée de la prestation entre la plateforme, le prestataire de livraison et le restaurateur.

Dans un deuxième temps, la plateforme theMcDAO va choisir, parmi une liste de restaurateurs franchisés, celui qui propose de préparer la commande au meilleur prix et dans les délais impartis. Tous les restaurateurs déclarent leur disponibilité et leur prix en temps réel à la plateforme. Dès lors que la commande est réceptionnée et préparée, le hamburger est placé dans une boite témoin spéciale, qui affiche un QRCode unique lorsque la température du contenant atteint une température inférieure à 40°C. Une fois tous les éléments prêts, le sac de transport est fermé et affiche alors un QRCode unique, flashé par le restaurateur. Celui-ci dépose la commande en attendant que le livreur vienne la chercher. Le livreur, qui s'est engagé à prendre en charge la commande dans un délai imparti, arrive au restaurant. Il flash le QRCode du sac de transport et va le livrer chez le client. Lors de la remise de la marchandise, le livreur flash le QRcode du client et termine sa livraison.

Dans ce scenario, le contrat lie le client, le restaurateur et le livreur. La liste des invariants est définie comme :

1. L'horodatage du livreur respecte les délais d'engagements de livraison (heure de réception, délai de transit)
2. L'horodatage de restaurateur respecte les délais de préparation (heure de service de la commande)
3. L'horodatage du client respecte le délai de livraison (heure de réception).
4. La boite témoin n'indique pas une température inférieure à la température cible.

Si, dans les 5 minutes suivant sa livraison, le client n'a pas fait de remarque sur la qualité de sa commande, les devises sont définitivement acquises et distribuées par le contrat au restaurateur et au livreur. Dans le cas contraire, le client peut faire appel du contrat de livraison, en prouvant soit que (1) le repas lui a été livré en retard à l'aide de l'horodatage du livreur ou (2) le contenu de la commande est arrivé froid comme en témoigne la QR code affiché dans la boite témoin. Une fois transmise l'une ou l'autre des informations à la procédure d'arbitrage, celle-ci pourra soit annuler toutes les transactions de valeurs pour cette commande, soit appliquer des pénalités au restaurateur ou au livreur.

Discussion

Le modèle que nous proposons permet de mieux cadrer deux cas où la validation d'un Smart Contrat dans la blockchain réduirait la confiance globale dans le système.

Premièrement, dans le cas où l'exécution du contrat dans le monde réel n'est pas conforme à la réalité des engagement pris par les parties prenantes d'un contrat. Dans ce cas, la procédure d'arbitrage doit statuer sur la violation *d'invariants exogènes* basé sur des oracles constitués de données hors-chaînes

et simple des transactions ou l'application de pénalités, amendant le transfert de valeur initial.

(c'est à dire non présentes dans la blockchain). Ces oracles visent à vérifier le bon déroulé des engagements réciproques pris par tous les acteurs. En découplant la gestion nominale (transfert de valeur) et la gestion conflictuelle supportée par un mécanisme d'arbitrage, nos propositions facilitent la lisibilité et la compréhension de l'exécution des contrats. A noter que l'arbitrage ne nécessite pas d'être entièrement automatisé, et peut supporter l'utilisation d'un oracle supplémentaire post-exécution, comme par exemple un expert humain.

Un autre aspect prometteur de notre approche et également de pouvoir placer des invariants sur les entrées du contrat, mais également sur l'état au sens large de la blockchain lors du démarrage de l'exécution du contrat (c'est à dire non limité aux seuls états des parties prenantes). De tels *invariants endogènes* pourraient porter par exemple sur le blocage d'un transfert de valeur trop importante (e.g. invariant de valeur absolue des transactions en jeux) ou l'exécution trop répétée d'un contrat (e.g. invariant de seuil de fréquence d'appel d'un contrat). Dans ce cas, la procédure d'arbitrage pourrait être de demander à d'autres parties prenantes de s'assurer de la validité des transactions en jeux *avant* que celles-ci deviennent opposables, c'est à dire avant d'être contraint de violer l'immutabilité de la blockchain. Dans le cas de TheDAO, la communauté aurait pu valider (ou invalider) le transfert massif de fond vers des comptes frauduleux, comme ce qui a été fait a posteriori avec le hard-fork.

Dans tous les cas, la nécessité de l'utilisation d'un procédure d'arbitrage doit relever de l'exception et non du fonctionnement nominal du système. En effet, le positionnement d'invariants trop restrictifs se déclenchant trop fréquemment restreindrait le débit des transactions publiées sur la blockchain en la surchargeant de contrats d'arbitrage et empêcherait l'utilisation de ressources placées dans des procédures d'appel faisant fonction de dépôt fiduciaire.

Notons également que cette approche peut être également couplée à des procédures de vérifications formelles au niveau du contrat, de la procédure d'arbitrage, ou de la procédure de vérification des invariants.

Code by Law

Les auteurs qui s'intéressent aux nouvelles technologies et à leurs relations avec le droit s'opposent généralement sur la question de savoir si le recours aux principes traditionnels des systèmes juridiques sont suffisants ou s'il convient de reconnaître que ces principes sont inadaptés au cyberespace (Raskin 2016). Toutefois, il semble que la majorité s'accorde pour reconnaître qu'il n'y a pas, et ne peut y avoir, d'indifférence du droit à l'égard de ces nouvelles formes d'interactions sociales que véhicule le numérique. Le droit contient suffisamment de principes largement formulés, de raisonnements alambiqués, pour se saisir de tout nouvel objet (sur la possible qualification d'actifs financiers ou de " commotidies " au regard de la législation américaine : (Hansen and Reyes 2017)). Le problème réside davantage dans l'adéquation du traitement juridique ainsi déduit. A cet égard, les Smart Contracts ne font absolument pas exception, ce qui pose donc la question de leur encadrement légal – soit par l'application de règles existantes, soit par l'invention de solutions adaptées à leurs spécificités.

Réfutation de l'hypothèse du non-droit

Le système juridique repose sur une loi fondamentale : celle de sa complétude. Au risque de faire du "panjuridisme", et à l'image de la nature qui a horreur du vide, le juriste, le juge, voir le fonctionnaire des finances publiques trouvera toujours des éléments de droit à appliquer à une situation de fait - fût-ce au prix d'une réflexion intense ou du recours à beaucoup d'imagination. Le droit n'est-il pas la plus puissante école de l'imagination nous disait Giraudaux ? Pour ne prendre qu'un exemple de cette réalité expansive et inclusive, des concepts tel que celui de lex numerica ou de lex informatica sont ainsi venus accompagner le développement du commerce électronique. Cette capacité inventive du droit, même dans des domaines

où l'Etat est a priori peu présent, va directement à l'encontre du sentiment de liberté, voire d'impunité, qui prédomine bien souvent chez les promoteurs du cyberespace (même l'existence d'enjeux économiques peut conduire, à l'inverse, les acteurs du numérique à appeler de leurs vœux une réglementation qui sera source de la confiance indispensable au développement de leur secteur d'activité).

Les promoteurs de the DAO le définissent ainsi comme un "Paradis libertaire", échappant à tout cadre juridique… dès lors que la communauté formée n'aurait pas formellement d'existence. Non seulement les initiatives juridiques proposant de reconnaître à ces communautés la personnalité juridique et donc une forme de responsabilité se sont multipliées, mais en outre le droit existant a trouvé à s'appliquer à travers la qualification de " titres financiers " des actifs acquis par les participants à the DAO (Pailler 2018) (Vamparys 2018). De manière générale, le rapport entre la technologie numérique et le monde juridique est bien souvent biaisé dans les analyses proposées par les promoteurs des Smart Contracts.

A titre d'exemple, Nick Szabo qui est un des précurseurs de cette technologie, considère que le contrat est un texte qui implicitement demande à un juge d'ordonner à une partie d'effectuer un paiement ou de livrer une prestation à l'autre sous certaines conditions. C'est une vision franchement très réductrice du contrat : elle prend l'hypothèse problématique de l'inexécution pour en faire l'essence du contrat. Pour reprendre les termes de Jean Carbonnier, elle n'envisage la vie du contrat que par la pathologie qui est susceptible de l'atteindre. Or, le Contrat a pour finalité première, de savoir à quoi l'on s'engage, que ce soit par écrit ou non, de pouvoir déterminer la teneur de son engagement et les modalités de sa mise en œuvre. Ensuite, cette vision contentieuse du contrat simplifie à l'extrême en prenant un cas très particulier, l'absence d'exécution, comme seule hypothèse de saisine du juge. Bien souvent, l'exécution n'est pas quelque chose de binaire mais un continuum et tout l'enjeu en cas de litige sera d'interpréter le contrat afin de savoir à quoi s'étaient engagées les parties.

Pour en revenir aux Smart Contracts, leurs promoteurs font une double erreur : ils pensent d'une part que l'auto-exécution évite toute possibilité de contentieux (Mik 2017), et d'autre part, que l'absence de contentieux est l'indicateur de l'a-juridicité de la situation " smart-contractuelle ". Ce qui est vrai, par contre, c'est que l'identification de la juridiction compétente, l'identification de la loi applicable, mais aussi celle de la qualification qu'elle implique, restent des questions épineuses et certainement source d'incertitudes pratiques. Si l'on peut alors parler de "non-droit", au sens d'" une baisse plus ou moins considérable de la pression juridique " (Carbonnier), cette baisse n'est qu'un phénomène transitoire et non pas un état stable comme le montre s'agissant des crypto-monnaies l'encadrement juridique croissant partout dans le monde (Blemus 2017).

Adaptation du droit

Même si l'hypothèse du vide juridique doit être écartée, la sécurité juridique, qui est la finalité de toute construction juridique, pourrait imposer une intervention en amont et en aval du déploiement d'un tel contrat automatisé et doté d'un mécanisme d'arbitrage recourant à un oracle. Les Smart Contracts suivraient ainsi le précédant causé par le développement du commerce électronique (dont certaines des adaptations qu'il a entraînées paraissent d'ailleurs utiles pour assurer l'exécution de Smart Contracts : (Cohen 2017)).

Au préalable, il faut évoquer une problématique importante, relative à la détermination du droit applicable. En principe, le droit d'un État voit son champ d'application subordonné à la satisfaction soit d'un critère personnel, soit d'un critère territorial. Pour des opérations " déterritorialisées ", comme celles qui prennent place au sein du cyberespace et qui mettent en relation des personnes susceptibles d'être de nationalités différentes, se pose donc une question quant au lien de rattachement avec le droit d'un État (Vauplane 2017). La question est a priori complexe, et peut difficilement être laissée à la seule appréci-

ation des parties. Face à la diversité des règles nationales, définies soit par le droit des affaires, soit par le droit de la consommation selon la qualité des différents cocontractants, le droit international privé offre les outils susceptibles de répondre à cette problématique – fût-ce par la conclusion de conventions internationales afin d'établir des solutions nouvelles et adaptées aux nouvelles technologies. On supposera par la suite que les problèmes de conflits de loi et de juridiction sont résolus ou du moins susceptibles de résolution. Et l'on supposera donc que la compétence des autorités nationales pour réguler en amont et en aval les Smart Contracts est établie.

En amont

Quatre points essentiels devraient justifier une intervention normative destinée à encadrer en amont la mise en place de Smart Contracts dotés de procédure d'arbitrage.

Il s'agit premièrement de la question du niveau de preuve de la fiabilité du système proposé. On peut penser que les Smart Contracts auront vocation à mettre en relation des parties entre lesquelles l'équilibre ne sera pas parfait, et qu'il sera parfois nécessaire de protéger celle qui se trouve en situation de dépendance. Asymétrie d'autant plus à craindre que le langage informatique du code est proprement incompréhensible pour le quidam, même s'il est supposément moins ambigu dans sa formulation (Mik 2017). Il faut donc non seulement réfléchir au niveau d'information requis et aux modalités de preuve qui doivent s'imposer à celui qui propose un Smart Contract (et par exemple s'interroger sur le recours à la validation formelle) dans l'optique d'assurer un consentement éclairé et non vicié (Raskin 2016) ; mais il faut aussi s'interroger sur l'existence possible d'une autorité compétente pour apprécier les informations fournies quant à la fiabilité du système et à la sincérité de l'information transmise, lorsqu'il est envisagé de proposer un certain type de Smart Contract aux particuliers.

Les mécanismes de contrôle prudentiel confiés à l'autorité des marchés financiers, qui veille à la protection de l'épargne et à l'information des investisseurs, ou encore le Mécanisme de surveillance unique mis en place par l'Union européenne depuis la crise financière de 2008 et qui repose sur une mise en réseau de la Banque Centrale Européenne et des autorités nationales pourrait servir d'exemple pour un contrôle permanent et répressif. Il est aussi possible de chercher à s'inspirer des mécanismes d'autorisations de mise sur le marché, délivrées dans le secteur du médicament ou des variétés végétales par exemple, afin d'envisager une procédure préalable d'autorisation ou d'agrément applicable aux Smart Contracts.

Le deuxième point concerne la liberté ou du moins l'égalité des parties dans le choix de recourir à un oracle. A cet égard, si l'on pense contrat d'adhésion et asymétrie du pouvoir contractuel, il apparaît évident que la décision sera souvent imposée aux particuliers cocontractants, dans la mesure où elle sera indissociable du contrat proposé. Comme en matière d'arbitrage, il ne serait donc pas inutile de réfléchir aux conditions qui permettent d'éviter l'imposition d'un mécanisme d'appel dont les caractéristiques – en particulier l'identité de l'arbitre, le choix des sources d'informations utilisées, et les procédés de décision – seront prédéterminées et potentiellement abusives. Il serait en ce sens cohérent avec les limites posées en matière d'arbitrage (“ arbitrabilité objective ”et“ arbitrabilité subjective ”) de prévoir des cas d'interdiction dans les cas relevant de l'ordre public, des lois de police ou de droits indisponibles. Là encore, une exigence d'information sur les droits et obligations des différentes parties pourrait s'avérer judicieuse.

Une troisième préoccupation tient à l'encadrement et au statut de l'information utilisée cette fois-ci au sein du mécanisme d'appel ou d'arbitrage. Lors du contrôle de l'exécution de l'obligation contractuelle, l'information fournie par l'oracle est primordiale. Pour le juriste, le constat que les faits sont parfois prépondérants dans l'application de la règle de droit n'est pas une découverte. C'est tout l'enjeu de l'appréciation des faits, puis de leur qualification juridique. Il faut donc s'interroger sur la nécessité et la possibilité de garantir la neutralité de l'information

utilisée au stade du contrôle de l'exécution, c'est-à-dire l'objectivité des données transmises.

Quatrièmement, le droit devrait s'intéresser *a fortiori* au statut de l'arbitre. D'abord et évidemment, il s'agit comme pour l'oracle de s'intéresser à son objectivité et à son indépendance vis-à-vis des parties. A nouveau, comme en matière de clauses compromissoires prévoyant le recours à l'arbitrage, la loi devrait interdire les clauses abusives, susceptibles par exemple le fait de confier l'examen de l'appel à une partie prenante au contrat.

Cette préoccupation peut d'ailleurs rejoindre la précédente. Ainsi, dans le cas d'un billet de train vendu par l'EPCI SNCF, l'État en tant qu'actionnaire unique n'aurait-il pas intérêt à ce que l'EPA Météo France soit généreux avec ses deniers quand il faudra déterminer si la non-exécution de l'obligation d'acheminement peut être justifiée par le caractère exceptionnel des conditions météorologiques ? Comme en matière d'information préalable à l'engagement contractuel, le recours à des autorités administratives indépendantes pourrait offrir des perspectives intéressantes, à moins de se satisfaire d'un contrôle ex-post, de nature judiciaire, en cas de contentieux persistant après l'inscription dans la blockchain de l'opération contractuelle.

Ensuite, il n'est pas inutile de se demander si l'arbitre ne doit pas être soumis à un certain nombre d'obligations déontologiques. Là aussi, la réglementation juridique existante en fournit des illustrations. Ainsi, dans le cas d'un contrat d'assurance vie, l'arbitre serait amené à manipuler des données personnelles couvertes en France par le secret médical, il faudrait donc en assurer le respect, ce qui peut être facilement obtenu en imposant le recours à un médecin-conseil comme c'est le cas pour l'assurance traditionnelle.

En aval

Si l'intervention ex ante de la règle de droit est de nature à éviter un certain nombre de complications, il n'empêche que l'exécution du Smart Contract, toute automatique qu'elle soit, peut être source de contestations. La question qui convient alors nécessairement d'envisager, en particulier dans le cas de contrats dotés de procédure d'arbitrage, c'est celle de leur force exécutoire (en cas d'une exécution impossible – par exemple, si le débiteur n'a plus de crédits disponibles sur la blockchain) ou de la contestation de cette dernière (en cas de contestation de l'exécution automatique et/ou de la décision prise sur appel par l'arbitre).

L'une des solutions avancées par les auteurs qui se sont penchés sur ce point, consiste à adosser au Smart Contract un contrat existant dans le monde réel, afin de lui appliquer en cas de besoin les voies légales prévues en matière d'exécution. A l'analyse, ce détour paraît tout à fait inutile et même contre-productif. Il revient à subordonner l'efficacité accrue du Smart Contract, en raison de l'automaticité de l'exécution, à l'absence de contentieux, alors que celle-ci a justement vocation à éviter celui-ci. Ensuite, cette proposition présuppose que le Smart Contract n'est pas susceptible d'être légalement reconnu comme un engagement contractuel, en raison de sa forme électronique notamment ; or un tel raisonnement repose sur une conception du droit laissant une part tout à fait excessive au formalisme (Raskin 2016) p321. Cela étant, l'analyse du Smart Contract comme un engagement contractuel ne résout pas toute difficulté que peut soulever la question de son exécution.

Si l'hypothèse de l'exécution impossible (pour des raisons techniques ou liées à la disposition des fonds nécessaires sur la blockchain par exemple) se résout sans difficultés particulières devant le juge du contrat, la situation dans laquelle c'est l'exécution qui est source de contentieux est davantage source d'interrogations.

En premier lieu, lorsqu'elle a été automatiquement assurée, l'exécution peut toujours être soumise à une contestation judiciaire (ce contentieux contractuel aurait toutefois pour particularité d'intervenir ex post, alors que le juge du contrat est usuellement saisi pour assurer le respect de l'obligation contractuelle et son exécution en nature, ou pour constater l'existence

d'un titre exécutoire). Tel sera le cas si la cause du contrat est illicite et contraire à l'ordre public, mais aussi en cas de divergence d'interprétation des obligations par les parties, ou entre les expectatives et les résultats produits par le Smart Contract.

Les problèmes juridiques pouvant surgir à ce stade sont nombreux : quel effet faut-il attacher à l'absence d'utilisation par l'une des parties de la possibilité de faire appel à l'arbitre avant inscription de l'opération dans la blockchain ? Faut-il voir, comme en matière d'arbitrage, dans cette passivité une renonciation rendant une saisine ultérieure du juge irrecevable ?

En second lieu, il faut également s'interroger sur l'effet du mécanisme d'appel mis en œuvre préalablement à l'inscription sur la blockchain du résultat du Smart Contract. Faut-il reconnaître à l'arbitrage intégré au Smart Contract l'autorité de chose jugée entre les parties ? Et si oui, quelles sont alors les conditions qui entourent l'hypothèse d'un appel devant l'autorité judiciaire compétente ? Quelle doit être l'étendue de la compétence du juge ? Convient-il de la limiter au respect de l'ordre public, ou de l'étendre aux constatations factuelles tirées de l'oracle ?

Enfin, dans différentes situations envisagées précédemment, au-delà de la reconnaissance par le droit du Smart Contract et de son caractère obligatoire, l'efficacité du mécanisme peut nécessiter la reconnaissance de sa force exécutoire. Ainsi, en cas de besoin et afin de transformer le en titre exécutoire, le droit devra déterminer si le Smart Contract possède en soi, comme l'acte authentique établi par un notaire, une force exécutoire, ou s'il est au contraire nécessaire de recourir à une procédure d'exequatur au préalable ?

Ouverture

Toutes ces questions sont ouvertes, car leurs réponses dépendront du niveau de protection que le droit souhaitera accorder aux utilisateurs des Smart Contracts. Les problèmes posés en aval de l'exécution témoignent en tout cas de l'importance pour les systèmes juridiques de se prononcer sur le régime juridique qui doit accompagner leur développent.

Certes, dans le cadre du commerce électronique international, on a pu observer la naissance spontanée d'une *lex electronica*, constituée de normes informelles et de codes de bonnes pratiques adoptés par les acteurs du secteur, généralement regroupés en organisations transnationales, plutôt que par les autorités étatiques. Cependant la nécessité de protéger le consommateur face à des professionnels a partout fait naître des réglementations spécifiques complémentaires, destinées à accompagner le développement d'internet et à insuffler la confiance nécessaire pour son développement. Le parallèle entre ce précédent et le cas des Smart Contracts n'est pas dépourvu de pertinence. On observe en effet, à travers des initiatives comme celle d'Hyperledger ou de clause.io, une volonté des acteurs du milieu de coopérer pour établir des standards. Cependant, il y a dans ces démarches des mobiles qui, sans être nécessairement incompatibles avec l'objectif de protection des utilisateurs, tiennent davantage d'une stratégie d'occupation. Cela étant, la diffusion, bien souvent en opensource, de standards techniques peut favoriser la sécurité juridique, dans la mesure où leurs promoteurs sont bien souvent soucieux d'intégrer des principes juridiques largement acceptés (Hansen and Reyes 2017). En accompagnant ce mouvement, le droit pourra ainsi éviter qu'il revienne aux seuls acteurs dominants du secteur de déterminer à travers leurs technologies les valeurs qui doivent régir le fonctionnement des Smart Contracts.

Références

Androulaki, Elli, Artem Barger, Vita Bortnikov, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart, et al. 2018. "Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains." In *Proceedings of the Thirteenth Eurosys Conference*, 30. ACM.

Atzei, Nicola, Massimo Bartoletti, and Tiziana Cimoli. 2017. "A Survey of Attacks on Ethereum

Smart Contracts (Sok)." In *Principles of Security and Trust*, 164–86. Springer.

Bhargavan, Karthikeyan, Antoine Delignat-Lavaud, Cédric Fournet, Anitha Gollamudi, Georges Gonthier, Nadim Kobeissi, Natalia Kulatova, et al. 2016. "Formal Verification of Smart Contracts: Short Paper." In *Proceedings of the 2016 Acm Workshop on Programming Languages and Analysis for Security*, 91–96. ACM.

Blemus, Stéphane. 2017. "Law and Blockchain: A Legal Perspective on Current Regulatory Trends Worldwide." *Revue Trimestrielle de Droit Financier*.

Bozic, Nikola, Guy Pujolle, and Stefano Secci. 2016. "A Tutorial on Blockchain and Applications to Secure Network Control-Planes." In *Smart Cloud Networks & Systems (Scns)*, 1–8. IEEE.

Breidenbach, Lorenz, IC Cornell Tech, Philip Daian, Florian Tramer, and Ari Juels. 2018. "Enter the Hydra: Towards Principled Bug Bounties and Exploit-Resistant Smart Contracts." In *27th {Usenix} Security Symposium ({Usenix} Security 18)*. USENIX} Association.

Cohen, Travis Parker, Alen West. 2017. "The Enforceability of Smart Contracts." *In Georgetown Law Technology Review* 1 (2).

Dannen, Chris. 2017. *Introducing Ethereum and Solidity*. Springer.

Grzegorczyk, Christophe. 2002. "Ordre Juridique Comme Réalité." *Droits*, no. 1: 103–18.

Hansen, J Dax, and Carla L Reyes. 2017. "Legal Aspects of Smart Contract Applications: Digital Asset Sales and Capital Markets." *Supply Chain Management, Land Registries, Government Records and Smart Cities, and Self-Sovereign Identity/Perkins Coie LLP//*.

Herbaut, Nicolas, and Daniel Négru. 2017. "A Model for Collaborative Blockchain-Based Video Delivery Relying on Advanced Network Services Chains." *IEEE Communications Magazine* 55.

Hirai, Yoichi. 2017. "Defining the Ethereum Virtual Machine for Interactive Theorem Provers." In *International Conference on Financial Cryptography and Data Security*, 520–35. Springer.

Jentzsch, Christoph. 2016. "Decentralized Autonomous Organization to Automate Governance." *White Paper, November*.

Kiayias, Aggelos, Alexander Russell, Bernardo David, and Roman Oliynykov. 2017. "Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol." In *Annual International Cryptology Conference*, 357–88. Springer.

Lessig, Lawrence. 1999. "Code Is Law." *The Industry Standard* 18.

Mik, Eliza. 2017. "Smart Contracts: Terminology, Technical Limitations and Real World Complexity." *Law, Innovation and Technology* 9 (2): 269–300.

Nakamoto, Satoshi. 2008. "Bitcoin: A Peer-to-Peer Electronic Cash System."

Pailler, Pauline. 2018. "Quel Encadrement Pertinent Pour Les Initial Coin Offerings ?" *In Revue Internationale Des Services Financiers*.

Raskin, Max. 2016. "The Law and Legality of Smart Contracts."

Szabo, Nick. 1997. "The Idea of Smart Contracts." *Nick Szabo's Papers and Concise Tutorials* 6.

Vamparys, Xavier. 2018. "Perspectives Américaines Sur La Régulation Des Crypto-Actifs." *Bulletin Joly Bourse*.

Vauplane, Hubert. 2017. "Blockchain and Conflict of Laws." *Revue Trimestrielle de Droit Financier* 4 (50).

Veronese, Giuliana Santos, Miguel Correia, Alysson Neves Bessani, Lau Cheuk Lung, and Paulo Verissimo. 2013. "Efficient Byzantine Fault-Tolerance." *IEEE Transactions on Computers* 62 (1): 16–30.

Xu, Xiwei, Cesare Pautasso, Liming Zhu, Vincent Gramoli, Alexander Ponomarev, An Binh Tran, and Shiping Chen. 2016. "The Blockchain as a Software Connector." In *Software Architecture (Wicsa), 2016 13th Working Ieee/Ifip Conference on*, 182–91. IEEE.

Thème B

« Systèmes autonomes et décision, droits fondamentaux »

Aide à la décision en matières médicale et judiciaire: quelle certification et quelles explications pour les algorithmes?

Sonia Desmoulin-Canselier[1]

Daniel Le Métayer[2]

Introduction

Depuis près de quarante ans, le droit relatif au traitement des données personnelles offre des ressources pour limiter l'automatisation des décisions adoptées à l'encontre des personnes. Dès 1978, les articles 2 et 3 de la loi Informatique et Liberté prévoyaient une prohibition de principe des décisions de justice automatisées et des décisions administratives ou privées prises sur le seul fondement d'un traitement automatisé de données portant sur le profil ou la personnalité de l'intéressé. Un droit à l'information et à la contestation sur « les informations et les raisonnements utilisés dans les traitements automatisés » était ainsi accordé à la personne concernée. La modification du texte, en particulier en 2004, n'a pas eu d'incidence significative sur ce droit : une « décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé de données destiné à définir le profil de l'intéressé ou à évaluer certains aspects de sa personnalité ». Le droit européen est venu conforter ces prévisions nationales, d'abord avec la Directive 95/46/CE du 24 octobre 1995[3] puis avec le Règlement 2016/679 du 27 avril 2016 (RGDP)[4]. Juristes et citoyens peuvent donc depuis longtemps trouver dans le droit positif des réponses apaisantes aux craintes suscitées par la perspective d'être un jour jugés ou gouvernés par des robots autonomes, ou soumis aux décisions d'algorithmes opaques. Pourtant, l'inquiétude n'a jamais été aussi forte. On n'a jamais autant parlé de « transparence » et de « loyauté » des algorithmes qu'aujourd'hui. Le droit offre-t-il un rempart véritablement protecteur contre la tentation de l'automatisation des décisions ? Outre que le droit européen est moins exigeant que le droit français, puisqu'il réserve l'exception du consentement explicite de la personne concernée là où la loi française cantonne cette exception au domaine contractuel, il faut bien constater que les textes évoqués ne visent que les traitements automatisés directement décisionnels. L'intervention humaine permet ainsi d'écarter les protections légales concernant les décisions fondées exclusivement sur un traitement automatisé. Or, il est aisé de prétendre limiter l'usage de l'outil algorithmique à un rôle d'aide à la décision. Une telle revendication, moins ambitieuse en apparence, est tout à fait compatible avec le développement d'un marché rentable : les producteurs peuvent espérer assumer moins de responsabilités en circonscrivant l'usage attendu du logiciel, et optimiser leurs chances de convaincre des utilisateurs rassurés de demeurer indispensables dans le processus décisionnel.

[1] CNRS, Université de Nantes, UMR 6297 DCS.
[2] Inria, Université de Lyon.
[3] Directive 95/46/CE 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, article 15.
[4] Règlement (UE) 2016/679 du Parlement Européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, article 22.

De fait, on constate depuis quelques années un accroissement de l'offre d'outils informatiques d'aide à la décision dans des secteurs clés comme ceux des activités médicales et judiciaires. Il s'agit, on le perçoit immédiatement, de domaines où les enjeux humains sont particulièrement importants, les décisions prises par les médecins et par les juges pouvant affecter significativement la vie des personnes. Ce sont également des secteurs où l'incertitude est difficile à accepter et où il est tentant de placer ses espoirs dans les algorithmes pour améliorer autant que possible la justesse des décisions. L'importance conférée aux données dans les deux domaines - état de l'art médical d'un côté, état du droit textuel et jurisprudentiel de l'autre – explique que des projets de conception de « systèmes experts » existent depuis plus de quarante ans[5]. L'objectif consistant à compléter les informations disponibles pour que le médecin, le personnel soignant, le juge et l'avocat puissent choisir la voie la plus appropriée est des plus légitimes et toute aide technique en ce sens est bienvenue. Le souhait de réduire la part de subjectivité pour limiter les risques de biais décisionnels irrationnels et lutter contre les inégalités territoriales (entre juridictions ou entre établissements hospitaliers) peut aussi motiver un recours systématisé aux outils d'aide à la décision, conçus comme des vecteurs d'harmonisation. Les systèmes d'aide à la décision disponibles sont très nombreux[6] et offrent des fonctionnalités variées allant de l'optimisation de l'information du décideur (par la présentation des informations, l'accélération de la recherche de données ou l'augmentation des masses de données traitées) à des suggestions de décision. Ils coexistent dans leur variété et, contrairement à une idée reçue, les dernières innovations en intelligence artificielle ne remplacent pas toujours des programmes informatiques de type « systèmes experts ».

Cependant, l'intervention d'une personne humaine, responsable de la décision et en charge de la motiver, suffit-elle à lever toutes les inquiétudes ? Qu'en est-il de la justification d'une décision humaine lorsqu'elle s'appuie sur un outil de traitement de données trop complexe pour que l'utilisateur puisse rendre compte des raisons pour lesquelles une décision a été rendue ? Dans des domaines aussi sensibles que la santé et la justice, peut-on accepter d'utiliser des outils dont on ne saurait ni détecter les biais, ni expliquer les résultats, si précis soient-ils ? Au-delà de la protection de la vie privée des personnes concernées et de l'utilisation qui peut être faite de données personnelles par ces outils algorithmiques se pose une question plus large et plus fondamentale : comment faire en sorte que les utilisateurs que sont les médecins et les juges (ou plus largement tout décisionnaire) fassent un usage pertinent des outils algorithmiques d'aide à la décision ? Pour répondre à ces questions, les textes relatifs à la protection de la vie privée et des données personnelles ne suffisent plus, même s'ils offrent quelques ressources. Nous suggérons une approche inclusive résumée par la notion de « redevabilité », inspirée de la notion anglo-saxonne d'*accountability*[7]. Elle permet d'insister sur le fait que le système algorithmique doit « rendre des comptes » sur ce qu'il fait, en s'en expliquant, mais aussi en permettant de vérifier s'il respecte ou non certaines exigences. Cette approche permet d'envisager une autre forme de protection des personnes concernées par la décision judiciaire ou médicale. Une protection indirecte mais décisive découlerait en effet de l'existence de dispositions juridiques visant à garantir de bonnes conditions d'utilisation des outils d'aide à la

[5] *Cf.* pour le droit : C. Thomasset et J. Vanderlinden, « Cantate à deux voix sur le thème « Une révolution informatique en droit ? », *Revue Trimestrielle de droit civil* 1998, p. 315 ; D. Bourcier, « À propos des fondements épistémologiques d'une science du droit », *in* Y. Aguila, *Quelles perspectives pour la recherche juridique ?*, Presses Universitaires de France « Droit et justice », 2007, p. 69-74.

[6] V. par ex. Le Cahier du « Monde », n° 22494 daté du mercredi 10 mai 2017.

[7] N. Diakopoulos, Accountability in algorithmic decision making, Communications of the ACM, Vol. 59, No 2, 2016.

décision. L'exigence de « redevabilité » que nous prônons recouvre plusieurs composantes complémentaires. La première concerne l'absence de biais pouvant aboutir à des discriminations et amène à envisager une obligation de certification (1). La seconde concerne le besoin d'explication et conduit à affirmer l'exigence d'intelligibilité (2).

1. Certifier l'absence de biais dans les traitements algorithmiques

Les textes européens et nationaux en matière de protection des données personnelles affirment que les données à caractère personnel doivent être traitées de manière « licite » et « loyale »[8]. Cette formulation amène à distinguer la « licéité » du traitement de la « loyauté », sans qu'il soit aisé de déterminer ce que recouvre exactement ce dernier terme : uniquement le respect des prévisions et des attentes contractuelles ou également le respect de certaines normes morales. Pour ce qui concerne les outils algorithmiques d'aide à la décision, il nous semble que la première exigence que l'on peut légitimement formuler est la possibilité de vérifier l'absence de biais problématiques.

Ce besoin de vérification pourrait se traduire juridiquement par une obligation de certification, définie comme l'obligation de démontrer, par exemple à une autorité de certification ou à un auditeur, que le système algorithmique satisfait un certain nombre de critères. Ces critères peuvent typiquement inclure la précision (pertinence des résultats), l'absence de discrimination, ou encore l'équité. Nous pensons que, s'agissant de domaines sensibles comme la santé ou la justice, il est impératif que des vérifications relatives au bon fonctionnement des systèmes d'aide à la décision et à leur conformité à ces critères puissent être effectuées, en amont et en aval de leur mise sur le marché. Cette perspective nous paraît réalisable, sur les plans techniques et juridiques.

Sur le plan technique, même si la question générale de la définition des critères à vérifier et des moyens d'opérer cette vérification est très complexe et donne lieu à des travaux de recherche, certaines solutions existent d'ores et déjà. A titre d'exemple, plusieurs définitions formelles de discrimination ont été proposées et des techniques ont été développées pour les évaluer[9]. Il est important de prendre en compte non seulement les discriminations directes mais aussi les discriminations indirectes, c'est à dire les traitements qui n'exploitent pas directement des attributs prohibés (origine ethnique par exemple) mais utilisent des données corrélées à ceux-ci (adresse du lieu de résidence par exemple). Des stratégies ont également été proposées pour assurer, par construction, qu'un algorithme respecte les critères de non-discrimination[10]. Elles reposent sur des traitements des données d'apprentissage (qui peuvent être elles-mêmes biaisées), des résultats (correction des biais éventuels) ou des algorithmes.

[8] RGPD, article 5 ; Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, plusieurs fois modifiée, article 6, 1.

[9] Andrea Romei and Salvatore Ruggieri. A multidisciplinary survey on discrimination analysis. The Knowledge Engineering Review, 29(5):582–638, November 2014. A. Datta, M. C. Tschantz, A. Datta, Automated Experiments on ad privacy settings, Proceedings of Privacy Enhancing Technologies (PETS), 2015.

[10] Dino Pedreschi, Salvatore Ruggieri, and Franco Turini. Discrimination-aware data mining. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery. Data Mining , KDD '08, pages 560–568, New York, NY, USA, 2008. ACM. Toon Calders and Sicco Verwer. Three naive bayes approaches for discrimination-free classification. Data Mining and Knowledge Discovery , 21(2):277–292, 2010.

Sur le plan juridique, il existe déjà des réglementations conditionnant l'accès au marché européen au respect d'une exigence de certification, sur lesquelles nous suggérons de prendre appui. Dans le domaine médical, le Règlement (UE) 2017/745 du 5 avril 2017[11], qui remplace la directive 93/42/CEE[12], exige une certification des dispositifs médicaux. Or, ce texte reconnaît désormais explicitement que les logiciels d'aide à la décision sont des dispositifs médicaux. Dès lors, il devient possible de s'appuyer sur ces dispositions et sur l'objectif européen de protection des personnes. Si, en l'état de rédaction, aucune disposition ne vient expressément formuler une obligation de vérification de l'absence de biais telle que nous la formulons, il est à la fois possible (en mobilisant les ressources offertes par l'interprétation dans un premier temps) et souhaitable (en suggérant une modification textuelle à moyen terme) de suggérer une intégration de cette problématique dans ce corps de réglementations. Certes, des difficultés pourraient se poser (tenant notamment à la variété des logiciels et aux interprétations antérieures des autorités nationales). Elles n'en rendent que plus nécessaire la prise de conscience de l'urgence d'intégrer la vérification de l'absence de biais dans les finalités du contrôle de la certification. Le droit du dispositif médical renforcerait ainsi opportunément les efforts faits en matière de traitement des données personnelles[13], tout en élargissant l'éventail des personnes bénéficiaires de cette exigence.

En matière judiciaire, il n'existe pas d'équivalent au système de certification des dispositifs médicaux, mais il existe des règles juridiques visant à assurer la fiabilité et la qualité des résultats d'expertise en matière d'analyses génétiques. Elles ne sont pas applicables aux logiciels d'aide à la décision judiciaire, mais elles illustrent le type de réglementation qui pourrait être mis en place. A défaut de règle spécifique, le principe de la libre appréciation du juge dans la détermination des sources d'information utiles au dévoilement de la vérité et au règlement du litige s'applique. A ce jour, il existe par exemple une variété de barèmes et de nomenclatures pour les chefs de préjudice. Leur usage est à la discrétion du juge et n'est donc ni obligatoire, ni harmonisé, ni soumis à vérification. L'idée de diffuser plus largement les barèmes, notamment sous la forme de référentiels, est un projet en cours depuis longtemps[14]. Le développement de logiciels utilisant des algorithmes de traitement de contenu pourrait rendre ces propositions de barèmes obsolètes, même si les premiers essais dans des cours d'appel françaises d'outils algorithmiques de traitement de contenu n'ont pas donné un résultat jugé satisfaisant. La diversité des offres sur le marché laisse, en effet, deviner un avenir prometteur. Or, un projet de systématisation du recours à ces outils, même utilisés à titre de référentiels, nous semble devoir être accompagné d'une réflexion sur le terrain de leur validation, voire de leur certification. Il serait paradoxal de préconiser l'utilisation de systèmes d'aide à la décision en vue de lutter contre les disparités judiciaires et de ne pas vérifier que lesdits systèmes ne contiennent pas des biais (visibles ou invisibles) et ne produisent pas des discriminations.

[11] Règlement (UE) 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux, modifiant la directive 2001/83/CE, le règlement (CE) n° 178/2002 et le règlement (CE) n° 1223/2009 et abrogeant les directives du Conseil 90/385/CEE et 93/42/CEE (*JO* L 117, 5 mai 2017, p. 1).
[12] Directive 93/42/CEE du Conseil du 14 juin 1993 relative aux dispositifs médicaux (*JO* L 169 du 12 juillet 1993, p. 1).
[13] RGDP, considérant 100.
[14] Voir par exemple la Réponse ministérielle du ministre de la Justice en date du 20 juin 2006 (J0 du 20/6/2006).

2. Rendre intelligibles les résultats de traitements algorithmiques

Le RGDP affirme, outre le nécessaire respect des principes de loyauté et de licéité, l'exigence de respecter un principe de « transparence » dans le traitement des données à caractère personnel[15]. Ce principe de transparence exigerait que « toute information et communication relatives au traitement de ces données à caractère personnel soient aisément accessibles, faciles à comprendre, et formulées en des termes clairs et simples »[16] et que « toute information adressée au public ou à la personne concernée soit concise, aisément accessible et facile à comprendre, et formulée en des termes clairs et simples et, en outre, lorsqu'il y a lieu, illustrée à l'aide d'éléments visuels ».[17] Bien que ces dispositions alertent sur le besoin de clarté et d'explications, elles présentent deux limites pour notre réflexion. Tout d'abord, il s'agit de faire en sorte que la personne dont les données sont utilisées soit informée, et non d'informer l'utilisateur professionnel. Ce sont les activités de collecte en ligne, par exemple à des fins de publicité ciblée, qui sont ainsi visées. Or, si besoin de « transparence » il y a, cela devrait couvrir un champ beaucoup plus vaste et concerner aussi les utilisateurs, qui peuvent être experts de leur domaine mais pas nécessairement en informatique. De plus, la notion polysémique de « transparence » peut conduire à des malentendus. Elle peut notamment donner à penser qu'une compréhension totale des processus est recherchée ou qu'aucun secret industriel (ou commercial) ne devrait être toléré. Or, une telle exigence ne paraît ni réaliste, ni véritablement souhaitable.

Le terme d'intelligibilité ou d'explicabilité (moins harmonieux) est, pour ces raisons, plus intéressant. Non seulement le concept d'intelligibilité est déjà connu en droit positif (notamment pour formuler les exigences constitutionnelles en termes d'intelligibilité de la norme), mais il explicite immédiatement l'enjeu essentiel. Une obligation d'explication, définie comme un moyen d'assurer une forme d'intelligibilité des systèmes algorithmes, devrait être introduite et déclinée selon les publics visés. Pour le citoyen sans compétence technique particulière, l'explication peut prendre la forme d'une justification d'un résultat (décision) qui le concerne. Un expert pourra être intéressé par des mesures plus globales, comme des explications sous forme d'arbres de décision ou d'autres représentations graphiques mettant en lumière les données prises en compte par l'algorithme et leur influence sur les résultats. L'explication ne se réduit donc pas à la publication du texte d'un algorithme ou du code source d'un logiciel qui peuvent demeurer opaques pour le commun des citoyens (et même pour des experts).

Nombre de recommandations, issues de comité d'éthique[18], d'autorité de régulation[19] ou d'institution européenne[20], ont souligné l'importance de l'intelligibilité des décisions reposant sur des traitements algorithmiques. Pourtant, en dehors des dispositions du droit relatif au traitement des données personnelles[21] qui ne concernent pas les utilisateurs de logiciels d'aide à la décision, aucune exigence spécifique n'est formulée en droit positif. Certes, des fondements

[15] RGDP, article 5.
[16] RGDP, considérant 39.
[17] RGDP, considérant 58.
[18] CERNA, *Ethique de la recherche en apprentissage machine*, juin 2017.
[19] CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, décembre 2017.
[20] Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)).
[21] Notamment: Directive 95/46/CE, art. 39 ; RGDP, art. 13.

juridiques existent : principe juridique et déontologique de responsabilité en matière médicale (décision « en conscience ») ; principes de motivation et de contradiction dans le champ judiciaire. Ces principes nous semblent nécessairement impliquer que des explications soient fournies à l'utilisateur sur le fonctionnement de l'algorithme d'aide à la décision. Néanmoins, il serait plus satisfaisant, même au-delà de ces domaines d'activité, de voir pleinement et clairement reconnaître une exigence juridique générale d'intelligibilité des outils algorithmiques d'aide à la décision, dès lors que leurs résultats peuvent affecter significativement la vie des personnes.

Reste ici un défi de taille : trouver les solutions techniques pour fournir des explications compréhensibles aux différents utilisateurs. Les explications des résultats d'un algorithme peuvent prendre des formes variées selon le domaine concerné, les techniques algorithmiques mises en œuvre (déterministes ou probabilistes, reposant sur l'apprentissage ou non, supervisées ou non, etc.) et le public visé (par exemple professionnel ou particulier). Cependant, il nous semble que ces explications devraient couvrir au moins deux aspects fondamentaux : les données utilisées et la logique conduisant au résultat. Il ne s'agit pas de rendre l'algorithme « transparent » au sens où l'utilisateur deviendrait capable de comprendre l'entier processus. L'objectif devrait plutôt être de fournir suffisamment d'informations sur les données utilisées et les choix opérés lors de leur traitement pour que l'utilisateur puisse porter un regard critique sur le résultat. Selon les techniques utilisées, la logique conduisant à un résultat peut être plus ou moins difficile à expliquer. Un nombre croissant de chercheurs travaille sur l'explication des résultats d'algorithmes dans la mouvance de ce qu'on appelle parfois l'« explainable AI ». Certains de ces travaux portent sur des analyses a posteriori, d'autres visent à concevoir des systèmes algorithmiques produisant, par construction, leurs explications. Si toutefois certaines techniques algorithmiques s'avéraient irréductibles à toute forme d'explication, on peut se demander s'il n'y aurait pas matière à considérer que la balance bénéfice/risque (présente explicitement notamment dans le droit des dispositifs médicaux) serait alors défavorable. On sait notamment que les algorithmes d'intelligence artificielle reposant sur l'apprentissage profond (« deep learning ») ont un fonctionnement particulièrement opaque pour des humains. S'il était impossible de fournir à l'utilisateur les informations nécessaires à l'exercice de son esprit critique, il pourrait s'avérer nécessaire d'écarter ce type d'outils, tout du moins temporairement, dans des domaines sensibles comme la santé et le droit.

Conclusion

En conclusion, il nous paraît important de ne pas focaliser l'attention uniquement sur les décisions prises *exclusivement* sur le fondement d'un traitement automatisé, mais de prendre en considération les exigences de « redevabilité » qui devraient s'appliquer aux outils algorithmiques d'*aide à la décision.* Pas plus que la loyauté, au sens de conformité des résultats d'un algorithme à son fonctionnement annoncé, la notion de transparence des algorithmes, au sens de la mise à disposition publique de leur code, ne saurait suffire à saisir les enjeux essentiels. Le véritable enjeu est plutôt de rendre compte, dans le sens de rendre possible la vérification d'absence de biais et rendre intelligible la logique qui a conduit au résultat suggéré. La notion de redevabilité, qui nous semble donc plus pertinente, pose cependant des questions complexes aussi bien sur le plan juridique que sur le plan technique. Ces questions sont d'une importance cruciale dans des domaines aussi sensibles que ceux de la santé et de la justice.

Julia SOURD
Avocat à la Cour
Docteur en Droit
12 rue du Professeur Demons
33000 BORDEAUX
julia.sourd@avocat-conseil.fr

L'encadrement juridique des traitements automatisés est il adéquat ?
En cas d'engagement de responsabilité civile.

« *Là où croit le péril croît aussi ce qui sauve* », la formule d'Hoderlin peut illustrer l'état du droit quant au traitement automatisé des données. En l'état actuel, il existe une forme de passivité juridique, la vision prospective restant à l'état de colloques, d'incubation. Mais, le droit de la responsabilité (ou plutôt les droits de la responsabilité) peut être suffisamment vivace pour absorber les potentiels contentieux, car les bases de réflexion ici avancées ne concernent que l'hypothèse d'un engagement de responsabilité à la suite d'un dommage corporel. Les autres aspects, notamment de propriété intellectuelle étant traité par d'autres intervenants.

L'objectif du droit dans nos sociétés modernes étant devenu comment le droit atteint un idéal d'indemnisation, gère le risque et ne bride pas la technologie, ces trois éléments semblent s'éloigner à toute allure de la faute pour se concentrer, quant à l'indemnisation du côté de la victime, et du côté du risque et de la technologie, vers l'innovation et le respect nécessaire que le droit doit avoir pour permettre à la société d'innover et de fonctionner.

Le risque doit donc être appréhendé comme une clé de répartition de la dette.

On a pu reprocher au risque de décourager les initiatives et l'esprit d'entreprise, mais c'est une justice de répartition financière *ubi emolumentum ibi onus* (là où est l'avantage, là doit être la charge).

En conséquence, il doit être déterminé ce que l'on fait, comment peut-on concilier un impératif éthique : l'indemnisation de la victime et un impératif sociétal : ne pas brider l'économie.

Du côté de l'indemnisation des victimes, il y aura à régler le problème de déterminer si toutes les victimes sont recevables à obtenir une indemnisation.
A l'intérieur de ces victimes et de ces différentes catégories, la question à se poser sera l'indemnisation doit-elle être intégrale ?

Et du côté de la notion de risque, il doit nécessairement être fait un choix, puisque sinon tout le droit de la responsabilité civile et tous les régimes spéciaux d'indemnisation seraient appelés en renfort d'une modeste tentative de systématisation.

L'un des philosophes de la responsabilité, Hans JONAS[1] mettait un terme à la croyance que dans le progrès tout est forcément positif et indiquait que « *nous sommes en danger permanent d'autodestruction collective* ».
Il disait également « *agis de telle sorte que les effets de ton action soient compatibles avec la permanence d'une vie authentiquement humaine* » (pages 30-31).
Ainsi, le simple fait qu'une technique soit potentiellement dangereuse, doit conduire à la suspendre.

Mais dans cette philosophie, l'équilibre avec la nécessité de permettre les innovations technologiques n'est pas suffisamment prise en compte.
Or, le droit c'est l'art de ce qui est juste et équitable (*jus est ars boni et aequi*).

Au sein de la réflexion globale juridique sur l'appréhension par le droit de l'IA, les responsabilités de type responsabilité médicale sont un domaine clef à appréhender voire à exporter.

Article L1142-1 Code de la Santé Publique :
« I. - *Hors le cas où leur responsabilité est encourue en raison d'un défaut d'un produit de santé, les professionnels de santé mentionnés à la quatrième partie du présent code, ainsi que tout établissement, service ou organisme dans lesquels sont réalisés des actes individuels de prévention, de diagnostic ou de soins ne sont responsables des conséquences dommageables d'actes de prévention, de diagnostic ou de soins* ***qu'en cas de faute.***

Les établissements, services et organismes susmentionnés sont responsables des dommages résultant d'infections nosocomiales, sauf s'ils rapportent la preuve d'une cause étrangère.

II. - Lorsque la responsabilité d'un professionnel, d'un établissement, service ou organisme mentionné au I ou d'un producteur de produits n'est pas engagée, un accident médical, une affection iatrogène ou une infection nosocomiale ouvre droit à la réparation des préjudices du patient, et, en cas de décès, de ses ayants droit au titre de la solidarité nationale, lorsqu'ils sont directement imputables à des actes de prévention, de diagnostic ou de soins et qu'ils ont eu pour le patient des conséquences anormales au regard de son état de santé comme de l'évolution prévisible de celui-ci et présentent un caractère de gravité, fixé par décret, apprécié au regard de la perte de capacités fonctionnelles et des conséquences sur la vie privée et professionnelle mesurées en tenant notamment compte du taux d'atteinte permanente à l'intégrité physique ou psychique, de la durée de l'arrêt temporaire des activités professionnelles ou de celle du déficit fonctionnel temporaire.

[1] Hans JONAS, le principe responsabilité 1979

Ouvre droit à réparation des préjudices au titre de la solidarité nationale un taux d'atteinte permanente à l'intégrité physique ou psychique supérieur à un pourcentage d'un barème spécifique fixé par décret ; ce pourcentage, au plus égal à 25 %, est déterminé par ledit décret. ».

Ce n'est pas l'aspect indemnisation de la loi qui va poser problème mais celui d'une éventuelle faute médicale, voire d'une faute d'une intelligence artificielle. Néanmoins, la plupart des techniciens pensent que la vraie intelligence artificielle (avec une part de conscience), malgré le deep Learning, n'existera pas .

La question de l'évolution d'un droit de la responsabilité face aux progrès technologiques se résout par la nécessité d'une réforme, de textes spéciaux ou le maintien des principes anciens.

Les juristes auront ils une attitude passive, isolée ou prospective en interface avec les techniciens et chercheurs qui seuls peuvent leur indiquer la potentialité du risque.

Le rapport Villani, qui devait être remis en janvier 2018, a été remis le 29 mars 2018, dans le cadre d'un grand rassemblement au collège de France. Le rapport devait exposer le choix juridique fait, ce qui n'a pas été réellement le cas. A noter que le rapport fait sous l'ancien gouvernement et rendu en mars 2017 avait conclu à la vivacité du droit positif actuel et à sa capacité à absorber les évolutions actuelles de l'IA.

I – sur la vivacité des règles existantes :

La nécessité de faire un choix et le maintien d'une discrimination, les exemples des principales réglementations :

Accidents du travail (loi de 1898 !, première loi à base de risque et d'indemnisation pesant sur celui qui tire profit du risque, mais indemnisation limitée).

Accidents de la circulation (sous les feux de la rampe avec les véhicules autonomes, un responsabilité pour risque mais la faute perdure dans la notion de conducteur, le conducteur reste la notion clef de la loi Badinter de 85, pourra t elle survivre à une vraie IA)

Responsabilité du faits des produits défectueux (la directive européenne de 1985, transposée en droit français en 1998, est sans doute l'exemple à suivre, en ce sens la résolution européenne : Règles de droit civil sur la robotique

Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)).

La France avait opté pour l'exonération pour risque de développement (le fabricant d'un produit défectueux ne peut être tenu pour responsable en cas de défaut d'un produit qui n'a pu être découvert ni évité car l'état des connaissances scientifiques et techniques objectivement accessibles lors de la mise en circulation du produit ne le permettait pas. En d'autres termes, ce n'est pas la défectuosité qui est future mais sa découverte. C'est le développement de techniques nouvelles qui permettra de déceler ce qui, avant elles, était indécelable. Mais ce n'est pas simplement un vice « indécelable », que des investigations exceptionnelles auraient permis de découvrir, c'est un vice qu'il est matériellement impossible pour le fabricant de connaitre car les techniques existantes sont insuffisantes).

Enfin, les **Accidents médicaux**. La loi Kouchner de 2004, pourra servir d'illustration, quant à la gestion de risque crée par des innovations médicales. Une indemnisation intégrale mais avec un seuil, une alliance fonds / assurance.

Même si en matière de responsabilité médicale, l'apport intéressant serait de voir un algorithme d'aide de prise à la décision de diagnostic médical ou un robot médical comme cause d'un accident médical identifié et de voir ce qu'en fait la jurisprudence, **pour l'instant il n' y a pas encore de cas.**

Les deux grands domaines proches pour l'utilisation des algorithmes (d'aide) de prise de décision, sont la matière médicale et judiciaire.

Deux Cours d'Appel tests (Rennes et Douai) ont appliqué le logiciel prédictive, pour rendre des jugements de manière autonome, c'est un échec total, pour le moment.

Il resterait peut être de se recentrer sur la responsabilité délictuelle, #1382 , le droit commun, issue du Code Napoléon, des temps des juristes comme Portalis, aurait pu être plus adapté et global : « *Tout fait quelconque de l'homme qui cause à autrui un dommage oblige celui par la faute duquel il est arrivé à le réparer* »

On aurait pu transformer le nouvel article 1240 en : « *tout fait de toute intelligence qui cause à autrui un dommage oblige celle par la faute de laquelle il est arrivé à le réparer* ». sauf les oppositions existantes quant à la nécessité de donner un patrimoine aux robots.

II – Sur la nécessité de réforme ?

Que ce soit le rapport de la commission sur le rapport précédent qui pensait à la vivacité du droit existant, certains articles, ou encore le rapport Villani quasi inexistant sur l'aspect juridique, La pensée majoritaire serait pour une absence de réforme.

Il y a un article fort intéressant d'Etienne VERGES, Professeur à Grenoble, sur les risques scientifiques et les innovations dans le droit de la responsabilité à la revue McGill publiée en juin 2014 (en Anglais !)[2], dans lequel il indique clairement que le droit commun de la responsabilité est suffisamment flexible pour absorber la responsabilité en matière de risques technologiques, sans avoir recours à des régimes spéciaux et à ce titre, il cite les exemples de la position de la Cour de Cassation, sur le distilbène[3] et du Conseil d'Etat sur l'amiante.
Dans ces deux cas, que ce soit le laboratoire ou l'Etat, il y avait des publications scientifiques qui indiquaient les risques envisageables.

On était donc face à une faute de négligence ou d'omission et la faute de négligence a été appréciée au regard du risque connu.

Le critère est donc un défaut de vigilance du laboratoire au regard du risque connu et de connaissances scientifiques et de risques identifiés.

Il expose certaines évolutions sur le fondement de l'article 1384 alinéa 1er, montrant cette grande capacité des principes généraux de la responsabilité civile à absorber une part des innovations techniques et scientifiques.

C'est ainsi que notamment en 2002, à propos de l'amiante, l'obligation de sécurité de résultat de l'employeur conduisant à la reconnaissance de sa faute inexcusable et donc à l'indemnisation complémentaire de la victime a été reconnue par la Cour de Cassation[4].
De la même manière lorsque la jurisprudence a distingué la garde de la structure et la garde du comportement[5].
Dans cet exemple, la Cour de Cassation décide d'attribuer la garde au fabricant, car il a gardé le contrôle sur le produit et la notion de garde de la structure est bien une évolution des notions de responsabilité civile par rapport aux technologies dangereuses.

---

[2] M. Etienne Verges, Risks and uncertainties of scientific Innovations in frenche liability law : between radical departure and continuity, Revue de droit Mc Gill, l'innovation Technologique et reponsabilité civile, juin 2014.
[3] Civ. 1ère, 07.03.2006, Bull. Civ. I, 131, n°143.

[4] Cass. Soc. 28.02.2002, JCP2, 614
[5] C. Cass.05.01.1956, 2ème Ch. JCP2, n°9095

Ces évolutions existent également par rapport aux dommages et ainsi la grande extension du préjudice moral[6]qui admet la réparation d'un préjudice moral constitué par la seule crainte de subir un préjudice.
Donc, nous passons du risque technologique à l'indemnisation d'un risque de dommages, même si cela contrarie les conditions du préjudice indemnisable qui doit être direct personnel et certain.
C'est le caractère certain du dommage qui est ici combattu.

Et en l'espèce, certaines juridictions du fond ont eu à se prononcer sur le problème des antennes relais et donc pour contourner ce problème d'absence de certitude scientifique, sont passé par le préjudice moral.

Il pourrait être envisagé des préjudices à naitre, par exemple les journalistes qui se déplacent sur des sujets avec des petites unités dont la puissance est quasi équivalente à des antennes relais pour avoir un 4G puissante et envoyer les sujets vidéo lourds à leur rédaction, en un temps record au plus près de l'actualité.
Si ces journalistes sont salariés d'une chaine de télévision, on passera par les accidents du travail, mais s'ils sont indépendants, il faudra passer par un autre biais (responsabilité du fait des produits défectueux ?).

L'évolution et la grande fraicheur et adaptabilité du droit commun, peut également se voir à travers la notion de causalité.

Normalement une orthodoxie juridique voudrait que lorsque l'on lâche sur une des conditions de la responsabilité civile, il faut être beaucoup plus dur sur les autres conditions de la responsabilité civile.
Ainsi, en cas de responsabilité sans faute on devrait avoir une causalité adéquate.

Or, on est passé par une torsion de la causalité et par le jeu des présomptions (article 1353 du Code Civil) en cas d'incertitude scientifique et notamment en 2009, à propos du lien entre la sclérose en plaques et l'hépatite B, les présomptions qui ont été retenues ont été une proximité de temps, une absence d'antécédents neurologiques et d'avoir écarté toutes les autres causes possibles.

Mais en général, ces évolutions jurisprudentielles sont un préalable à une mise en place légale plus structurée. Ainsi, notamment avec l'émergence d'une présomption de causalité légale dans le code de la santé publique par rapport au HIV ou à l'hépatite C, cette présomption de causalité se retrouve également en cas d'accident nucléaire.

[6] C. Cass., 19.12.2006, n°0515719

Ces évolutions qui ont été listées de manière non exhaustive sont parfaitement normales car les évolutions technologiques s'accompagnent toujours de choses incertaines.

Pour le Professeur Etienne VERGES, il y a eu une évolution du droit de la responsabilité civile qui s'est d'une part constituée par une rupture (faute inexcusable et obligation de sécurité de résultat) et reconnaissance des dommages écologiques sans victime directe et d'autre part, une continuité avec deux principes qui émergent, le principe de précaution et l'indemnisation du risque de dommages.

Le droit français n'est pas un droit anglo-saxon, oral, mais écrit, basé sur des principes qui datent de Montesquieu, comme celui du rôle du juge, qui n'est que « *la bouche qui prononce les paroles de la loi* » (même si réforme du droit des obligations, d'octobre 2016) permet de faire rentrer le juge dans le contrat).

L'avocat, comme le médecin face à un nouveau virus, s'adapte.

C'est au niveau européen que la réforme juridique doit se décider et se faire. C'est la condition pour avoir un marché suffisant (il faut se rappeler que google ne pouvait être développé qu'aux Etats –Unis, au regard de la taille du marché).

Or, la résolution du parlement européen impose des législations.

Ce qui est une bonne direction, la meilleure façon de ne pas brider l'évolution c'est de légiférer (tout ce qui n'est pas interdit est permis) surtout cela apporte une sécurité juridique, une activité qui est assurable. Un risque quantifié.

Il ne reste plus qu'à choisir le modèle.

Une responsabilité générale d'un droit des robots? quid de la complexité des intervenants et de la confusion des responsabilités ? puisqu'y compris l'état peut intervenir, on peut envisager le cas des véhicules autonomes (l'existence de routes avec balises et le fait que l'état impose le véhicule autonome peut permettre d'envisager la responsabilité de l'Etat).

On peut lire en ce sens, (*Recueil Dalloz 2018 p.129*
Véhicule autonome : vers une autonomie du régime de responsabilité applicable ?
Marjolaine Monot-Fouletier, Maître de conférences à la Faculté de droit de l'Université catholique de Lyon et de Marc Clément, Rapporteur à la Cour administrative d'appel de Lyon), qui plaident pour un fonds de garantie pour les véhicules autonomes et pour qui :
la Résolution du Parlement européenne relative aux règles civiles de la robotique : une fois les parties responsables en dernier ressort identifiées, leur responsabilité devrait être proportionnelle au niveau réel d'instructions données au robot et à l'autonomie de celui-ci, de sorte que, plus un robot

est autonome, plus sa capacité d'apprentissage est grande, et plus sa période de formation a été longue, plus grande devrait être la responsabilité de la personne qui l'a formé ; la résolution précise que lorsqu'il s'agit de déterminer qui est la personne réellement responsable du comportement dommageable du robot, les compétences acquises par un robot au cours de sa formation ne devraient pas être confondues avec les compétences strictement dépendantes de sa capacité à apprendre de manière autonome.

Dans cette logique, l'intensité variable de l'autonomie du véhicule est susceptible d'influer sur les conditions d'intervention du fonds : plus le véhicule pourra être considéré comme autonome, plus le fonds interviendra à titre principal dans la compensation du dommage subi ; moins le véhicule sera autonome, plus l'on restera sur le terrain d'engagement d'une responsabilité classique nécessitant la démonstration d'une imputabilité, d'un lien de causalité.

Donc il y a, sans doute, l'obligation de passer par un fonds mais qui finance le fonds et comment se retourne t-on contre les responsables in fine ? Un fonds général et commun comme l'exemple Néo-Zélandais qui fonctionne depuis plus de 40 ans ?

On prend l'exemple de la responsabilité médicale ? le plus récent ? le moins bridant. Sauf que la nécessité d'indemniser le risque créé en l'occurrence est justifié par l'obligation de sauver une vie, pas pour plus de confort (vitesse, fatigue etc.), sauf que l'effet induit de l'IA est censé être plus de sécurité (quid des logiciels programmés pour hiérarchiser des vies, en fonction de l'âge etc.).

L'exemple des accidents du travail ? l'idée du risque profit et le renfort de la faute inexcusable et des obligations de sécurité de résultat.

Mais, il n' y a pas que du risque profit, en matière d'IA.

L'Exemple de la circulation ? sur le risque créé , mais on n'a plus de conducteur et on a surtout du fabricant.

Fort à parier que l'on se dirige vers responsabilité du fait des produits défectueux, avec exonération pour risques de développements.

Sauf que toute la responsabilité des robots sera t elle à base d'un produit défectueux ? (Dalloz IP/IT 2016 p.287)
« Robots intelligents et responsabilité : quels régimes, quelles perspectives ? »
Georgie Courtois, Avocat Associé, De Gaulle, Fleurance & Associés, qui identifie comme éléments de responsabilités :

En tout état de cause, lorsque des incidents impliquant des robots vont survenir, la détermination des responsables potentiels requerra une réflexion sur le rôle :

- du propriétaire ;

- de l'utilisateur ;

- du concepteur ;

- du fabricant ;

- du programmeur du logiciel intégré au robot ;

- *du concepteur de l'IA.*

Il conclut par : *Les régimes de responsabilité du fait des choses et des produits défectueux semblent aujourd'hui adaptés au niveau d'autonomie des robots. À mesure de l'évolution de l'IA et de la prise d'autonomie des robots, un régime spécial similaire à celui de la responsabilité du fait des animaux pourrait être adopté. Il sera nécessaire de maintenir une responsabilité sur le propriétaire ou l'utilisateur du robot autonome. Dans ce cadre, la création d'une personnalité juridique du robot semble inappropriée.*

Ils sont néanmoins nombreux à rappeler : le respect des lois d'Asimov : Runaround (1942), d'Isaac Asimov. Les trois lois sont :

Un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger.

Un robot doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec la Première loi.

Un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la Première ou la Deuxième loi.

In fine étant donné que les acteurs politiques disent qu'il faut toujours un humain pour valider la décision finale, il restera tout le mécanisme traditionnel de la faute civile et pénale, pour moraliser les comportements.

Il sera intéressant de voir comment se résout l'accident mortel provoqué par la voiture autonome Uber ayant tué un piéton, les avocats saisis du dossier pour le compte de la fille, sont spécialisés en réparation du dommage Corporel, ce qui n'est pas anodin. Mais l'exemple américain, de droit anglo- saxon, ne sera pas transposable *in extenso*.

Enfin, comme indiqué précédemment, le rapport Villani du 29 mars 2018 a laissé un goût d'inachevé quant à l'approche juridique.
Après avoir mis en exergue les avancées médicales grâce à l'IA, en terme de diagnostic, de soins, de base de donnée médicale, la seule vision en terme de responsabilité relève de ces quelques lignes :

- -« *clarifier la responsabilité médicale des professionnels de sante en cas d'utilisation de technologies d'intelligence artificielle : à l'heure actuelle, la responsabilité médicale d'un médecin peut être engagée en raison d'une faute ou d'un manquement déontologique (généralement conçu*

comme un manquement à des obligations d'information et au droit du patient à consentir de manière claire à l'acte médical). En l'absence de la reconnaissance d'une personnalité juridique autonome pour l'algorithme et le robot, il serait envisageable de tenir le médecin pour responsable de l'utilisation des programmes, algorithmes et systèmes d'intelligence artificielle, sauf défaut de construction de la machine ».

Thème C

« Numérique et pratiques juridiques »

« 50 nuances de mots : du robot ou du juriste, qui porte la cravache ? »

Comment l'intelligence sémantique peut être utilisée dans une perspective de « robotisation » de la justice ?

Florian Laussucq
Doctorant en droit public
CRDEI, Université de Bordeaux
florian.laussucq@u-bordeaux.fr

Ouassila Labbani Narsis
Maître de conférences au laboratoire CIAD
Université de Bourgogne – Dijon
ouassila.labbani@u-bourgogne.fr

Les rapports entre droit et numérique sont bien souvent abordés sous l'angle de la rivalité, une opposition dans l'exercice d'une prérogative lourde, celle de dire le droit, et donc le juste. Pour autant, il est parfois possible de trouver une complémentarité réelle entre les deux. L'intelligence artificielle, et en particulier, l'étude de la distance sémantique est indéniablement un terrain d'élection pour travailler sur ce paradigme nouveau. Pour dégager ce champ de recherche, il convient auparavant d'en délimiter clairement les contours, exercice d'autant plus impérieux qu'un tel travail d'une part concerne des publics par définition peu avertis (chaque communauté scientifique « découvrant » l'autre) et d'autre part, car il serait ironique de ne pas faire un tel effort sémantique alors même qu'il s'agit de l'objet de l'étude.

Le droit est ici entendu dans un sens large, dans la mesure où la thèse défendue est celle de l'applicabilité et de l'apport de la notion de « distance sémantique » à l'ensemble des aspects de ce que l'on appelle le droit. Autrement dit, la présente étude se veut généraliste quant aux branches du droit concernées, et de la même manière, les conséquences seront abordées tant sur un plan théorique que pratique (contentieux).

Il faut dès lors se pencher sur l'idée de distance sémantique. La distance sémantique est un concept qui se matérialise par la différence entre le niveau d'abstraction offert par le langage d'interface et le niveau de conceptualisation de sa tâche par l'usager. Autrement dit, le concept de distance sémantique s'inscrit dans le cadre des relations hommes-machines, en tenant compte du fait que les deux ne parlent pas la même langue, ce qui crée de fait une distance. La réduction de distance sémantique vise alors à épargner un effort intellectuel à l'usager (l'humain) en utilisant des notions qui lui sont familières, aussi bien pour lui permettre de s'exprimer, que pour lui présenter les informations.

Rapporté au cadre du droit, il s'agira de faire cesser l'ambiguïté entre deux termes pour une machine, c'est-à-dire la possibilité pour un mot d'avoir plus d'un sens à la fois, la machine plus encore que l'être humain ne pouvant accepter une telle situation. En effet, que l'on parle de robot, d'ordinateur ou de machine, il est à chaque fois fait référence à un objet, nécessitant d'être dirigé dans sa compréhension du monde qui l'entoure.

A ce stade, il faut maintenant préciser notre sujet quant à son étendu, sur au moins deux aspects : d'une part, la présente étude se concentre sur les solutions (et les problèmes) de distance sémantique en matière de justice prédictive, celle-ci s'entendant dans un sens d'IA semi forte, et non d'un robot juge relevant du fantasme à moyen ou court terme. Et d'autre part, il ne s'agira que d'un survol des solutions techniques, un tel travail

n'ayant pas la prétention de présenter ces techniques, ni de manière détaillée, ni de manière exhaustive.

L'intérêt est ici multiple : dans un premier temps, il s'agit d'éclairer d'un jour nouveau les questionnements des juristes quant aux perspectives, et aux limites du robot juriste. Il en ressortira que les craintes si elles sont légitimes, n'en sont pas nouvelles : pour paraphraser Hobbes, c'est bien l'homme derrière la machine qui demeure un loup pour l'homme.

Il en ressort dans un second temps une interrogation parallèle, celle de l'influence de la machine sur le raisonnement du juriste, c'est-à-dire ce que nous apprend la « robotisation » du raisonnement juridique précisément sur ce même raisonnement. Une solution pouvant se trouver dans une évolution de la logique juridique de manière à devenir complémentaire de la machine, en étant moins dogmatique et moins… robotique.

La question de recherche est donc la suivante : le concept de distance sémantique est-il applicable dans le domaine du droit et dans quelle mesure peut-on y trouver des motifs de rapprochement plutôt que des concurrences entre les professions du droit et du numérique ?

I. L'applicabilité de la distance sémantique en droit

Il s'agit de savoir ici dans quelle mesure la notion de distance sémantique, qui s'inscrit dans un paradigme de recherche en sciences du numérique, est transposable en droit, tant sur le plan théorique (A) que pratique (B).

A) La transposition théorique du cadre conceptuel de distance sémantique en droit

La distance sémantique en tant que champ de recherche est un sujet largement étudié par la science du numérique, mais également par d'autres champs scientifiques[1]. C'est cette transversalité qui laisse à penser qu'il serait possible de le transposer dans la recherche scientifique, puis en droit positif.

Sans aller aussi loin dans l'analyse, il s'agira ici de démontrer, en partant d'un exemple de distance sémantique, que la transposition en droit est possible.

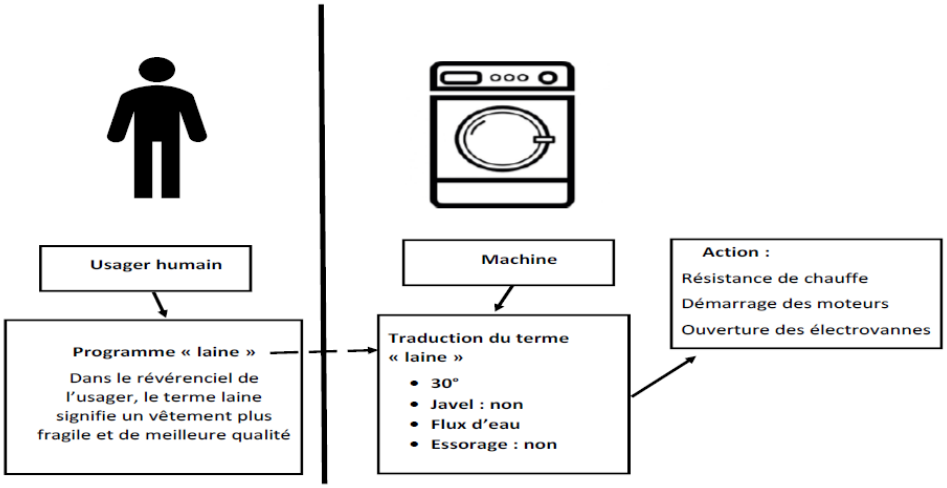
1) Application standard du concept de distance sémantique

Par exemple, La machine à laver dispose d'une touche "couleur" et d'une touche "blanc". Or il s'agit de deux abstractions qui appartiennent au référentiel de l'usager. Elles seront alors transposées dans la machine en température de lavage, protocole de lavage, adjonction ou non d'eau de Javel, ... ; et actionneront les résistances de chauffe, les moteurs et les électrovannes. Exprimer les réglages de lavage selon les notions "couleur" et "blanc" réduit la distance sémantique de l'interface de la machine à laver, par rapport à une expression sous la forme de détails du protocole de lavage, difficilement compréhensible par un usager

[1] Etienne Brunet, « Peut-on mesurer la distance entre deux textes ? », Corpus [En ligne], 2 | 2003, mis en ligne le 15 décembre 2004, consulté le 06 septembre 2018. URL : http://journals.openedition.org/corpus/30

«standard». Le schéma suivant illustre parfaitement ce phénomène : l'usager humain va utiliser le programme «laine», ce terme renvoyant dans son esprit à une série de concepts intériorisé par le cerveau humain[2], mais pour la machine, le terme «laine» ne signifie rien en soit. La machine fonctionne en effet, de manière schématique, par le biais d'algorithme, c'est-à-dire d'une suite d'instructions simples dont la réponse est «oui» ou «non» (vrai ou faux en réalité). Il y a donc un décalage entre le concept de «laine» utilisé par l'usager humain, et ceux reconnus par la machine : c'est la distance sémantique (illustrée ici par la flèche en pointillés).

Schéma n°1 :



2) Application du concept en droit

Il faut encore une fois préciser ici que la présente étude se situe dans l'hypothèse d'une robotisation du droit, tant au niveau du conseil que du contentieux (le premier existe déjà, le second relève pour l'heure de la prédiction). Dans une telle hypothèse, le cas de figure précédemment posé se présente à nouveau, lors d'une interaction entre d'une part, un usager humain (le juriste, voir le client directement[3]), et d'autre part, la machine. Comme dans le cas précédent, la machine est ici programmée par le biais d'algorithme, c'est-à-dire d'une suite d'instructions simples[4]. Il y a donc une distance entre les concepts sémantiques tels qu'ils existent dans l'esprit du juriste, et ceux utilisés par le programmeur (qui n'est pas forcément un juriste).

Le schéma 2 reprend les éléments précédents pour illustrer de manière simplifiée ce que pourrait être le fonctionnement de l'interaction homme machine, et les problèmes d'ambiguïté sémantique qui pourraient en découler.

Une telle problématique peut sembler très abstraite, mais elle est pour autant d'ores et déjà applicable,

---

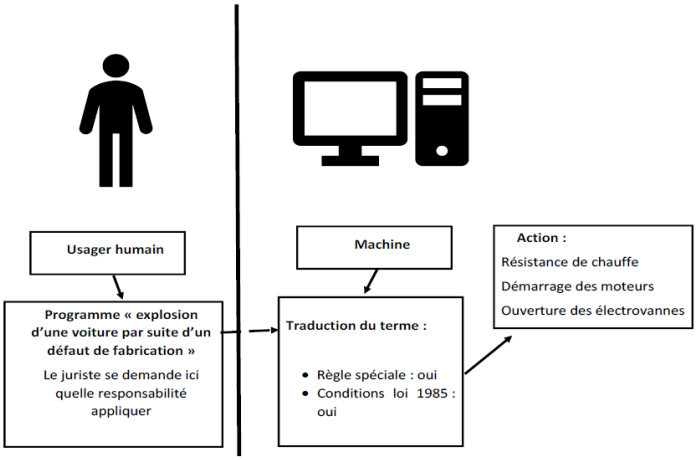
[2] Concepts qui sont parfois trop vagues et doivent à ce titre être développés, cf. infra

[3] Eve d'Onorio di Méo, *Le rôle des Avocats dans le défi numérique* consultable sur : https://www.village-justice.com/articles/role-des-Avocats-dans-defi-numerique,24774.html#efucDPEKzTSoJSvs.99;

[4] Un exemple d'algorithme appliqué au droit est proposé en annexe

notamment dans le cas de ce que l'on appelle le « web sémantique ».

Schéma n°2 :



B) Application pratique dans le cadre du web sémantique

Il s'agit en effet de démontrer que la distance sémantique entre les termes utilisés par des usagers différents (homme et machine, machine et machine) pose d'ores et déjà des questions juridiques. Il ne s'agira pas d'en dresser une liste exhaustive, mais simplement une présentation schématique dont l'objectifs et de mettre en exergue l'importance pour les juristes de s'intéresser aux questions de sémantiques dans le numérique. Ainsi allons-nous aborder ici la question du *spamdexing* et des conséquences qu'il peut avoir en termes de contrefaçon ou de parasitisme. Pour bien comprendre la problématique ici ainsi que son lien avec la distance sémantique, quelques précisions doivent être faites au préalable.

Tout d'abord, il faut comprendre comment fonctionne un moteur de recherche. Un moteur de recherche est une application qui va extraire du web les sites répondant à une requête effectuée par un utilisateur, qu'il liste dans des pages de résultats. L'élément clé à retenir dans le fonctionnement d'un moteur de recherche est qu'il classe ses résultats : il ordonne les sites présentés selon des critères qui lui sont propres. Le premier de la liste est donc, selon le moteur, le plus pertinent pour répondre à la requête, d'où l'importance d'un tel classement sur le plan économique. Or, parmi ces critères figure la popularité, qui se mesure aux nombres de « clics » reçus par le site. D'où la tentation de tromper l'algorithme de Google, en gonflant artificiellement ce nombre, c'est ce que l'on appelle le «spamdexing». Ce dernier, également appelé «spam indexing» se définit comme un *« ensemble de techniques consistant à tromper les moteurs de recherche sur la qualité d'une page ou d'un site afin d'obtenir, pour un mot-clef donné, un bon classement dans les résultats des moteurs*[5]*»*.

Les techniques habituelles de référencement abusif sont de deux types, avec d'une part la création de sites

[5] Ibrahim Coulibaly, *Contrefaçon, concurrence déloyale et parasitisme sur Internet : entre spamdexing, cybersquatting et typosquatting,* consultable sur : https://www.village-justice.com/articles/Contrefacon- concurrence-deloyale,17655.html#DbtF4ByBA8iOT0mk.99

internet dont le but n'est autre que d'héberger des liens (*farm links*) qui en redirigeant les uns vers les autres, créent du « click » et améliorent ainsi leur classement via le référencement « naturel » de Google. L'autre type de référencement abusif consiste en un enregistrement massif de noms de domaines, étant précisé qu'en soi, aucune de ces pratiques n'est abusive[6]. Il en découle que ces pratiques sont tout à fait légales, à condition de ne constituer ni un parasitisme, ni une concurrence déloyale, voir une contrefaçon.

Afin d'illustrer le présent propos, il sera uniquement fait référence au problème juridique soulevé par le premier référencement, que l'on qualifie de référencement naturel.

En effet, dans ce cas, il s'agit bien de tromper un algorithme programmé en fonction d'une distance sémantique prédéfinie par Google, qui va ordonner l'affichage du terme «botte» (et les sites commerciaux relatifs à ce terme) dès lors que l'on tape « santiags ». Il s'agit en effet de noms de domaines (« bottes.com ») qui présentent la particularité d'être descriptifs, c'est dire de se confondre avec un objet. Ne pouvant faire l'objet d'une protection par le droit commun, la personne exploitant un nom de domaine descriptif ne peut alors interdire son utilisation par des tiers.

L'utilisation est donc libre, à condition toutefois que le tiers qui le réutilise n'agisse pas de façon déloyale. Dans un tel cas, le droit commun s'applique et une telle pratique est sanctionnable devant les tribunaux[7]. Par exemple, la Cour d'appel de Paris, a pu qualifier de «comportement déloyal» l'utilisation d'un nom de domaine car il «traduit une volonté délibérée de tirer profit à moindre coût des investissements d'autrui en créant dans l'esprit du public un risque de confusion sur l'origine du service offert[8] ». En l'espèce, il s'agissait d'une société proposant des annonces de locations saisonnières d'avoir choisi un nom de domaine quasi identique, « annoncesvacances.com », à celui adopté trois mois plus tôt par une société concurrente, «annonces-vacances.com ».

On constate alors l'importance pour les juristes de comprendre et mesurer les mécanismes d'indexation des moteurs de recherches, c'est-à-dire les algorithmes visant à traduire le langage du client potentiel dans le langage du moteur de recherche, et ce d'autant plus que la responsabilité de Google a été définitivement écartée par la CJUE en 2010[9] (ce qui est contestable car Google maitrise la programmation des domaines notamment via la sémantique).

Nous l'avons vu, le concept de distance sémantique a une pertinence dans le droit, tant sur le plan théorique que pratique. Il convient alors d'en envisager la portée, c'est-à-dire les conséquences.

---

[6] CA Paris, pole 5, ch. 1, 26 septembre 2012, n° 10/22304, SA CRM COMPANY GROUP c/ SARL ADICTEL

[7] Sur les vertus du droit commun de la responsabilité, Muriel Chagny, Concurrence déloyale et nom de domaine. Usez mais n'abusez pas des noms de domaine !, Communication, commerce électronique, février 2012, pp.27-28

[8] CA Paris, 8 oct. 2003, D. 2004. Somm. 1157, obs. Y. Auguet

[9] CJUE, 23 mars 2010, aff. C-236/08 à 238/08

II) La portée du concept de distance sémantique en droit

La question de la distance sémantique intéresse directement les juristes, en ce qu'elle concerne l'interaction homme/machine. Ainsi, le juriste de demain se devra de maitriser non seulement les outils, mais également la mécanique interne, sous peine d'en devenir prisonnier (A). Mais au-delà, la question de l'ambiguïté sémantique nous parait intéresser la manière même dont les juristes pensent les concepts et les manient au quotidien (B).

A) L'apport direct : réduire la distance sémantique pour améliorer l'utilisation du numérique en droit

1. Les apports d'une bonne interface juriste/robot

Tout d'abord, il convient de préciser qu'au-delà des apports, et donc des possibilités qu'ouvre le numérique, il risque de s'agir également d'une obligation. En effet, à l'heure de la robotisation du droit, il parait nécessaire pour les juristes non seulement de manier les outils du numérique, mais au-delà, de pouvoir en être pleinement acteur. Or pour se faire, il nous parait nécessaire de maitriser également les mécanismes sous-jacents, «invisibles», du numérique, sous peine de devenir dépendants d'algorithmes dont on peine à connaitre l'origine, et dont certains voudraient aujourd'hui faire croire qu'ils ne sont le fruit d'aucune intervention humaine[10].

Mais dans une perspective plus optimiste, l'amélioration de l'interface homme/machine ouvre de nombreuses perspectives pour les juristes, perspectives déjà maintes fois citées et que nous ne faisons ici que reprendre : possibilité de se concentrer sur les tâches les plus complexes, allégement des délais de justice[11], ouverture aux justiciables qui pour de petits litiges ne souhaitent pas procéder par avocat, etc…

Toutefois, il est possible de constater que le développement du numérique peut engendrer une transformation profonde de l'activité du juriste. Or il nous semble faux de penser que cette évolution ne puisse pas être positive.

2. Une opportunité pour les juristes

La multiplication des interactions entre juristes et robot ne se fera pas aisément, pour la simple raison que les machines amenées demain à faire un travail juridique, n'ont pas eu la «programmation», aussi appelée parfois «formation universitaire», que les humains ont eue. Or, pour programmer des robots juristes, qui mieux que des juristes ? Pour illustrer notre propos, prenons l'exemple simple d'une conversation entre deux personnes :

[10] Quand l'intelligence artificielle transformera notre vie au bureau, the economist, cité par courrier international du 28/03/2017

[11] « A titre d'exemple, en 2011, le délai moyen pour obtenir une décision de justice était de 12 mois devant le juge

« - Tu connais un juriste ?

- Oui je connais un avocat.

- Merci. »

Dans une conversation aussi banale, la première personne a généralisé sa requête au concept de juriste, qui représente la catégorie la plus abstraite recouvrant toutes les formes de réponses acceptables. La deuxième a, probablement sans même y prêter attention, utilisé sa taxinomie de concepts pour en déduire qu'un avocat est un juriste, et que par conséquent sa réponse est pertinente. Le fait que cette connaissance taxinomique soit partagée est implicite, puisque la deuxième personne suppose que sa réponse sera comprise sans préciser qu'un avocat est un juriste, et que c'est effectivement le cas. Le recours à des conceptualisations partagées et aux inférences qu'elles permettent est donc au cœur d'activités aussi simples que cet échange. Par exemple, si l'on recherche des «lois» concernant des «sociétés». Si le système d'information se contente de travailler au niveau textuel, avec les mots clefs «loi» et «société», vont alors apparaître plusieurs problèmes :

- Le bruit : le système ne saura pas faire la différence entre les lois déjà votées, les projets de loi ; les lois de transposition d'une directive, une loi interprétative (et donc rétroactive), une loi de ratification d'un traité, etc…
- Le silence : le système, s'il rencontre le terme « E-N-T-R-E-P-R-I-S-E », ne saura pas qu'il est pertinent pour votre requête, car il cherche le mot « S-O-C-I-E-T-E ».

Si maintenant, vous expliquez au système quelques aspects de notre réalité sur le système juridique (les SAS et les SARL sont des sous-types de société, qui est lui-même un sous-type d'entreprise), les documents (Lois et règlement sont des sous-types de droit national, qui est lui-même un sous-type de Droit applicable) et les relations entre les deux, avec leurs signatures (par exemple, il existe une relation champ d'application, qui peut s'établir entre un Droit applicable et un Société.)

On constate donc que l'ambiguïté sémantique, si elle est résolue automatiquement par un cerveau humain, ne l'ai pas de la même manière par une machine. Il faut donc «apprendre» à la machine à résoudre ce problème, ce qui revient à calquer le raisonnement humain. Et c'est là que les juristes trouvent pleinement leur place; qui mieux qu'un avocat pour programmer les instructions de l'algorithme (voir annexe), via une étroite collaboration avec les professionnels du numérique.

B) <u>Une portée indirecte : l'ambigüité sémantique en droit</u>

Nous l'avons vu, la distance sémantique est un concept qui vise à étudier un phénomène constitué précisément par l'écart qui existe entre deux concepts, en s'efforçant de réduire ainsi l'ambigüité sémantique. En effet, le juriste connait mieux que quiconque les dangers de l'ambigüité entre les termes: les exemples sont légions de jurisprudence où il fut question pour les juges suprêmes de préciser les contours d'un concept donné, et ce afin

de pouvoir distinguer les régimes juridiques, c'est-à-dire les conséquences attribuées à un terme ou à un autre : ainsi en est-il de la définition du déchet[12] ou du cours d'eau[13].

De plus, la question de la distance sémantique se situe dans le champ plus vaste de la sémantique, lui-même compris dans celui d'ontologie : la question de l'ontologie, c'est-à-dire du classement des concepts les uns par rapport aux autres, en vertu de leurs composants propres, est une question théorique et pratique des sciences du numériques[14].

Bibliographie :

Antoine Garapon, *Les enjeux de la justice prédictive (*JCP, éd. G., 2017, 31)

Andon Tchechmedjiev, *État de l'art : mesures de similarité sémantique locales et algorithmes globaux pour la désambiguïsation lexicale à base de connaissances*, Actes de la conférence conjointe JEP- TALN-RECITAL 2012, volume 3 : RECITAL, pages 295–308, Grenoble, 4 au 8 juin 2012.

MAEDCHE A. & STAAB S. (2002). Measuring Similarity between Ontologies. ekaw, p. 251-63.

Mounira Harzallah, Emmanuel Blanchard, Giuseppe Berio et Pascale Kuntz, *Mesures sémantiques pour la comparaison des « constructs » des langages de modélisation d'entreprise*, 8èmes Journées Francophones, Extraction et Gestion des Connaissances, Sophia Antipolis 29 janvier 2008

L. Léger*, C. Tijus, *L'effet de l'hétérogénéité sémantique dans la détection de mots*, Psychologie française 52 (2007) 367–385

[12] Cass. Comm, 3 mars 2015, Communauté urbaine de Marseille Métropole

[13] CE, 21 octobre 2011, Ministre de l'Ecologie, du Développement Durable, des Transports et du Logement C/ EARL CINTRAT (requête n° 334-322 publiée au Lebon),

[14] Jérôme Euzenat, *Quelques pistes pour une distance entre ontologies*, 8èmes Journées Francophones Extraction et Gestion des Connaissances, Sophia Antipolis 29 janvier 2008 ; Latifa Baba-Hamed, Réda Soltani et Kamel Sabri, *Construction d'une ontologie pour la recommandation de films à un utilisateur*, atelier disponible à l'adresse suivante : http://wwwlimics.smbh.univparis13.fr/GBPOnto/data/documents/2010/6bababahamedpreusentation.pdf

Annexe : exemple d'algorithme appliqué au droit

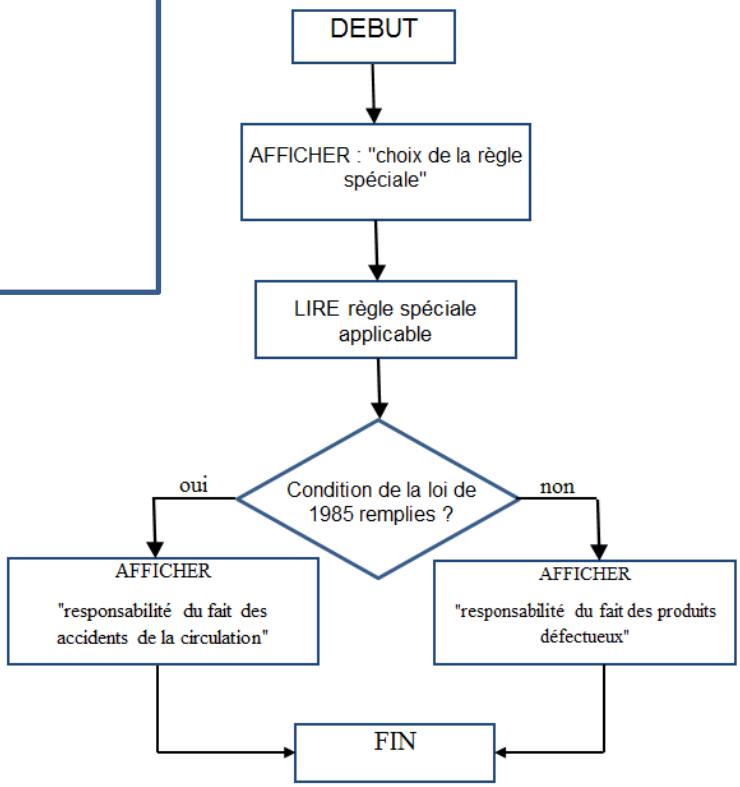
Prenons un exemple classique en droit, le droit de la responsabilité civile. Pour une question de responsabilité, 3 algorithmes sont nécessaires (nous le verrons plus tard, il s'agit ici d'un postulat arbitraire, résultant d'un choix humain) :

- Algorithme 1 : détermination de la règle applicable : règle générale ou règle spéciale ?
- Algorithme 2 : en supposant l'application de la règle spéciale, quelle règle spéciale appliquer ?
- Algorithme 3 : sur la base de la règle spéciale, quelle solution donner au litige ?

Si l'on applique l'algorithme 1, voici ce que l'on peut écrire :

```
Variable règle applicable en Entier
Début
   Ecrire « Il y a-t-il un produit défectueux ? »
   Lire règle applicable
   Si produit (vrai) ET Défaut (vrai)
   Alors
      Ecrire "régime spécial, loi de 1998"
   Sinon
      Si Véhicule terrestre à moteur (vrai) ET Impliqué (vrai)
     Alors
         Ecrire "régime spécial, loi de 1985"
      Sinon
         Ecrire "règle générale"

FinSi
   FinSi
Fin
```



Logique juridique et logique mathématique : quelles convergences ?

Guillaume Aupy & Sébastien Platon*

Récemment, un phénomène tout à fait notable s'est fait jour en matière de droit : la « justice prédictive » [1], c'est-à-dire l'utilisation des nouvelles technologies, et en particulier du Big Data et de l'intelligence artificielle, pour prédire une décision de justice par l'analyse automatisée de la jurisprudence. Aux Etats-Unis, la start-up Legalist a développé un modèle économique fondé sur un algorithme capable de déterminer en quarante-huit heures les chances de succès d'un client et la durée probable des procédures à partir d'une base de données de quinze millions de dossiers sur les vingt-cinq dernières années [2]. Des logiciels sont utilisés par des juridictions américaines pour évaluer les risques de récidive des prévenus [3]. Dans certaines provinces chinoises, des systèmes d'intelligence artificielle proposent même des décisions aux juges [4]. En France, plusieurs sociétés, comme Case Law Analytics, Tyr Legal ou Prédictice, proposent des outils permettant, par exemple, d'estimer le montant de dommages et intérêts ou d'une pension alimentaire, d'obtenir des statistiques sur les chances de gagner une procédure, d'afficher graphiquement l'état d'un contentieux ou les arguments les plus souvent utilisés, etc. Le barreau lillois travaille, en partenariat avec une startup, au développement d'un logiciel qui sera non seulement un moteur de recherches spécialisées mais aussi une moulinette à données permettant « d'enrichir la stratégie judiciaire » [5]. Et les possibilités s'accroitront nécessairement avec la loi du 7 octobre 2016 pour une République numérique [6]. Celle-ci prévoit, en son article 20, que les jugements « sont mis à la disposition du public à

*Respectivement INRIA et Université de Bordeaux et Professeur de droit public, Université de Bordeaux

1. Il y a sur cette question une littérature de plus en plus foisonnante quoiqu'essentiellement composée d'articles relativement courts. Pour une vue d'ensemble plus large, l'on renverra plus particulièrement aux Actes du colloque du 12 Février 2018, organisé par l'Ordre des avocats au Conseil d'Etat et à la Cour de cassation, La justice prédictive, Dalloz, 2018.

2. Garapon A., « Les enjeux de la justice prédictive », Revue pratique de la prospective et de l'innovation n° 1, Octobre 2016, dossier 4

3. Prévost S. et Sirinelli P., « Génération Legaltech », Dalloz IP/IT 2017, p. 65.

4. « Justice prédictive : où en est-on ?», Village de la justice.com, 21 juillet 2016, `https://www.village-justice.com/articles/Justice-predictive-est,22683.html`

5. « Pionniers en Europe, les avocats lillois testent "la justice prédictive" », La Voix du Nord, 20 janvier 2017, `http://www.lavoixdunord.fr/106005/article/2017-01-20/pionniers-en-europe-les-avocats-lillois-testent-la-justice-predictive`

6. Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, JORF n° 235 du 8 octobre 2016.

titre gratuit dans le respect de la vie privée des personnes concernées » - c'est-à-dire avec obligation préalable d'anonymisation. Il en résultera une augmentation considérable des données disponibles, et donc, a priori, de la fiabilité des analyses. L'IA va-t-elle mettre les juristes au chômage ? Cela dépend, en partie, de la capacité du raisonnement juridique à être réduit à une raisonnement de pure logique mathématique.

Avant de répondre à ces questions, il convient de se poser une première question plus générale : *Qu'est-ce que l'intelligence artificielle ?*. D'un point de vue d'informaticien, cette définition n'a, astucieusement, jamais été définie très formellement [7], ce qui permet d'englober un peu tout et n'importe quoi.

Dans cette lignée, plutôt que de définir l'IA, nous allons nous concentrer sur certaines techniques utilisées pour créer des « machines capables de simuler l'intelligence » [8]. En particulier sue les deux techniques les plus courantes et présentes dans les médias ou l'esprit des gens quand on parle d'IA.

— Les techniques dites *computationelles* : résoudre des problèmes avec des données issues d'exemples et de l'apprentissage (réseaux de neurones, deep learning).
— Les techniques dites *symboliques* : résoudre les problèmes avec de la connaissance, être capable de créer un raisonnement (automates, script, systèmes multi-agents etc)

L'idée d'*intelligence computationelle* est de créer un système qui à partir d'un énorme ensemble de *données* (ensembles de faits (*entrées*) et de jugements (*sorties*)), va créer une *boite noire*, qui peut prendre une nouvelle entrée, et créer elle même une nouvelle sortie. Du point de vue de l'utilisateur de ce système, aujourd'hui il n'est pas possible de comprendre le raisonnement qui a amené la sortie. La création de ces boites noires posent de nombreux problèmes éthiques, et juridiquement, il est peu souhaitable que de telles machines soient amenées à légiférer. Pour des discussions à ce sujet, nous encourageons la lectrice ou le lecteur à lire le rapport de 2017 de la CNIL[7].

Par opposition, l'idée d'*intelligence symbolique* est d'essayer créer un raisonnement juridique, ou encore *une preuve*, grâce à un formalisme mathématique, la *logique des prédicats* (voir Section 2.1). Il est alors possible pour l'humain·e de suivre et interpréter ce raisonnement pour dire si il y a erreur ou non. Quand il est question de formaliser une logique juridique, c'est à ce deuxième cas que l'on s'intéresse : une intelligence symbolique.

Pour discuter de l'éventuelle formalisation mathématique de la logique juridique, nous proposons de procéder en deux temps. Dans un premier temps, nous nous interrogerons sur le point de savoir si les systèmes juridiques sont axiomatisables. Dans un second temps, nous étudierons l'applicabilité des systèmes de déduction au raisonnement juridique.

7. « Comment permettre à l'Homme de garder la main ? », Rapport CNIL, Décembre 2017 `https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf`
8. Définition du Larousse, Oct. 2018

1 Les systèmes juridiques sont-ils axiomatisables ?

Avant de réfléchir à l'axiomatisation des systèmes juridiques, nous allons essayer de définir un cadre semi-formel aux systèmes de preuves mathématiques.

Deux blocs de base d'une preuve mathématiques sont les suivants :

— Les axiomes

— Le système de preuve, de déduction

Évidemment d'autres éléments rentrent en compte dans une preuve mathématiques, mais il est important de comprendre ces deux blocs pour comprendre la formalisation de la logique mathématiques. Dans cette section nous allons discuter ce que sont les axiomes et certaines propriétés d'un système axiomatique. Nous discuterons dans la Section 2 des systèmes de preuve.

1.1 Les axiomes et systèmes axiomatiques

Dans un système mathématique, les **axiomes** sont les propositions **non démontrées** du système étudié. Elles sont donc admises et servent de support de construction à tous les résultats.

On peut prendre comme exemple dans le modèle de géométrie vue en classe de 6è l'axiome :

Deux droites parallèles ne sont pas sécantes.

Attention, il faut faire une différence entre la *définition* des droites parallèles, par exemple : «Deux droites sont dites parallèles si elles ont la même direction.», et un axiome qui est une propriété de ces droites parallèles (ici, ne pas être sécantes).

La différence entre un axiome et une définition est qu'ici nous pouvons décider de supprimer cet axiome de notre modèle mathématiques. En effet, tout une branche des mathématiques est basé sur le fait qu'on puisse avoir des droites parallèles sécantes et non confondues ![9]

La conséquence est que certaines propriétés vraies dans un système axiomatique (par exemple *Par un point extérieur à une droite, il passe toujours une parallèle à cette droite, et une seule*[10] dans notre géométrie de 6è) sont fausses dans d'autres systèmes !

De manière générale et ce qu'il faut retenir, c'est qu'il n'y a pas de résultats vrais ou faux en mathématiques. Il y a des résultats vrais dans des systèmes axiomatiques précis, mais qui peuvent être faux dans d'autres systèmes axiomatiques.

C'est quelque chose qu'on pourra retrouver dans le législatif, on peut imaginer un système juridique ayant pour axiome « Il ne faut pas voler.». Dans ce système, jamais cet axiome ne pourra être remis en cause (par définition de

9. On appelle cette branche la Géométrie non Euclidienne `https://fr.wikipedia.org/wiki/Géométrie_non_euclidienne`

10. À noter que dans certains systèmes axiomatiques, cette propriété est considérée comme l'axiome de base. Plus d'informations à ce sujet sont disponibles `https://fr.wikipedia.org/wiki/Axiome_des_parallèles`

l'axiome). On aura naturellement des lois du type : « Il est interdit de voler, quitte à mourir de faim.». Dans un autre système juridique, on peut imaginer l'axiome : « On a le droit de ne pas mourir de faim.». Dans ce nouveau système, une loi dira : « On a le droit de ne pas mourir de faim, quitte à voler.». La question naturelle est quels sont les axiomes de notre système. Nous discuterons de ça dans la Section 1.2.

Se posent alors deux questions :

— Que se passe-t-il si on a les deux axiomes : « Il ne faut pas voler.» et « On a le droit de ne pas mourir de faim.» dans notre système ?

— Est-ce que je peux répondre à toutes les questions dans chaque système ?

Incohérence L'une des questions naturelles d'un système axiomatique est sa cohérence. Que ce passe-t-il si l'un des axiome est « Il ne faut pas voler.» et un autre axiome est « Il faut voler.». Dans ce cas là on dit que le système est *incohérent*, c'est à dire qu'il permet de prouver n'importe quoi (et donc qu'en général il est inutile mathématiquement).

Dans le cas de « Il ne faut pas voler.» et « Il faut voler.» l'incohérence est une conséquence directe des axiomes. L'incohérence d'un système n'est pas toujours simple à montrer. Par exemple dans un système avec deux axiomes : « Il ne faut pas voler.» et « On a le droit de ne pas mourir de faim.», on pourrait imaginer un raisonnement qui arrive au résultat : « Il est autorisé de voler si c'est pour ne pas mourir de faim et que toutes autres solutions ont été envisagées.». Ce résultat contredit l'axiome 1 et dans ce cas on revient au fait que notre système est incohérent.

Indécidabilité Dans les écrits juridiques on trouve certaines mésinterprétations de la logique mathématiques, qui justifieraient pourquoi elle ne peut s'appliquer à la logique juridique. La première (et la plus fréquente ?), serait que la logique mathématiques est binaire : vrai ou faux.

En fait c'est beaucoup plus compliqué que ça. Sachant un système axiomatique, il y a effectivement des résultats vrai et des résultats faux. Mais il y a aussi des résultats qui ne sont ni vrai ni faux, ils sont indécidables : on ne peut pas décider si les résultats sont vrais ou faux au vu des axiomes du système. On dit que le système est incomplet.

Concrètement, imaginons un système où l'unique axiome est « On a le droit de ne pas mourir de faim.». Dans ce système, on peut se demander si voler est autorisé ou non. On a vu qu'on peut sûrement montrer que dans ce système (i) on peut voler pour ne pas mourir de faim. On doit aussi pouvoir montrer (ii) qu'on ne peut pas voler si notre vol entraîne que la victime du vol va mourir de faim. Mais en dehors de ces deux cas, est-il autorisé de voler ? A priori le simple système axiomatique ne permet pas de trancher, et donc ce n'est ni autorisé ni interdit. Le système est incomplet.

Au passage, un résultat très fort montré par Kurt Gödel est que *la plupart* des sytèmes mathématiques sont incomplets [11], c'est à dire qu'il existe dedans

11. Pour un énoncé plus précis, voir : `https://fr.wikipedia.org/wiki/Théorèmes_d'`

des énoncés qui ne sont ni démontrables (vrais), ni réfutables (faux) ! Dans ce cas là, on peut prendre cet énoncé comme ... axiome de notre nouveau système [12] !

C'est ce qu'il se passe en terme de droit, lorsque les axiomes ne permettent pas de trancher : une nouvelle règle est ajoutée au système.

1.2 Peut-on axiomatiser un système juridique ?

La question de l'axiomatisation du système juridique a déjà été traitée en doctrine, qui considère souvent que celle-ci n'est possible que dans une certaine mesure. Cependant, il n'est pas certain que les concepts de la logique mathématique pertinents aient été systématiquement employés à bon escient. Par exemple, si l'on admet que l'axiomatisation d'un système logique requiert l'admission de termes premiers indéfinissables, il peut être tentant d'assimiler à ces derniers, dans le champ juridique, les notions dites « floues » [13]. Le droit connaît un certain nombre de notions floues ou indéterminées qui sont pourtant centrales dans la mécanique du droit. On peut citer de nombreux exemples : la bonne foi, les bonnes mœurs, l'ordre public, etc. Il existe même des notions que l'on a pris l'habitude de qualifier de « fonctionnelles », par emprunt à Georges Vedel, c'est-à-dire des notions que l'on est en peine de définir logiquement, qui sont vagues, contradictoires et sans unité, excepté la fonction qu'elles jouent dans le droit [14]. Cependant, ces termes « flous » sont peut-être moins des termes « premiers » au sens de la logique mathématique, des termes non définis et indéfinissables, que des termes mal définis – et ce, à dessein, afin de permettre leur adaptation à des contextes non nécessairement anticipés a priori.

Mais c'est surtout l'identification des « propositions premières », ou axiomes, qui est délicate dans l'axiomatisation du système juridique. Certains auteurs [15] ont pu considérer comme des axiomes certaines notions utilisés par les théoriciens du droit comme des « clés de voûte » de leur système de pensée.

A ce titre, l'une des difficultés récurrentes de la théorie du droit est en effet d'expliquer à quel moment le « non-droit » devient du droit. On peut illustrer cette difficulté par la théorie du droit de Hans Kelsen, l'un des théoriciens du droit les plus influents du XXème siècle. Son but théorique a été de construire ce qu'il qualifiait lui-même de « théorie pure du droit » [16], c'est-à-dire un sys-

incomplétude_de_Gödel

12. Le mot de la fin de cette section est pour Zach Weinersmith (2018) : https://www.smbc-comics.com/comic/problem

13. V. not Buffelan J.-P., « Le droit, l'informatique et la mathématique », Journal de la société statistique de Paris, tome 115 (1974), p. 301.

14. Vedel G., « De l'arrêt Septfonds à l'arrêt Barinstein (La légalité des actes administratifs devant les tribunaux judiciaires) », JCP 1948. I. 682 ; « La juridiction compétente pour prévenir, faire cesser ou réparer la voie de fait administrative, JCP, 1950. I. 851. Pour une analyse critique de l'utilisation de la notion de « notion fonctionnelle » v. Tusseau G., « Critique d'une métanotion fonctionnelle. La notion (trop) fonctionnelle de notion fonctionnelle », Revue française de droit administratif, 2009, p. 641.

15. Jobart J.-Ch., « La logique juridique : de l'ordre et du désordre au chaos », Droit prospectif, revue de la recherche juridique, 2006-4 (1), p. 1873.

16. C'est le nom de son ouvrage le plus célèbre : Théorie pure du droit, traduction Ch. Eisenmann, Paris, Dalloz, 2ème édition, 1962.

tème théorique qui permet d'identifier et de décrire un système juridique sans avoir recours à un quelconque système de valeurs (ceci est bien, ceci est juste) mais uniquement en ayant recours à une structure formelle. Dans son approche, une norme juridique est valide dès qu'elle a été adoptée conformément à une autre norme, elle-même valide. Il y a donc dans la théorie kelsénienne une chaîne de régression de la validité, qui implique une hiérarchie entre les normes juridiques (une norme adoptée conformément à une autre norme lui est inférieure). L'ensemble des normes existant à un moment donné forme donc une pyramide. Cette image de la pyramide est la plus célèbre de la « géométrie juridique » et est infligée aux étudiants dès la première année de droit. Ladite pyramide remonte jusqu'à la norme « suprême », la Constitution. Mais le problème de cette approche, c'est que la Constitution, pour être une norme juridique valide, doit elle-même être conforme à une norme supérieure – sauf à considérer que la Constitution n'est pas elle-même une norme juridique. Pour résoudre cette aporie, Kelsen a posé le concept de Grundnorm. Cette Grundnorm n'est pas une norme juridique réelle et identifiable, comme l'est la Constitution, mais une norme hypothétique fondamentale, de nature logico-transcendantale, considérée comme une supposition nécessaire de l'esprit juridique pour assurer la cohérence de l'ordre juridique.

La Grundnorm est-elle alors un axiome ? Cela n'est pas certain. Un axiome doit pouvoir servir de base à une démonstration – et donc, adapté au système juridique, permettre de déterminer qu'une norme est valide ou bien qu'elle ne l'est pas. Or, la Grundnorm, parce qu'elle n'a pas de contenu, ne peut fonder une démonstration logique : il est impossible de prouver qu'une norme est invalide à l'aide de la Grundnorm. Elle est davantage une hypothèse nécessaire pour délimiter un système juridique qui est clos (sa validité ne dépend de rien d'autre que de lui-même) et complet (toutes les normes du système juridique sont des normes juridiques, y compris la norme suprême). Elle s'apparente donc plus, d'un point de vue logique, à une décision : « on » décide que le système juridique existe.

Y a-t-il alors de véritables axiomes dans le système juridique ? Si l'on considère comme axiome dans le système toute proposition qui est donnée mais non démontrable, alors on peut considérer comme axiome toute proposition juridique adoptée à un moment donné, qui s'inscrit dans le champ des possibles délimité par les propositions juridiques préexistantes mais n'en découle pas inéluctablement. C'est l'hypothèse, courante en droit, de la marge d'appréciation ou du pouvoir discrétionnaire en droit administratif : il est rare qu'une norme juridique prédétermine totalement les normes juridiques subséquentes (cela peut arriver : on parle, dans un tel cas, de « compétence liée »). Souvent, les normes supérieures déterminent la procédure d'adoption des normes inférieures, posent des interdits, fixent des objectifs voire des principes. L'adoption de la norme inférieure implique alors une décision, laquelle ajoute alors un nouvel axiome dans le système. Le système juridique peut alors être considéré comme axiomatique, à condition d'accepter que les axiomes dont il est question ne sont pas fixées une fois pour toutes « à l'origine » mais évoluent constamment, certains apparaissant, d'autres disparaissant, le tout restant cohérent à condition de demeurer

dans le champ des possibles délimité par le système. Sur la base de ces axiomes, il est possible de déduire des théorèmes posant des règles juridiques non axiomatiques mais résultant des axiomes. Ainsi, s'il existe une règle (axiome) selon laquelle toute personne majeure ayant la nationalité française a le droit de vote, sauf privation de droits civiques, et une autre règle (axiome) selon laquelle l'âge de la majorité en France est 18 ans, alors il est possible de construire un théorème selon lequel toute personne âgée d'au moins 18 ans ayant la nationalité française a le droit de vote, sauf privation de droits civiques.

On peut probablement identifier un certain nombre d'axiomes structurels, c'est-à-dire un certain nombre de proposition juridiques qui ne sont pas elles-mêmes démontrables sur la base d'autres propositions du système, et qui donc en ce sens sont décidées mais non démontrées, et qui sont essentielles à son bon fonctionnement. On peut ici citer les règles relatives aux résolutions des conflits entre normes juridiques. Que faire, en effet, lorsque deux normes juridiques se contredisent mutuellement ? La réponse a été systématisée, notamment, par le théoricien du droit italien Norberto Bobbio [17]. D'abord, il convient de regarder si l'une des normes n'est pas « spéciale » par rapport à l'autre : par exemple, une norme professionnelle applicable aux médecins libéraux uniquement est spéciale par rapport à une norme applicable aux professions libérales en général. Dans ce cas, la règle spéciale peut déroger à la règle générale. Ensuite, si un tel rapport général / spécial n'existe pas, il convient de faire jouer le rapport hiérarchique : la norme supérieure l'emporte sur la norme inférieure (par exemple, une loi sur un décret). Enfin, si les deux normes antagonistes sont de même niveau hiérarchique, la norme la plus récente prévaut sur la norme la plus ancienne. Ces trois règles de résolution des conflits de normes juridiques, ainsi par ailleurs que l'ordre de leur examen, peuvent être considérés comme des axiomes du système juridique.

Reste une dernière question : le système juridique est-il cohérent ? Il faut ici examiner une objection qui découle, là encore, de la définition de la validité chez Kelsen. Chez Kelsen, une norme est valide à partir du moment où elle a été produite par l'autorité compétente. Une telle norme peut cependant être incompatible avec une norme supérieure, en ce que la prescription qu'elle contient est contraire à une prescription de rang supérieur, ou encore parce que les formes de son adoption ne sont pas conformes aux formalités imposées à un rang supérieur. Pour autant, chez Kelsen, cette norme demeure valide tant qu'elle n'a pas été « purgée » du système selon les modalités que celui-ci prévoit (par exemple, un décret contraire à une loi peut être contesté devant le juge administratif, lequel peut alors l'annuler). Cela ne génère pas une véritable incohérence dans le système, pour deux raisons. D'une part, d'un point de vue dynamique, il existe précisément la plupart du temps des moyens de purger les contradictions. D'autre part, même lorsqu'il n'en existe pas (ou, ce qui revient au même, qu'il en existe mais qu'ils n'ont pas été utilisés dans le délai imparti), on peut considérer qu'il existe une règle permettant de telles contradictions.

17. Bobbio N., « Des critères pour résoudre les antinomies », Communication faite à Bruxelles, au Centre national de recherches de logique, le 2 février 1963, texte reproduit in Essais de théorie du droit, Bruxelles et Paris, Bruylant et LGDJ, 1998, p. 89.

2 L'applicabilité des systèmes de déduction au raisonnement juridique

L'une des autres pierres angulaires du système de raisonnement mathématique avec les axiomes est le système de déduction. Il sera présenté brièvement dans les développements qui suivent, avant que de s'interroger sur son applicabilité au raisonnement juridique.

2.1 Les systèmes de déduction

Comment, formellement à partir de plusieurs axiomes puis-je obtenir un résultat ? Pour cela on utilise des règles d'inférences, des règles de raisonnement admises. Pour une liste complète des règles d'inférences, le mieux est de suivre un cours ou de lire la page wikipedia `https://fr.wikipedia.org/wiki/Règle_de_résolution`.

Nous présentons deux exemples ici.

Le Modus Ponens :

$$((A \implies B) \wedge A) \implies B$$

Cette règle peut être comprise : si on sait que le fait A entraîne le fait B, **et** qu'on a le fait A, alors on a le fait B.

Le Tiers Exlus :

$$(A \vee \bar{A})$$

Cette règle peut être comprise : soit A est vrai, soit A est faux. C'est le principe de toute preuve par l'absurde, supposons une propriété vraie, montrons que l'on obtient une contradiction, ça montrera que la propriété initiale ne peut pas être vraie. Par le principe du tiers exclu, elle est donc fausse.

Ce qui est intéressant c'est que ces règles d'inférences paraissent triviales mais ont mis longtemps à être formalisées correctement [18]. L'autre chose intéressante est que même si ces règles paraissent triviales pour n'importe quel raisonnement logique, elle reste des axiomes, c'est à dire sans preuves. Il y a d'ailleurs toute une branche des mathématiques [19] qui s'intéresse à l'intérêt du tiers-exclu : quels sont les résultats qui sont prouvables sans tiers-exclu ? Est-ce qu'ils sont les mêmes que ceux prouvables avec le tiers exclu ?

Un *raisonnement logique* est ensuite une combinaison des résultats vrais (ou faux) du système, en partant des axiomes et des résultats qui ont pû être montrés, combinés aux règles autorisées dans notre système de déduction.

Petite parenthèse technique, là où ça devient intéressant, c'est que techniquement un système de déduction est un ensemble d'*axiomes* (par exemple le

18. Pour une histoire de la formalisation de la logique, on pourra lire *Logicomix* de Apostolos, Papadimitriou, Papadatos, et Di Donna. (2010).

19. Logique classique contre logique intuitionniste

Modus Ponens). Ces axiomes étant admis (par définition), on peut les utiliser comme système de déduction pour créer des théorèmes, qui sont d'autres méthodes de notre système de déduction. Toute une branche des mathématiques est basée sur ces réflexions [20].

2.2 Application des connecteurs logiques au raisonnement juridique

Certains connecteurs logiques s'intègrent assez harmonieusement dans le raisonnement juridique. De façon tout à fait évidente, la conjonction et la disjonction peuvent s'appliquer aux critères cumulatifs / alternatifs en droit. Les critères d'application de la règle de droit sont dits cumulatifs quand il faut que toutes les conditions posées par la règle soient réunies pour que la règle produise ses conséquences. Par opposition, les critères d'application de la règle sont dits alternatifs quand il suffit qu'une des conditions posées par la règle soit réalisée pour que la règle s'applique. De même, la récurrence s'applique assez bien au phénomène des « illégalités en cascade ». Il s'agit là d'une conséquence du caractère hiérarchisé du système juridique, dont il résulte que l'illégalité d'une norme, dans certains cas, implique l'illégalité de toutes les normes inférieures qui en découlent. On peut donner un exemple assez caricatural : une personne A passe un concours pour devenir fonctionnaire et réussit. Elle adopte alors un certain nombre de décisions. Elle peut habiliter une autre personne à prendre des décisions en son nom, ou bien il se peut que des fonctionnaires subalternes prennent des décisions sur la base de la décision adoptée par A. Or, le décret organisant le concours qu'a passé A est déclaré illégal et est annulé ou invalidé. L'annulation / invalidation est rétroactive en principe : A n'a jamais été fonctionnaire, il n'a jamais eu le pouvoir de prendre les décisions qu'il a prises, qui sont donc elles aussi nulles pour motif d'incompétence, et la nullité de ces dernières contamine à son tour toutes les décisions prises sur cette base. Nous qualifions cet exemple de « caricatural » car, en réalité, il existe des moyens de limiter cet effet domino, notamment la théorie du fonctionnaire de fait [21] ou la modulation des effets dans le temps d'une annulation [22]. Il n'en reste pas moins valide en théorie.

Mais c'est surtout l'implication qui est précieuse en droit, où elle s'applique sous la forme du syllogisme judiciaire, c'est-à-dire l'application de la règle à un cas donné :

$$\Big(\text{majeure} \bigwedge \text{mineure}\Big) \implies \text{conclusion}.$$

Exemple : le vol est puni de prison, X a commis un vol donc X va en prison. La majeure est la règle. On pourrait considérer qu'elle est vraie par hypothèse et ne pas l'intégrer dans la formule (X a commis un vol $\implies$ X va en prison).

20. Pour plus d'information, je recommande la page `https://fr.wikipedia.org/wiki/Calcul_des_propositions`.

21. Cour de cassation, Chambre civile, 7 août 1883, Sirey 1884-1, p. 17. Dans le même sens, Conseil d'Etat, 2 nov. 1923, *Association des fonctionnaires de l'administration centrale des postes et télégraphes*, Leb. p. 699.

22. CE, Ass., 11 mai 2004, *Association AC !*, req. n° 255886.

Cependant, ce n'est pas si simple. Par exemple, si la règle est incompatible avec une règle supérieure (ex : une loi contraire à la Constitution ou à une convention internationale) elle n'est pas applicable. A toutes fins utiles, on peut la considérer comme « fausse » dans le cadre d'un tel raisonnement. La mineure est l'inconnue, que doit déterminer le juge ou le jury.

Il est à noter que la majeure elle-même peut être considérée comme une proposition avec une implication : le vol est puni d'une peine de prison par exemple : Vol $\implies$ prison. Ce qui donne ((Vol $\implies$ prison) $\bigwedge$ « X a commis un vol ») $\implies$ X va en prison. Autre exemple : en vertu de l'article 1240 du code civil, « Tout fait quelconque de l'homme, qui **cause** à autrui un **dommage**, oblige celui par la **faute** duquel il est arrivé à le réparer », ce qui peut être formalisé ainsi : ((Faute $\implies$ dommage) $\implies$ obligation de réparation). Appliqué dans un cas concret, (((Faute $\implies$ dommage) $\implies$ obligation de réparation) $\bigwedge$ « X a commis une faute » $\bigwedge$ « Y a subi un dommage » $\bigwedge$ « la faute de X a entraîné le dommage de Y ») $\implies$ X doit réparer le dommage de Y.

Et l'on peut encore complexifier, car la mineure est en réalité double : c'est à la fois le fait et la qualification juridique du fait. On peut alors essayer de synthétiser tout cela dans la formule générale suivante :

$$\Big(\big(\text{condition d'application de la règle} \implies \text{conséquence que la règle y attache}\big) \bigwedge \text{le fait x est établi} \bigwedge \text{le fait x peut être qualifié juridiquement comme correspondant à la condition d'application de la règle}\Big) \implies \text{la conséquence prévue par la règle est applicable.}$$

Pour autant, l'extension des connecteurs logiques au raisonnement juridique est-elle toujours possible ? C'est là une question qui agite les théoriciens du droit de longue date. Des auteurs comme Georges Kalinowski ont dédié leurs recherches à l'étude de la logique déontique, c'est-à-dire pour résumer l'application de la logique formelle aux systèmes normatifs. D'autres auteurs, comme le théoricien du droit belge Chaïm Perelman, ont considéré que le raisonnement juridique n'était pas soluble dans la logique mathématique, notamment en opposant déduction et argumentation. Dans un article publié en 1960 dans la revue Logique & Analyse [23], et dans la continuité de son Traité de l'argumentation [24], il a estimé que l'argumentation n'est pas un simple calcul mais l'appréciation de la force d'un raisonnement découlant de la pesée, à la fois subjective et contingente, d'arguments qui ne peuvent pas être appréciés de façon binaire en « vrai » ou « faux » mais de façon graduée : « Un argument n'est pas correct et contraignant ou incorrect et sans valeur, mais est relevant ou irrelevant, fort ou faible, en fonction de raisons justifiant son emploi en l'occurrence ».

Il n'est pas question de trancher ici ce débat. L'on se contentera ici de réfléchir à l'applicabilité des connecteurs logiques aux trois types d'argument les

23. Perelman Ch., « Logique formelle, logique juridique », Logique Et Analyse 11, 1960, p. 226.

24. Olbrechts-Tyteca L. et Perelman Ch., Traité de l'argumentation, la nouvelle rhétorique, PUF, 1958, rééd. Université de Bruxelles, 2008.

plus classiques du raisonnement juridique (et plus précisément judiciaire), ceux que l'on utilise quand il n'existe pas de règle clairement applicable au litige : l'argument a pari causa, l'argument a fortiori et l'argument a contrario.

L'argument par analogie, ou argument a pari causa consiste à appliquer à une situation non régie par un texte les dispositions applicables à une situation analogue. Par exemple, dans la mesure où l'annulation d'un mariage « ressemble » au divorce, la prestation compensatoire prévue en cas de divorce peut s'appliquer en cas d'annulation. Autre exemple : si les voitures de métro sont interdites aux chiens et aux chats, elles le sont aussi aux lapins et aux furets. Cet argument n'est pas utilisé dans toutes les disciplines juridiques, et notamment pas en droit pénal, en raison du principe d'interprétation strict des textes.

L'argument par analogie rend aussi compte, dans une certaine mesure, de l'influence de la jurisprudence. Tout juge s'efforce en effet, en principe, d'être fidèle à la jurisprudence. Cela signifie que la solution définitive qui a été donné à un problème juridique dans un cas précédent (de préférence par une cour suprême) doit en principe être adoptée dans un cas similaire. C'est sur ce principe-là, en particulier, que s'appuient les solutions de justice prédictive.

L'argument « a fortiori » permet d'étendre une règle, à une hypothèse non prévue mais où elle se justifierait a plus forte raison. Par exemple, soit une règle disant qu'il est possible de tuer quelqu'un pour protéger sa vie. A fortiori, cela implique qu'il est possible de blesser quelqu'un (ce qui est moins grave) pour protéger sa propre vie.

L'argument « a contrario » mène à considérer que lorsque le législateur applique une règle particulière à un cas déterminé, cette règle ne s'applique pas aux cas non prévus. Ainsi, selon l'article 6 du code civil, « on ne peut déroger, par des conventions particulières, aux lois qui intéressent l'ordre public et les bonnes mœurs ». A contrario, cela signifie qu'on peut déroger par conventions particulières aux lois qui n'intéressent ni l'ordre public ni les bonnes mœurs.

Ch. Perelman, dans l'article précité, a pu mentionner ces arguments pour estimer que le raisonnement juridique n'est pas réductible à la logique mathématique : « Ces arguments ne peuvent, en effet, servir à une démonstration rigoureuse et l'on ne voit pas de machine capable de les manier, car leur usage nécessite, chaque fois, une prise de position, qui justifierait leur application dans des circonstances déterminées ». C'est sans doute vrai. Pour autant, il est peut-être possible de structurer cette prise de position dans un cadre plus ou moins formalisé.

Par exemple, l'argument a pari causa est implicitement fondée sur le postulat que la règle dont on étend l'application à un cas non prévu est la manifestation, ponctuelle, d'une règle ou d'un principe plus général. L'identification de ce principe est probablement une prise de position. Mais une fois celui-ci identifié, c'est à nouveau le syllogisme standard qui s'applique. L'argument a fortiori

est aussi une forme de raisonnement par implication, qui implique cependant au préalable de formaliser les règles permettant d'identifier le "a fortiori" (par exemple, la règle selon laquelle "tuer" est plus grave que "blesser").

Enfin, une structuration un peu formalisée de l'argument a contrario peut permettre d'en déminer les pièges, que le juriste connaît bien. L'argument a contrario est en effet toujours d'un maniement délicat : ce n'est pas parce que la règle 1 s'applique dans le cas A qu'elle ne s'applique pas dans le cas non-A. Souvent, l'utilisation (ou non) de l'argument a contrario vient en renfort, et nécessite une réflexion sur la règle elle-même à qui on souhaite l'appliquer. Cela nécessite une appréciation assez difficilement automatisable. Pour autant, cette appréciation peut éventuellement être informée par des connecteurs logiques. Ainsi, l'applicabilité ou non de l'argument a contrario peut être guidée par les notions d'implication et d'équivalence : si le juge, eu égard au contexte de la règle et à son but, estime qu'elle pose une implication, alors le raisonnement a contrario ne fonctionne pas. En revanche, si la règle juridique pose une équivalence, alors le raisonnement a contrario marche. Là encore, cependant, les connecteurs ne permettent pas de répondre à la question, mais peuvent la reformuler, sous une forme de guide qui peut être standardisé.

3 Conclusions et directions

À notre sens, le problème de l'écriture formelle de la logique juridique est un non-problème qui pourra être résolu rapidement. Il fut un temps pas si lointain (1900) où on prenait les personnes qui entendaient formaliser la logique mathématique pour des fous, puis vinrent Whitehead et Russell et leurs Principia Mathematica. Cela veut aussi dire qu'avant , la logique mathématique n'était pas rigoureusement formalisée !

Les difficultés qui se posent sur l'automatisation de la loi (traitement automatique du langage, création des "démonstrations", etc) sont des problèmes qu'on peut réduire à d'autres problématiques d'IA (en particulier sur le raisonnement en IA), et non pas à la formalisation du raisonnement juridique. Nous considérons cela aussi comme un non-problème, non pas parce que ce n'est pas difficile, mais parce que beaucoup de gens s'attellent à ça, et ce ne sont pas des problèmes spécifiques au droit.

Il faut maintenant se concentrer sur les difficultés propres au système juridique, parmi lesquelles on trouve le biais des bases de données (pour une IA basée sur le "Big Data"). Par exemple, la loi pour une république numérique précise que « la mise à disposition du public est précédée d'une analyse du risque de ré-identification des personnes ». Or, c'est là une réserve potentiellement importante, selon comment elle est interprétée : n'est-il pas finalement quasiment toujours possible d'identifier les protagonistes, même après anonymisation, sur la base des éléments du litige ? Et si l'anonymisation « réelle » aboutit à gommer des éléments du litige qui sont pertinents pour la prédiction, comme par exemple les fonctions des protagonistes, ne perd-on pas alors en fiabilité ? D'autre part, toutes les juridictions ne sont pas concernées : les juridictions administratives

spécialisées, par exemple, semblent exclues. Or, pour certaines – on pense notamment à la Cour nationale du droit d'asile – un outil prédictif serait tout à fait précieux et protecteur de la sécurité juridique d'une catégorie de justiciables particulièrement vulnérables, à savoir les demandeurs d'asile. Pour autant, en cassation, les décisions des juridictions administratives spécialisées sont examinées par le Conseil d'Etat, dont les décisions, elles, seront rendues publiques. Est-ce que cela peut générer un biais ? Faut-il exclure le contentieux des juridictions administratives spécialisées de la justice prédictive dans la mesure où ce contentieux ne sera pas parfaitement public ? Peut-on aménager cette asymétrie vraisemblable de données ? S'y ajoutent les problématiques de "comprendre une décision" vs "accepter une décision". Aujourd'hui, en deep learning / réseaux de neurones, il est admis que les algorithmes prennent d'excellentes décisions, mais que personne ne sait ce qui a poussé à prendre la décision. Une décision de justice "assistée par ordinateur" peut-elle être acceptable si ses motivations sont impénétrables, quand bien même on aurait la certitude qu'il s'agit d'une "bonne" décision ?

Notes finales

Ce document a été fait dans le cadre des journées de Convergence du Droit et du Numérique de l'université de Bordeaux `http://cdn.u-bordeaux.fr/`, n'a pas été relu et a une vocation de vulgarisation. Guillaume Aupy n'est pas expert en logique mathématique et il est possible que quelques faux sens se soient glissés. Pour aller plus loin scientifiquement, il existe de nombreuses personnes travaillant sur la formalisation du droit (voir par exemple les publications de l'association International Association for Artificial Intelligence and Law `http://www.iaail.org/`.

Vers une remise en cause de la légalité du FNAEG ?

Ousmane Gueye* François Pellegrini†

Introduction

La naissance de la police scientifique, à la fin du XIXe siècle, est indissociable de celle de la biométrie. L'objectif des travaux de précurseurs tels que Bertillon [1] et Vučetić était de quantifier les caractéristiques physiques des individus et d'en établir une classification efficace. Le but de cette classification était, d'une part, de permettre la transmission de signalements fiables entre brigades et, d'autre part, d'accélérer l'identification des criminels récidivistes par comparaison entre les signalements et/ou traces laissées sur des scènes de crimes et les données de référence contenues dans des fichiers de police (les fameux « sommiers »). Cette quantification, sous forme nativement numérique (taille, poids) ou de valeurs discrètes (couleur des yeux, type de pigmentation, catégories de minuties des empreintes digitales), se prête naturellement à un codage sous forme numérique. C'est ainsi que l'essor de la mécanographie (traitement mécanisé de l'information), entamé sensiblement à la même période, a pu contribuer à nourrir les projets sociétaux de numérotation et de catalogage des personnes qui ont caractérisé le XXe siècle. La naissance de l'informatique moderne, au tournant de la Seconde Guerre mondiale, a accru de façon significative les capacités de traitement, permettant la réalisation de « croisements » entre fichiers, là où la mécanographie ne permettait la réalisation que de simples tris. C'est d'ailleurs en réaction à la tentative de croisement, par l'administration, de l'intégralité des fichiers en sa possession, au moyen d'un identifiant unique des personnes, dans le cadre du projet SAFARI [2], que fut créée en 1978 la Commission nationale de l'informatique et des libertés (CNIL).

La découverte de l'ADN [3] en tant que support principal de l'hérédité, et l'évolution des techniques de séquençage permettant d'exposer le code porté par les chromosomes, a profondément révolutionné la biométrie. Toute trace corporelle d'une personne contenant du matériau génétique peut désormais être associée à un individu donné avec une probabilité presque certaine. L'usage de la biométrie génétique pour créer un fichier de police a donc émergé, poussé tant par l'abaissement du coût des analyses que par l'émoi suscité par des crimes particulièrement abjects.

C'est ainsi qu'en 1998 a été créé en France le Fichier national automatisé des empreintes génétiques (FNAEG) [4]. Comme en d'autres occasions, le législateur a entendu renforcer les pouvoirs procéduraux

*Doctorant en droit du numérique, Université de La Rochelle, Centre d'Études Juridiques et Politiques, Faculté de Droit, Science Politique et Gestion, 45 rue François de Vaux de Foletier, 17024 La Rochelle cedex 1, France. ousmane.gueye@univ-lr.fr

†Professeur d'informatique, Université de Bordeaux, LaBRI & INRIA Bordeaux Sud-Ouest, 351 cours de la Libération, 33405 Talence, France. francois.pellegrini@labri.fr

1. Voir par exemple : « Alphonse Bertillon et l'identification des personnes (1880-1914) », Criminocorpus, https://criminocorpus.org/fr/expositions/suspects-accuses-coupables/alphonse-bertillon-et-lidentification-des-personnes-1880-1914/. Consulté le 9 septembre 2017.

2. Le projet SAFARI (« Système automatisé pour les fichiers administratifs et le répertoire des individus ») visait à l'interconnexion des fichiers nominatifs de l'administration française. Il fut révélé au public en mars 1974 par le quotidien *Le Monde*, dans l'article intitulé « SAFARI ou la chasse aux Français » de Philippe BOUCHER.

3. Pour « acide desoxyribo-nucléique ». L'ADN est constitué de deux brins enroulés en double hélice, chaque brin codant l'information génétique sous la forme de séquences construites au moyen des quatre bases nucléiques identifiées par les lettres A, C, G et T.

4. Loi n° 98-468 du 17 juin 1998 relative à la prévention et à la répression des infractions sexuelles ainsi qu'à la protection des mineurs.

de la police judiciaire en lui accordant des moyens dérogatoires du droit commun [5]. Depuis lors, les lois destinées à renforcer les prérogatives de la police judiciaire n'ont jamais cessé de se succéder, chacune visant à étendre le périmètre et les moyens d'actions des enquêteurs dans la recherche d'éléments de preuve [6], au détriment des libertés individuelles [7]. Ainsi, une loi de 2003 a permis d'inscrire sur ce fichier tout individu suspecté d'un simple délit [8], conduisant à une augmentation substantielle de sa taille. En 2016, le nombre d'échantillons contenus au sein du fichier était de plus de 3 422 786 [9], représentant près de 5 % de la population française.

Les données génétiques sont des données biométriques très particulières, en ce qu'elles ne renseignent pas seulement sur la personne qui en est le porteur, mais sur l'ensemble de sa lignée : ascendants, descendants et collatéraux. Les techniques de séquençage ne permettent donc pas seulement d'identifier une personne, mais aussi de donner, avec une très forte probabilité, son lien de parenté avec un échantillon de référence. Cette technique, dite de recherche « en parentèle », est occasionnellement utilisée dans le cadre de litiges en paternité ou en succession. Elle ne faisait pas partie des modalités d'usage initiales du FNAEG, celui-ci ayant été présenté comme uniquement destiné à renvoyer une réponse positive ou négative (« *hit / no hit* ») quant à la présence d'un profil génétique dans le fichier. La recherche en parentèle fut pourtant mise en œuvre au sein du FNAEG en 2011, pour une recherche « en aveugle », à l'occasion d'une affaire qui a particulièrement ému l'opinion publique [10]. Ce n'est cependant qu'en 2016 que le législateur a autorisé cette pratique. Pour autant, cette autorisation suscite des interrogations au regard de l'article 9 du Code civil et de l'article 8 de la Convention de sauvegarde des droits de l'Homme et des libertés fondamentales (aussi dite « Convention européenne des droits de l'Homme »). La recherche en parentèle en aveugle dans le cadre d'enquêtes criminelles n'est-elle pas en contradiction avec la jurisprudence de la Cour européenne des droits de l'Homme (CEDH) ? La CEDH a déjà eu l'occasion de préciser que « *la protection offerte par l'article 8 de la Convention serait affaiblie de manière inacceptable si l'usage des techniques scientifiques modernes dans le système de la justice pénale était autorisé à n'importe quel prix et sans une mise en balance attentive des avantages pouvant résulter d'un large recours à ces techniques, d'une part, et des intérêts essentiels s'attachant à la protection de la vie privée, d'autre part* [...] » (§ 112) [11]. Le FNAEG, tel qu'il est actuellement utilisé, notamment dans le cadre de la recherche en parentèle en aveugle, ne porte-t-il pas atteinte de façon disproportionnée au droit à la vie privée et à la protection des données à caractère personnel ? Le FNAEG, malgré les efforts fournis par les autorités pour l'intégrer dans le dispositif légal à travers de multiples réformes suscitées par des événements particuliers, est utilisé pour atteindre des finalités différentes de celles qui lui sont attribuées par les textes qui l'instituent (partie II). Cette légèreté blâmable traduit la tentation du fichage généralisé qui inquiète les défenseurs des droits et libertés individuelles, surtout dans un contexte d'état d'urgence quasi permanent (partie I).

Partie I : La tentation du fichage généralisé

La révolution numérique a permis à l'État de moderniser son action, mais aussi de rendre plus efficace le contrôle qu'il exerce sur ses populations. Pour éviter que l'usage de l'informatique ne débouche sur des dérives et des pratiques attentatoires aux droits et libertés fondamentaux, une loi du 6 janvier

5. Code de procédure pénale, articles 706-73 à 706-102. Voir : C. LAZERGES, « La dérive de la procédure pénale », *Revue de science criminelle*, n° 3, 2003, pp. 644–654.

6. Loi n° 2004-204 du 9 mars 2004 portant adaptation de la justice aux évolutions de la criminalité, dans son article 49 ; loi du 12 décembre 2005 sur la récidive des infractions pénales, dans son article 18.

7. Loi n° 2001-1062 du 15 novembre 2001 relative à la sécurité quotidienne, dans son article 56.

8. La loi n° 2003-239 du 18 mars 2003 sur la sécurité intérieure (dite « loi Sarkozy »), dans son article 29, élargit l'usage du FNAEG à de simples délits (vol, tag, arrachage d'OGM, etc.) et permet aussi d'inclure de simples suspects.

9. Marion VAN RENTERGHEM, « La tentation du fichage génétique de masse », *Le Monde*, 25 septembre 2006, http://www.lemonde.fr/a-la-une/article/2006/09/25/la-tentation-du-fichage-genetique-de-masse_816576_3208.html .

10. Voir *infra*.

11. CEDH, Grande Chambre, 4 décembre 2008, S. et Marper c. Royaume-Uni, Req. n° 30562/04.

1978, dite « Informatique et Libertés », l'une des premières du genre, a été adoptée en France.

Aujourd'hui, le respect des dispositions de cette réglementation par la puissance publique se pose, en raison notamment de sa volonté de plus en plus manifeste de ficher toute la population au travers des fichiers qu'elle met en œuvre en un nombre sans cesse croissant. Ces fichiers font craindre l'ère d'une surveillance généralisée. Deux facteurs peuvent témoigner de ce fait : il s'agit d'une part de l'extension des fichiers administratifs en fichiers de police (I) et, d'autre part, de l'extension du fichier FNAEG aux personnes impliquées dans les délits banals (II).

1. Extension des fichiers administratifs en fichiers de police

Les pouvoirs publics et l'autorité judiciaire rivalisent d'ingéniosité pour collecter des données personnelles au sein de fichiers susceptibles d'usage à des fins de contrôle et de surveillance des individus. C'est ainsi que le fichier TES (« Titres électroniques sécurisés ») a été étendu pour regrouper, au sein d'une base de données centralisée, non seulement les informations nominatives de tous les Français détenteurs d'un passeport ou d'une carte nationale d'identité, mais aussi leurs données biométriques, telles que la photographie du visage et les empreintes digitales. Ce fichier n'est cependant que l'un des nombreux fichiers existentiels tenus par les autorités administratives ou les organismes publics et parapublics, auxquels s'ajoutent plus de 80 fichiers de police [12].

La nouvelle version du fichier TES constitue un dispositif particulièrement impressionnant, permettant en pratique la collecte des mêmes données et pour les mêmes usages que la loi de mars 2012 « relative à la protection de l'identité » [13] qui avait été censurée par le Conseil constitutionnel. Saisie suite à l'émoi suscité chez les défenseurs des libertés par l'instauration d'un fichier des « gens honnêtes », la Haute juridiction avait en effet déclaré contraires à la Constitution certaines dispositions de ladite loi, pour trois raisons liées au droit au respect de la vie privée :

- l'ampleur des données collectées sur « la quasi-totalité de la population française », et notamment les empreintes digitales, considérées comme « particulièrement sensibles » ;
- les caractéristiques techniques du fichier, qui le rendraient utilisable « à d'autres fins que la vérification de l'identité d'une personne » ;
- le fait que la loi autorise l'utilisation du fichier « à des fins de police administrative ou judiciaire ».

La lutte contre la fraude documentaire et l'usurpation d'identité, qui ont été invoquées pour justifier la nécessité du TES, constitue également l'une des finalités du Fichier automatisé des empreintes digitales (FAED). Ce fichier commun à la police et la gendarmerie nationales, créé par le décret n° 87-249 du 8 avril 1987 (modifié en 2005), permet :

- d'identifier les traces digitales et palmaires relevées sur les scènes d'infraction afin de rechercher et d'identifier les auteurs de crimes ou de délits ;
- de détecter les usurpations d'identité et les identités multiples.

L'enregistrement systématique dans le FAED de tout individu mis en cause dans une enquête, de façon directe ou indirecte, a lui aussi suscité des réactions visant à préserver les libertés publiques. Dans une décision du 18 avril 2013 [14], la Cour européenne des droits de l'Homme a sanctionné la France, en se fondant sur le fait que « *la conservation des empreintes digitales d'un ressortissant non condamné dans un fichier automatisé, constitue une atteinte disproportionnée au droit à la vie privée, cette mesure ne pouvant être considérée comme nécessaire dans une société démocratique* ».

Aujourd'hui, au prétexte de la simplification du déroulement de la procédure pénale, le législateur permet à la police et à la gendarmerie, dans le cadre d'une enquête, de mener toute action qui parait utile à la résolution de l'affaire. Cela autorise en pratique la consultation de tous les fichiers de l'ad-

12. Voir : « Rapport 2008 du groupe de contrôle des fichiers de police et de gendarmerie », http://www.lemonde.fr/mmpub/edt/doc/20081211/1129541_rapport_fichiers_v17.pdf .
13. Loi n° 2012-410 du 27 mars 2012 relative à la protection de l'identité.
14. Décision CEDH n° 19522/09 du 18 avril 2013, M. K. c/ France.

ministration, ainsi que la mise en œuvre des moyens publics pour traquer l'individu dans ses moindres faits et gestes, parfois en violation des droits et libertés individuelles.

Dans un rapport de 2008 commandé par le ministre de l'Intérieur, le Groupe de contrôle des fichiers de police et de la gendarmerie en charge d'examiner la mise en œuvre des dispositifs, après avoir constaté la nécessité de ces fichiers dans l'exercice des fonctions des autorités judiciaires, avait recommandé, entre autres :

- d'améliorer la procédure de création ou de développement des fichiers de police et de gendarmerie (qui reste très nébuleuse), notamment en institutionnalisant le groupe de contrôle sur les fichiers ;
- de fournir à la population une information pédagogique sur ces fichiers ;
- de définir les modalités de destruction, d'archivage et de transfert des fichiers ;
- de mieux contrôler l'utilisation des fichiers ;
- de désigner un expert « Informatique & Libertés » au sein des services de police et de gendarmerie.

Ces propositions de mesures, destinées à favoriser la transparence et à veiller au respect des droits et libertés individuelles, n'ont pas reçu l'assentiment de l'État, nonobstant tous les manquements relevés dans ce rapport. Les conditions de création et de mise à disposition des fichiers laissent penser que les gouvernements s'inscrivent désormais dans la logique de faciliter le recours dans les procédures pénales à l'ensemble des bases de données disponibles, dans la perspective de l'élucidation des affaires. C'est sans doute pourquoi les autorités de police et de gendarmerie étendent en pratique le fichier aux personnes impliquées dans des délits banals.

2. L'extension du fichier FNAEG aux personnes impliquées dans les délits banals

Le FNAEG a été créé par la loi du 17 juin 1998 relative à la répression des infractions sexuelles ainsi qu'à la protection des mineurs et mis en œuvre par un décret du 18 mai 2000. Sa finalité principale était d'identifier les récidivistes des infractions les plus graves à l'aide de leur profil génétique, ainsi que les personnes disparues et les cadavres non identifiés.

La collecte de l'empreinte génétique d'une personne requiert un prélèvement biologique sur celle-ci, ce qui constitue a priori une atteinte à son intégrité physique. À cela s'ajoutent les craintes pour le droit au respect de la vie privée que suscitent la conservation et l'exploitation de ces empreintes. C'est pourquoi, afin de garantir la proportionnalité du dispositif et de le rendre acceptable par la population, les possibilités d'enregistrement dans le FNAEG étaient à l'origine très limitées : seules les empreintes génétiques des personnes reconnues coupables d'une infraction à caractère sexuel ou de certaines atteintes aux mineurs pouvaient être conservées. Ces modalités ont été progressivement élargies par des réformes successives : d'abord aux principaux crimes d'atteinte aux personnes et aux biens par une loi de 2001[15], puis, par la loi n° 2003-239 du 18 mars 2003 pour la sécurité intérieure et enfin par la loi n° 2004-204 du 9 mars 2004 portant adaptation de la justice aux évolutions de la criminalité. Le législateur, par un décret de mars 2004[16], a modifié le Code de procédure pénale (CPP) pour y intégrer les dispositions de ces textes et également faciliter les modalités d'alimentation et de consultation du FNAEG.

En pratique, la police relève aujourd'hui systématiquement l'ADN de toute personne en garde à vue, nonobstant la présomption d'innocence qui est un des piliers du droit pénal. Sachant que le simple fait de conserver les empreintes génétiques d'une personne est susceptible de la stigmatiser ou de créer une discrimination à son encontre, se pose la question de savoir si ces pratiques ne constituent pas une atteinte disproportionnée à la protection des droits et libertés des personnes.

En 2015, 3 006 991 profils génétiques étaient enregistrés au sein du FNAEG. Ce nombre particulièrement important se répartit comme suit :

15. Loi n° 2001-1062 du 15 novembre 2001 relative à la sécurité quotidienne.

16. Décret n° 2004-470 du 25 mai 2004 modifiant le Code de procédure pénale (deuxième partie : décrets en Conseil d'État) et relatif au Fichier national automatisé des empreintes génétiques.

- 472 505 personnes condamnées, soit 16 % du nombre de personnes inscrites au fichier ;
- 2 280 448 personnes mises en cause au cours d'une enquête, soit 76 % des personnes concernées ;
- 254 038 traces de personnes inconnues, soit 8 % du total.

Ainsi, alors qu'il ressort des textes qui l'instituent que le FNAEG est un fichier d'identification qui n'a pas pour objectif de conserver les antécédents judiciaires, comment justifier l'enregistrement de l'ADN de 76 % des profils dans le fichier en dépit du fait que les auteurs n'ont pas été condamnés pour les infractions citées à l'article 706-55 du Code de procédure pénale ? Quelle que soit l'efficacité de cette pratique dans la résolution des crimes, l'enregistrement systématique des empreintes génétiques des auteurs des infractions mineures ou, dans certains cas, de simples mis en cause, en l'absence de tout encadrement, ne constitue-t-il pas un détournement de finalité ?

Partie II : De la légalité au détournement de finalité du FNAEG

Le principe de la légalité des délits et des peines, qui est une garantie fondamentale des droits de la personne devant les juridictions répressives, suppose la nécessité pour le législateur de prévoir des dispositions législatives ou réglementaires qui précisent le champ d'application du FNAEG. Jusqu'à présent, cette exigence semble faire défaut en France, d'autant plus que le Code de procédure pénale ne contient pas de prescriptions particulières qui déterminent notamment, le régime de général de la recherche en parentèle (I). Ce manquement résulte de la législation du FNAEG elle-même, qui reste ambiguë à plusieurs égards, comme semble l'indiquer la jurisprudence de la Cour européenne des droits de l'Homme (II).

1. La problématique de la recherche en parentèle

La recherche en parentèle est une technique permettant de déterminer le degré de parenté (ou son absence) entre deux échantillons. Comme pour la recherche directe, elle est basée sur la recherche de correspondances entre deux échantillons d'ADN.

En l'état actuel de la science, on considère que l'ADN est constitué de segments codants, qui servent à produire les protéines du vivant, et de segments non codants, qui ne semblent pas participer à cette production. Les segments codants présentent une faible variabilité génétique, car la plupart de leurs mutations conduisent à la production de protéines dysfonctionnelles, aux conséquences létales car inaptes à participer au développement de l'embryon ou de l'enfant qui les porte. À l'inverse, les segments non codants présentent une très forte variabilité car, du fait que leurs mutations ne semblent pas létales, peu de pression de sélection s'exerce sur eux. Il existe donc une bien plus forte probabilité que deux personnes différentes ne possèdent pas les mêmes segments non codants, et ce sont donc ces derniers qui sont utilisés à l'heure actuelle pour l'identification des personnes. Plus les deux échantillons comparés possèdent de segments non codants identiques et plus le degré de proximité parentale entre les deux est élevé, jusqu'à la correspondance parfaite lorsqu'il s'agit de la même personne ou de son vrai jumeau. Il est à noter qu'en dépit d'une pression toujours plus forte des services d'enquête, le législateur s'est pour le moment refusé à considérer l'étude des segments codants dans les enquêtes criminelles, car lesdits segments peuvent révéler des renseignements très intimes sur les personnes : couleur des yeux, des cheveux, de la peau, maladies génétiques portées, etc. [17] L'article 706-56-1-1 du Code de procédure pénale dispose que la liste des segments non codants devant être exclusivement considérés pour l'identification des personnes, y compris en parentèle, est fixée par arrêté, dont le dernier en date est celui du 6 juillet 2016 [18].

Alors qu'elle est utilisée depuis longtemps aux États-Unis, la recherche en parentèle est une pratique récente en France dans le cadre des enquêtes criminelles. Elle y fut employée pour la première fois

17. Cependant, dans une affaire récente en France, le magistrat enquêteur a demandé qu'un profil phénotypique d'un suspect soit établi à partir du profil génétique de l'échantillon collecté, ouvrant ainsi une brèche jurisprudentielle.

18. Voir : https://www.legifrance.gouv.fr/eli/arrete/2016/7/6/JUSD1618384A/jo/texte .

en 2011, dans le cadre de l'affaire « Élodie Kulik », une jeune femme violée et étranglée par plusieurs agresseurs le 10 janvier 2002, dans la Somme. Les enquêteurs furent mis sur la piste d'un premier auteur en 2011, une fois établie la correspondance entre l'ADN retrouvé sur la scène de crime et celui de son père, dont le profil génétique était enregistré dans le FNAEG pour avoir été condamné en 2001 pour des faits d'agressions sexuelles. Bien que le mis en cause soit mort dans un accident en 2003, son complice présumé, mis en cause suite à l'identification du principal protagoniste, est confondu. L'affaire est considérée comme élucidée le 16 janvier 2012, et ledit complice est renvoyé le 6 avril 2017 par les juges devant la cour d'assises. Les conséquences de l'absence de base légale sur la validité des preuves présentées ne sont à l'heure actuelle pas encore connues.

Pour continuer à mettre en œuvre cette technique redoutablement efficace, un protocole fut signé le 1er octobre 2012 entre la Direction des affaires criminelles et des grâces, le directeur général de la police nationale et le directeur général de la gendarmerie nationale, notamment pour entériner le régime de l'expertise en parentèle et plus généralement de l'analyse des empreintes génétiques.

Bien que ce protocole n'a fait l'objet d'aucune publication, la 1ère section de la Chambre de l'instruction de la cour d'appel de Paris l'a invoqué comme base légale dans son arrêt du 12 décembre 2016, pour rejeter une demande de nullité d'actes de la procédure dans une affaire de viols aggravés et d'agressions sexuelles aggravées dont les preuves ont été obtenues principalement par la recherche en parentèle en aveugle.

La particularité de cette décision est qu'elle reconnait la validité d'une recherche en parentèle ayant permis de cibler des collatéraux dont l'un sera finalement identifié comme porteur de l'ADN retrouvé à l'époque des faits en 2009. Pour autant, le Code de procédure pénale ne contient aucune disposition particulière permettant l'identification des collatéraux par leurs empreintes génétiques. L'article 706-56-1-1 du CPP, créé par l'article 80 de la loi du 3 juin 2016, qui reconnaît la recherche en parentèle, ne l'autorise expressément qu'« *aux fins de recherche de personnes pouvant être apparentées en ligne directe* [...] » [19], excluant ainsi les fratries et cousins [20]. Saisie en dernier ressort, la Chambre criminelle de la Cour de cassation, dans un arrêt du 28 juin 2017, après avoir constaté que la cour d'appel n'était pas fondée à appuyer sa décision sur le protocole susmentionné, a décidé que « *l'arrêt n'encourt pas pour autant la censure dès lors que les articles 81, 706-54 et suivants du Code de procédure pénale permettaient au juge d'instruction d'ordonner une expertise ayant pour objet l'identification et la recherche des auteurs des crimes et délits mentionnés par l'article 706-55 dudit code en sélectionnant, par une comparaison avec le profil génétique identifié comme étant celui de l'auteur de l'infraction* ».

Cette jurisprudence reste discutable en raison du fait que la Cour de cassation semble se substituer à la Chambre de l'instruction de la cour d'appel de Paris pour trouver une base légale à sa décision alors qu'en sa qualité de juge de droit, son rôle se limite à dire si les juges du fond ont fait une exacte application de la loi au regard des données de fait. La volonté de désigner un coupable après des années d'enquêtes criminelles n'aurait-elle pas pris le dessus sur les fondamentaux du droit pénal ?

De surcroît, le simple fait que la recherche en parentèle qui a été menée dans cette affaire, ait permis d'identifier un collatéral et non un descendant en ligne directe, comme le prévoit expressément l'article 706-56-1-1 du CPP précité, jette le discrédit sur cette jurisprudence, notamment quant au respect du principe de l'interprétation stricte de la loi pénale par le juge, prévue à l'article 111-4 du Code pénal. Le raisonnement de la Cour de cassation conduit en effet à faire de l'article 705-54 une base légale de la recherche en parentèle, alors que cette disposition semble viser les conditions pour lesquelles il est autorisé, dans le cadre d'une enquête judiciaire, de procéder à un rapprochement de l'empreinte de l'auteur présumé d'une infraction avec les données du fichier.

Cette jurisprudence confirme les propos de Catherine Bourgain, généticienne à l'Inserm, qui a déjà

19. Article 706-56-1-1 du Code de procédure pénale, créé par l'art. 80 de la loi n° 2016-731 du 3 juin 2016.

20. On ne peut s'empêcher de penser que cette limitation a été posée afin de restreindre ce dispositif aux recherches en paternité, lorsque les deux personnes sont connues. Son usage pour les recherches en aveugle est plus que discutable, voir *infra*.

dénoncé les écueils possibles du FNAEG en indiquant qu'« *on peut faire dire aux données ADN autre chose que ce pourquoi elles étaient prévues* » c'est-à-dire « *que si l'un de vos proches est fiché, vous l'êtes aussi en partie* » [21]. Cette situation soulève également des inquiétudes sur les atteintes à la vie privée de personnes qui ne sont pas directement concernées, dont le nombre est très important. En effet, la présence de l'ADN d'une personne renseigne de façon fiable sur ses deux parents biologiques, ainsi que sur ses enfants. En supposant qu'en France le nombre moyen d'enfants par femme est d'environ 2,1 [22], qu'il y a à peu près autant d'hommes que de femmes dans la population, et que les deux parents et les enfants de chaque personne sont effectivement connus [23], on peut considérer que chaque personne possède environ 2,1 descendants directs en moyenne et donc que le coefficient multiplicateur entre le nombre de personnes présentes et celles qu'il est possible d'impliquer directement ou indirectement est de 5,1 [24]. Sur la base des chiffres de 2015, le FNAEG permet ainsi d'identifier directement et indirectement plus de 14 millions de personnes. En incluant la recherche en parentèle indirecte sur la fratrie, une personne ayant en moyenne (2,1 - 1) = 1,1 sœurs ou frères, ce coefficient peut s'élever jusqu'à 6,2 [25]. Les réserves du Conseil constitutionnel, dans sa décision du 16 septembre 2010, doivent être interprétées comme une mise en garde à l'encontre du législateur pour l'obliger à compléter et à préciser l'encadrement légal du FNAEG afin d'éviter toute interprétation extensive, surtout lorsque des situations non prévues sont soumises aux juges.

2. Une législation ambiguë

Le Conseil constitutionnel a été déjà saisi par la Cour de cassation d'une question prioritaire de constitutionnalité (QPC) pour examiner la conformité à la Constitution des dispositions ayant institué le FNAEG. Dans sa décision, la Haute juridiction a jugé conformes les articles 706-54, 706-55 et 706-56 du Code de procédure pénale, dans leur rédaction antérieure à la loi n° 2010-242 du 10 mars 2010 tendant à amoindrir le risque de récidive criminelle et portant diverses dispositions de procédure [26], sans manquer de soulever deux réserves d'interprétation fondées sur l'article 9 de la Déclaration de 1789 qui, en matière de procédure pénale, proscrit « *toute rigueur qui ne serait pas nécessaire* ».

La première de ces réserves concerne les infractions permettant un prélèvement d'ADN aux fins de rapprochement avec les données du fichier, énoncées dans le 3ème alinéa de l'article 706-54 du CPP. Le Conseil a spécifié que l'expression « crime ou délit » employée par le législateur dans ledit 3ème alinéa devait s'interpréter comme limitant ce prélèvement à l'égard des personnes soupçonnées d'avoir commis les infractions énumérées à l'article 706-55 du CPP. Par conséquent, la commission d'une simple contravention ou d'un délit non visé par cet article ne peut donc pas conduire à un tel prélèvement (cons. 19).

21. Simon PIEL, « Comment l'enquête sur le meurtre d'Elodie Kulik a été relancée par l'ADN d'un parent », *Le Monde*, 21 février 2012, http://www.lemonde.fr/societe/article/2012/02/21/comment-l-enquete-sur-le-meurtre-d-elodie-kulik-a-ete-relancee-par-l-adn_1642851_3224.html .

22. Source INSEE, voir : https://www.insee.fr/fr/statistiques/1379743 .

23. Le pourcentage d'enfants nés d'adultère n'est pas négligeable, sans pour autant qu'il soit possible de disposer de chiffres fiables en la matière. Le fichage génétique systématique promu par certains pourrait incidemment permettre au responsable du fichier d'identifier tous les enfants nés de dons de gamètes ou d'adultère, ainsi que les géniteurs dont les empreintes seraient contenues dans le fichier.

24. Cette estimation fait abstraction de nombreux autres biais, tels le fait que la natalité des catégories socio-professionnelles inférieures, les plus représentées parmi les auteurs d'infractions actuellement susceptibles d'inscription au FNAEG, est supérieure à la moyenne, ou encore que si deux enfants sont fichés, ni eux ni leurs parents ne doivent être comptés deux fois. Pour autant, son ordre de grandeur doit être considéré comme pertinent pour le nombre d'inscrits considéré.

25. Cette dernière valeur n'est que théorique, car la fiabilité d'identification des fratries est bien inférieure à celle en ligne directe. En effet, selon la manière dont les segments non codants des parents sont répartis dans leurs gamètes, deux enfants peuvent avoir en commun moins de segments non codants que deux personnes non apparentées. De façon générale, le risque d'erreur d'interprétation des résultats de correspondances génétiques doit inciter les enquêteurs et magistrats à la plus grande prudence quant à l'usage de ces résultats dans leurs procédures.

26. Décision n° 2010-25 QPC du 16 septembre 2010.

La seconde de ces réserves porte sur la fixation de la durée de conservation des empreintes au fichier. Le Conseil constitutionnel a jugé qu'il appartient au pouvoir réglementaire de préciser par décret cette durée, de la « *proportionner compte tenu de l'objet du fichier, à la nature ou à la gravité des infractions concernées tout en adaptant ces modalités aux spécificités de la délinquance des mineurs* » (cons. 18).

Cette décision n'est pas sans incidence sur la croissance du nombre de profils génétiques au FNAEG, l'élargissement des crimes et délits concernés en étant la principale cause. Conçu spécifiquement pour ficher les criminels sexuels, le fichier ratisse large aujourd'hui puisqu'il va des auteurs des crimes contre l'humanité aux vols simples et aux arracheurs d'OGM.

Le prélèvement d'empreintes génétiques suppose en principe le consentement exprès de son auteur, mais l'article 706-56 du CPP semble se passer de ce consentement puisqu'il fait du refus de s'y soumettre un délit continu qui expose son auteur à un an d'emprisonnement et 15 000 euros d'amende. Cette peine est doublée lorsque les faits sont commis par une personne condamnée pour crime. À cela s'ajoutent les prélèvements par surprise que peuvent faire les officiers de police. La pratique consiste généralement à récupérer, à l'occasion d'un interrogatoire, la salive déposée sur un gobelet, ou l'un des cheveux de la personne, parfois à son insu. Un magistrat du TGI de Paris indique ainsi que « *rien n'oblige les officiers de police judiciaire à prévenir du fichage ADN lors d'un prélèvement clandestin* » [27].

La CEDH est récemment intervenue pour sanctionner la France pour fichage abusif. Dans sa décision du 22 juin 2017, la Cour a rappelé que le simple fait de mémoriser des données relatives à la vie privée d'un individu constitue une ingérence au sens de l'article 8. La CEDH a estimé que la condamnation d'une personne pour son refus de prélèvements d'empreintes génétiques, était contraire au droit au respect de la vie privée. « La Cour considère que le régime actuel de conservation des profils ADN dans le FNAEG n'offre donc pas, en raison tant de sa durée que de l'absence de possibilité d'effacement, une protection suffisante à l'intéressé et ne traduit donc pas un juste équilibre entre les intérêts concurrents, publics et privés, en jeu. [28] » Cette décision semble aujourd'hui inspirer la jurisprudence française, notamment quant à l'appréciation du « délit continu » de refus de prélèvement d'ADN. Ainsi, par un jugement du 3 novembre 2017, le tribunal de Paris a prononcé la relaxe d'un militant anti-publicité poursuivi pour ce délit, au motif que les faits qui lui sont reprochés ne justifient pas une inscription au FNAEG.

À la lumière des éléments ci-dessus, la proportionnalité du FNAEG, sous sa forme et son usage actuels, peut donc être questionnée à plusieurs titres.

En premier lieu, si la recherche en parentèle directe peut faire sens dans les recherches en paternité, pour laquelle les deux individus sont connus, la recherche en parentèle à l'aveugle dans le FNAEG, directe ou indirecte, constitue un détournement manifeste de finalité. En effet, le FNAEG n'est dans ce cas utilisé que comme un fichier « de circonstance », de « gens honnêtes », puisqu'en dépit de leurs infractions passées les personnes fichées n'ont en l'espèce pas commis l'acte faisant l'objet de la recherche. La recherche en parentèle ne présenterait aucun intérêt pratique si le FNAEG ne contenait que quelques milliers de criminels ; son efficacité est donc conditionnée par l'existence préalable d'un fichage de masse. Qui plus est, les infractions susceptibles de conduire au fichage sont celles généralement commises par des individus appartenant aux groupes sociaux les moins favorisés, alors que la criminalité en col blanc est exclue de ce type de fichage. La recherche en parentèle en aveugle induit donc une inégalité devant la loi au détriment des « gens honnêtes » les moins aisés.

En deuxième lieu, le fait que la recherche en parentèle en aveugle ne permette pas l'identification d'une unique personne, mais renvoie l'information que l'un des descendants d'une personne présente puisse être incriminé, semble manifestement contrevenir aux dispositions de l'article 12 de la directive 2016/680 [29], selon lequel : « *le traitement des données génétiques, des données biométriques* » est restreint « *aux fins d'identifier une personne physique de manière unique* », ce qui n'est pas le cas.

27. Évelyne SIRE-MARIN, « La société du soupçon - Pucés, filmés, fichés, quelles alternatives? », in *Contre l'arbitraire du pouvoir - 12 propositions*, La Fabrique, Paris, février 2012, ISBN : 978-2-358-72026-7, pp. 97-120, http://lafabrique.fr/contre-larbitraire-du-pouvoir/ .

28. Communiqué de presse CEDH n° 215 (2017).

29. Directive (UE) 2016/680 du 2 avril 2016 du Parlement européen et du Conseil, relative à la protection des personnes

En troisième lieu, le fait que le FNAEG contienne actuellement les données de près de 5 % de la population, couplé à la possibilité de recherche en parentèle indirecte, étend le pouvoir d'atteinte de ce fichier à une fraction considérable de la population française. On se trouve donc en présence d'un « méga-fichier » par destination, ciblant spécifiquement les « gens honnêtes ». Ces modalités d'usage devraient donc en toute logique être censurées par le Conseil constitutionnel, selon le même raisonnement que celui employé en 2012. Indépendamment de la problématique de la recherche en parentèle en aveugle, la proportionnalité du FNAEG peut être questionnée sur la seule base du nombre important de personnes qu'il contient.

Il résulte de ces considérations que la législation du FNAEG révèle des incohérences et mérite à cet effet d'être révisée pour assurer la protection des libertés individuelles telles qu'elles sont garanties par les textes fondamentaux. Une application stricte du principe de proportionnalité devrait également conduire à une purge significative de son contenu.

Conclusion

La réglementation de l'usage des technologies numériques reste un défi pour le droit. Là où les « méga-fichiers », et en particulier le FNAEG, sont perçus comme une menace pour le droit à la vie privée, l'État les considère comme des outils indispensables aux enquêtes criminelles, permettant de facto d'assurer pleinement sa mission de protection des populations. Face à la médiatisation de la criminalité et au spectre du terrorisme islamiste, les pouvoirs surenchérissent en créant des banques de données biométriques de plus en plus importantes, et en mettant en place, dans le cadre du traité de Prüm [30], un échange facilité de ces données.

Pour autant, ce n'est pas parce qu'une technologie est rendue possible par le progrès scientifique et technique qu'elle doit être utilisée. Face au risque de leur dévoiement, la main du législateur ne doit pas trembler. C'est ainsi qu'en 2011, le Royaume-Uni, pourtant grand adepte des technologies de surveillance de masse, détruisit son National Identity Register. De même, les Pays-Bas interdirent la création de tels méga-fichiers, que la centralisation rend qui plus est extrêmement vulnérables. La France, sous le double regard des exemples voisins et de la législation européenne, saura-t-elle retrouver le sens de la mesure ?

physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil.

30. Le traité de Prüm a été signé le 27 mai 2005 à Prüm, en Allemagne, par sept États membres de l'Union européenne : l'Allemagne, l'Autriche, la Belgique, l'Espagne, la France, le Luxembourg et les Pays-Bas. Il prévoit notamment l'échange de données génétiques, d'empreintes digitales et autres données à caractère personnel.

Systèmes d'information pour les chercheurs en droit

Alex Chauvet, Annie Foret

alex.chauvet@u-bordeaux.fr, univ. Bordeaux
foret@irisa.fr, univ. Rennes 1 & IRISA

Convergences du Droit et du Numérique, 2017

Résumé

Cette étude concerne les systèmes d'information en droit et propose un nouveau prototype. Nous discutons les systèmes d'information actuellement utilisés par les chercheurs en droit, en précisant leurs limites. Nous présentons les principes d'un nouveau prototype pour un meilleur système. Ce travail s'accompagne d'une première réalisation concrète, un système à facettes sémantiques, résultat de notre chaîne de traitement sur un ensemble de décisions du Conseil Constitutionnel.

Contenu :

1 Introduction

La recherche juridique est aujourd'hui largement fondée sur un positivisme : il s'agit d'étudier le droit tel qu'il est. Il s'ensuit que les études menées sont non seulement assises sur les textes mais surtout sur leur application. A ce titre, les arrêts et décisions de justice occupent une place prédominante dans les sources des travaux.

Cependant, ce prisme contentieux n'est pas nouveau. Ainsi, l'importance de la compilation des décisions de justice pour en favoriser l'étude était déjà à l'origine de la création de recueils de jurisprudence dont certains ont allègrement dépassé le siècle d'existence (recueil Lebon, Bulletin de jurisprudence de la Cour de cassation).

Pour autant, la problématique est aujourd'hui renouvelée par au moins deux facteurs : l'explosion du nombre de décisions depuis une trentaine d'années[1] ; l'ouverture et l'accès croissant aux décisions de justice. Il y a donc une aubaine pour la recherche mais aussi une source de difficultés qu'elle n'a pas encore su appréhender de manière satisfaisante. S'il existe des systèmes d'informations compilant la jurisprudence, ceux utilisés par les chercheurs en droit sont largement perfectibles et demeurent limités dans leurs possibilités (voir Figure 4). En particulier, ils n'exploitent pas suffisamment les caractéristiques de ces textes juridiques qui se prêtent pourtant assez bien à un traitement informatique, surtout dans un pays où la motivation des décisions de justice est caractérisée par une rédaction sèche et reprenant la forme d'un syllogisme. En effet, plus que des énoncés littéraires, ils sont des propositions articulant des informations (énoncés porteurs de structures, redondance des termes et formules stéréotypées, marqueurs du discours, articulation logique).

Pourtant, la recherche juridique n'a pas investi ce champ consistant à produire elle-même ses bases de données et systèmes d'information. Elle ne produit complètement ni ses données ni les outils permettant de les exploiter. Sur ce point, elle s'oppose par exemple à l'économie, qui est pourtant une science humaine voisine et qui a longtemps été enseignée avec le droit. Ainsi, les chercheurs en droit subissent les systèmes d'information plus qu'ils ne participent à leur conception, alors même que ceux-ci ne sont pas toujours adaptés. Les raisons sont multiples. La première réside sans doute dans un déficit de compétence et de formation aux techniques afférentes. Mais plus fondamentalement, la doctrine juridique n'est pas assez sensibilisée sur les potentialités de tels outils. En particulier, elle ne perçoit pas forcément que ceux-ci ne sont pas seulement un moyen d'archiver des arrêts et décisions, mais aussi un moyen de les présenter selon une perspective. Ces systèmes peuvent éventuellement permettre une superposition des perspectives doctrinales sur un même objet. Par exemple, une décision du Conseil constitutionnel sur une loi relative au Code du travail intéressera les spécialistes du droit constitutionnel mais aussi les spécialistes du droit du travail. Leurs visions seront probablement différentes mais complémentaires et aboutiront à des classifications différentes. En l'état, il peut être assez difficiles pour l'un des groupes d'accéder rapidement à la perspective de l'autre. Or les juristes affectionnant tout particulièrement les classifications et le droit étant de moins en moins compartimenté et plus en plus transversal, il y aurait là matière à gagner en efficacité des recherches.

Par ailleurs, d'autres arguments d'ordre méthodologique devraient aussi attirer l'attention des juristes. Ils ressortissent de biais épistémologiques trop peu souvent examinés.

Un premier biais épistémologique a trait à l'autonomie et à l'indépendance de la recherche. La plupart des systèmes d'information aujourd'hui utilisés par la doctrine sont construits par les juridictions elles-mêmes. Les chercheurs sont donc dans une relative dépendance vis à vis des organes qu'ils sont chargés d'étudier voire parfois de critiquer... Pendant longtemps, cette dépendance était nécessaire dans la mesure où la diffusion des décisions de justice était limitée voire verrouillée. Les juridictions ne donnaient vraiment à voir que ce qu'elles voulaient. Aujourd'hui, elle n'a plus de raison d'être. Ainsi, continuer d'utiliser uniquement ces systèmes d'information, c'est refuser de prendre de la distance avec ce qui constitue régulièrement un objet de recherche. La chose est connue des chercheurs en droit pour ce qui est de

[1] http://www.justice.gouv.fr/statistiques-10054/chiffres-cles-de-la-justice-10303/

l'autonomie conceptuelle. Car c'est là une spécificité du droit que d'avoir un objet qui n'est pas muet. Le soleil ne se qualifie pas lui même d'étoile ; la terre ne se désigne pas comme une planète. Or les objets juridiques prétendent très fréquemment appartenir à des catégories. Par exemple, lorsqu'une juridiction refuse de qualifier un mécanisme de peine, elle ne le fait pas forcément selon des critères objectifs et scientifiques. Elle peut le faire pour ne pas attacher à ce mécanisme le régime juridique qui va avec une peine (sa proportionnalité, son absence de caractère rétroactif notamment). Cela ne signifie pas que le chercheur ne peut pas qualifier lui ce mécanisme de peine. Aussi contre-intuitive qu'elle puisse paraître, cette autonomie des concepts est nécessaire car doctrine et juridictions (plus généralement autorités normatives) sont dans des situations radicalement différentes : la juridiction applique le droit et tout son travail de classement est guidé par cet impératif ; le chercheur vise à rendre le chaos du monde un peu plus intelligible et cohérent. Ainsi, de la même manière qu'il n'est pas scientifique de ne pas s'affranchir d'une notion produite par un juge lorsque celle-ci ne convient pas, il n'est pas scientifique non plus de continuer de se reposer sur les systèmes d'information fournis par les juridictions sans envisager de s'en affranchir. Il y a là un biais de la recherche. En effet, ces derniers ne présentent pas des données brutes mais bien des informations, avec une perspective. Celle ci est très remarquable lorsque sont reprises les classifications des anciens recueils papiers ou lorsqu'il n'est possible de filtrer les décisions que selon des propriétés limitées et choisies par le service de documentation de la juridiction.

Un autre biais épistémologique vient d'une tendance assez naturelle des études doctrinales à se focaliser sur des corpus limités de décisions. Le plus souvent, il s'agit d'étudier la jurisprudence des juridictions les plus importantes du pays (Conseil d'Etat, Cour de cassation, Conseil constitutionnel) et par forcément en intégralité. La raison est assez simple : ces juridictions contribuent plus activement à la production du droit que les juridictions inférieures dont la mission est plutôt de juger des affaires "du quotidien". A cela s'ajoute que les décisions de ces juridictions sont plus facilement accessibles car leur diffusion est mieux assurée et plus systématique. Mais procéder ainsi consiste à ne s'intéresser qu'à un spectre très restreint du droit en postulant éventuellement qu'il est identique au droit appliqué par les autres juridictions. Pour les chercheurs du début du XXème siècle, ce mode opératoire était assez justifié car il y avait alors beaucoup moins et de juridictions et de décisions. Mais aujourd'hui, compte tenu des délais, de la complexité croissante des procédures et de leur coût, nombreuses sont les affaires qui ne parviennent pas à ces juridictions. Or la meilleure connaissance des arrêts et décisions de principe des juridictions les plus prestigieuses du pays ne permet pas forcément de savoir comment une affaire sera tranchée en première instance... L'intérêt d'utiliser des systèmes d'information est alors de dépasser le champ d'étude actuel des décisions contentieuses et d'élargir le spectre : d'abord en intégrant plus largement les décisions des juridictions suprêmes, puis en y intégrant les décisions des juges "ordinaires". Ce faisant, on aurait des indications sur les masses et ordres de grandeur et une modélisation sans doute plus fidèle du droit tel qu'il est appliqué concrètement.

Dernière enjeu épistémologique : la pérennité et l'accessibilité des recherches entreprises. Dans toute étude contentieuse, la démarche du chercheur commence nécessairement par un recensement et une classification des décisions à étudier selon des critères propres à sa perspective. Actuellement, ce travail est souvent perdu à l'issue de la recherche ou ne fait l'objet que d'un tableau plus ou moins précis dans l'appareil critique. Le conserver à l'aide d'un système d'information permettrait qu'il soit directement utilisé par d'autres qui ne perdraient donc pas de temps ou beaucoup moins à retrouver le corpus de décisions pertinent. Accessoirement, elle faciliterait aussi la revue par les pairs et autoriserait le développement de méthodes quantitatives et statistiques sur les décisions de justice.

Si les perspectives sont nombreuses pour le juriste, le développement de ces outils n'en demeure pas moins une tâche difficile qui requiert donc une aide technique. A cet égard, pour les chercheurs en informatique concernés par la représentation des connaissances et le traitement automatique des langues naturelles, travailler sur un objet comme le droit permet de tester la pertinence de certains modèles et de démarches associées. Nous considérons ici l'approche des systèmes d'information logique avec l'objectif de redonner du pouvoir à l'utilisateur. Dans le domaine du droit, nous envisageons des systèmes

d'aide à un usager et une chaîne de traitement allant des textes bruts ou partiellement structurés, à des représentations formelles et à facettes logico-sémantiques, pour traiter et valoriser des données.

Au-delà de la validation d'une approche, ce type d'étude peut aussi mener à raffiner et à adapter les modèles initiaux et ceci à différents niveaux d'analyse (logico-formel, linguistique, traitement des données, interactions).

L'étude suivante se veut donc une démonstration par l'exemple. Son objet est volontairement limité et consistera à présenter des décisions du Conseil constitutionnel pour un juriste se donnant pour objet de recherche les élections législatives.

2 Illustration avec un système compilant les décisions du Conseil constitutionnel relatives aux élections législatives.

2.1 Choix effectués

Il serait illusoire de vouloir immédiatement produire un système d'information abordant un très grand nombre de décisions de justice. En effet, en dépit d'une certaine proximité dans leurs structures et formulations, celles-ci présentent une assez grande variabilité qui rend leur traitement moins facile qu'il peut paraître de prime abord. Dans un premier temps, il est donc nécessaire de faire un choix répondant à des impératifs multiples.

D'abord, puisque l'objet principal est juridique, le vocabulaire peut être un obstacle. En effet, selon le domaine concerné, les termes et expressions employés peuvent être plus ou moins abscons et constituer des facteurs de ralentissement. Le choix de la thématique est donc important. Par ailleurs, celle-ci doit autoriser une amélioration des outils existants ou au moins des ajouts propres au chercheur. Ce point n'est pas anodin car si le corpus de décisions étudié est trop hétérogène ou inversement trop homogène, il ne permet aucune systématisation et donc aucune présentation rationnelle. Là encore, il y a donc un choix à opérer, notamment de la part du juriste. Il consiste à déterminer ce qu'il faut archiver et comment. Enfin, vient le problème majeur, celui de la difficulté technique, certains critères pouvant être extraits ou récupérés de sources ouvertes, d'autres devant être complètement construits.

Dans le cadre de ce projet, le choix de la thématique s'est ainsi porté sur les élections législatives. La juridiction concernée était donc principalement le Conseil constitutionnel.

Bien que sa conception soit conforme à une utilisation standard, le site de cette juridiction peut présenter des limites pour un chercheur (voir annexe). Les données qu'il contient peuvent faire l'objet d'ajouts et d'une présentation différente de celle dont l'utilisateur est captif. En effet, sur la question des élections législatives, il apparaît rapidement que les critères formels retenus sont insuffisants. Ces élections peuvent être évoquées dans des décisions DC, QPC (pour leurs règles), AN (pour le déroulement des élections), I et D (pour les fins avant terme des mandats). De même, pour le juriste, cette classification ne permet par forcément d'accéder à l'information recherchée. Par exemple, pour les décisions I et D relatives aux incompatibilités et déchéances de mandat, il importe surtout de savoir quelle est la cause de cette fin de mandat avant terme. Souvent, il s'agit de l'exercice d'une profession ou d'une fonction incompatible. Idem pour les décisions AN relatives au contrôle de l'élection : ce contrôle peut porter sur le déroulement du scrutin, sur les comptes de campagne ; il peut aboutir à une validation ou une invalidation etc. Dans toutes ces situations, il n'existe pas de propriété qui permettrait d'ordonner les décisions étudiées et la limitation de la recherche textuelle devient alors handicapante.

D'un point de vue technique, le choix de cette thématique et de cette juridiction était aussi assez opportun. En effet, les sites des juridictions et légifrance, qui sont les sources principales de décisions, sont assez inégaux dans la façon dont les données sont structurées. La plupart du temps, les décisions sont archivées comme du texte brut avec un balisage html qui vise principalement à leur mise en forme sur la page internet. En revanche, le site du Conseil constitutionnel va un peu plus loin en offrant des décisions légèrement structurées permettant une extraction un peu plus simple de certaines informations.

2.2 Scenarios d'utilisation.

La conception du système d'information impose de s'entendre sur un cahier des charges. Il s'agit ici de définir ce que l'on veut pouvoir faire et que le site du Conseil constitutionnel ne permet pas.

Scenario 1 : amélioration de l'ergonomie.

> Le système d'information proposé doit d'abord permettre une navigation équivalente à ce qu'autorise la recherche experte sur le site du Conseil constitutionnel. Il en reprend donc les critères (type, date, résultat etc). Dans le cadre d'une telle utilisation, son intérêt réside principalement dans le caractère intuitif des questions posées et dans le regroupement de certaines propriétés que le site du Conseil n'exploite pas complètement. Par exemple, les dates des décisions peuvent être regroupées par années (comme le site du Conseil), mais aussi par mois voire jours.

Scenario 2 : gain en précision des recherches.

> Le système d'information proposé doit ensuite permettre une navigation plus fine sur les décisions, notamment en ne traitant plus une décision comme une unité d'information insécable. Il s'agit ici de descendre au niveau du "considérant" (paragraphe). Le système permettant l'étiquetage des éléments (l'ajout d'une nouvelle propriété), le chercheur peut ainsi identifier les "considérants" importants des décisions et les regrouper.

Scenario 3 : choix du point de départ de la recherche ; ajout de perspectives propres au chercheur.

> Le système d'information proposé doit enfin permettre, en fonction des critères intégrés, de commencer la recherche depuis un point qui ne serait pas autorisé par le site du Conseil constitutionnel. Il doit permettre au chercheur d'organiser la recherche en fonction de la perspective qu'il s'est donnée. En s'appuyant sur des catégories établies par le juriste, il doit ainsi faciliter l'isolation d'un corpus de décisions présentant la même propriété.
>
> Si, par exemple, on s'intéresse à l'organisation et aux principes des élections législatives, il faut pouvoir accéder aux décisions présentant la propriété identifiant cette thématique, *indépendamment de la structure du site du Conseil constitutionnel* (notamment en décisions DC, QPC, I, AN etc.). En l'occurrence, une telle requête consiste à sélection la question sur `SousDomaine ?` puis Organisation et principes des élections législatives.
>
> Idem, par exemple, pour les décisions I (incompatibilité) qui sont souvent présentées selon un critère personnel : incompatibilité de M. ou Mme X. Or s'il travaille sur les incompatibilités en général, le chercheur préférera une information sur la profession en cause plutôt que sur la personne en cause. Ici, une telle requête consistera à naviguer dans les décisions I en fonction des questions `ProfessionCompatible ?` et `ProfessionVisee ?`

Afin de répondre à ce cahier des charges, le prototype requiert le choix de critères qui dépendent de la nature des données et des relations des éléments à classer (contenant/contenu, un à plusieurs, plusieurs à plusieurs). Ils sont détaillés dans la section 3.4 sur le choix des descripteurs.

3 Chaîne de traitement

3.1 Vue d'ensemble

Dans cette expérience, nous ciblons la réalisation d'un prototype, permettant une recherche flexible, portant sur un ensemble de *décisions de justice*.

Nous suivons pour cela la chaîne de traitement illustrée dans la Figure 1 pour une mise au point à partir d'une sélection d'exemples, appliquée ensuite à un ensemble plus complet de documents.

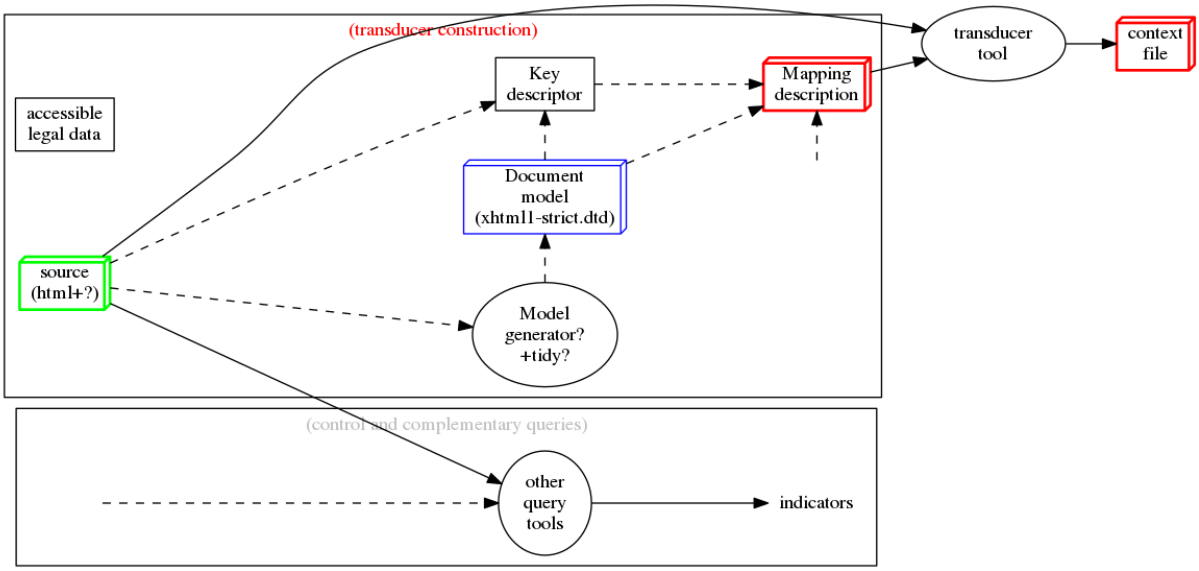


Figure 1: schéma de développement

La chaine vise à obtenir un "fichier de contexte logique" (cadre à droite de la Figure 1), qui pourra être chargé dans un "système de gestion de contexte logique". *Ceci est analogue au chargement d'une base de données dans un système de gestion de base de données, mais un contexte logique offre d'autres possibilités.* Nous avons utilisé Camelis (version 1) http://www.irisa.fr/LIS/ferre/camelis/ basé sur l'analyse de concept logique (S. Ferré and O. Ridoux), une approche qui étend l'analyse de concept formel (B. Ganter and R. Wille).

Cette réalisation, s'est appuyée sur des données légales accessibles au format XML / HTML, avec un modèle de document associé. Ce format permet alors un certain type de traitement uniforme. La construction est guidée à la fois par un modèle d'entrée et un modèle de contexte logique visé avec : un choix de types d'objets ; une sélection de propriétés ; un lien entre les deux modèles et la spécification d'une clé dans le source.

3.2 Systèmes d'information logique

3.2.1 Définitions

Formellement, un *contexte logique* est défini par un ensemble fini O d'objets o_i (de label l_i) et pour chaque o_i, un ensemble fini de descriptions logiques $d(o_i)$ pour un langage logique L.

Un *système de gestion de contexte logique* permet de charger et d'exploiter un tel contexte, d'interroger par des requêtes logiques (explicites dans L, ou interactives) ; la réponse est alors un sous-contexte d'objets satisfaisant cette requête.

3.2.2 Utilisation pour l'exploration de contextes.

Une vue en contexte logique permet d'explorer des informations de manière flexible, avec des facettes sémantiques, des possibilités d'inférences logiques, sans rédaction de requête a priori et d'obtenir aussi des indications sur la qualité des données. Un tel contexte peut être enrichi par d'autres informations

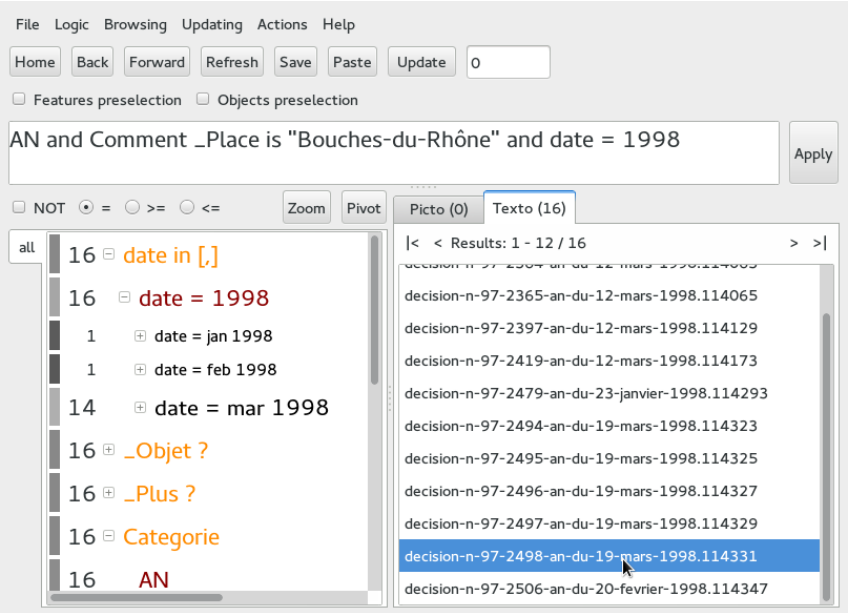


Figure 2: Sous-contexte sélectionné dans Camelis (3 fenêtres synchronisées)

(de natures diverses), en le reliant à d'autres applications (par des actions associées selon des arguments dans le contexte).

Présentation de contexte. Un contexte est présenté avec trois fenêtres synchronisées (voir Figure 2) : une question (une requête éditable) dans la fenêtre en haut, un index multi-hiérarchique de propriétés (les liens cliquables) dans la fenêtre à gauche, une présentation des objets (des exemples sélectionnables) dans la fenêtre à droite.

Navigation dans des sous-contextes. L'utilisation typique du système est une navigation dans un contexte menant du contexte initial *All* (comprenant tous les objets prévus dans le contexte) à un sous-contexte C_i ($i > 0$) comprenant les objets satisfaisant un ensemble de propriétés particulières (cf video[2]). Les propriétés du sous-contexte C_i sont résumées par la formule en haut (requête éditable), cette requête n'a pas besoin d'être rédigée, elle peut être construite par clics et sélections dans l'index de propriétés dans la fenêtre à gauche (les objets présentés seront alors ceux vérifiant la propriété cliquée) ou par sélection dans la fenêtre des objets.

Langage de requêtes. La connaissance du langage de requête n'est pas nécessaire. Cependant elle peut être lue ou éditée pour un usage en mode expert.

Le langage permet une combinaison avec les opérations booléennes : and, or, not de propriétés selon le type d'attribut ; il permet d'intégrer des questions de la forme :

- sur les chaînes de caractères : ... is "une chaîne" ;

  et aussi (où l'initiale du mot-clé suffit, c pour contains etc.) :

[2] http://www.irisa.fr/prive/foret/DEMO/VIDEOS/Screencast7sep2017-LIS-DecisionsConseilConsitutionnel.webm

... contains "une chaîne" ; ... beginswith "une chaîne" ; ... endswith "une chaîne" ;

... match "expression régulière (de type unix)" ;

exemple (nombre avec signe optionnel) : ... match "[+-] ?[0-9]+"

- sur les entiers : ... = un_entier ; ... in un_intervalle

  exemple : in 2000 .. 2015 ; ex : in 2000 ..

- sur les dates : ... = une_date ... in un_intervalle

  exemple : date in [may 2016,jul 2017]

Le langage de requête permet aussi des questions sur les heures, les droits sur fichier.

3.2.3 Nouveau prototype

Notre réalisation concrète sur des données du Conseil Constitutionnel, illustre les avantages d'une telle approche pour des informations juridiques. D'autres formats cibles sont aussi possibles. Le contexte que nous avons construit présente les caractéristiques suivantes.

- Les objets sont des décisions du conseil constitutionnel sur les élections.
- Les facettes sont des propriétés choisies, avec extraction de valeurs automatisée ; un ajout de tags personnels est possible.
- Il s'agit d'un contexte modulaire, pouvant comporter des données hétérogènes.
- Le système permet une navigation/sélection flexible et sûre, même sans connaissance du langage.

Nous détaillons par la suite certaines étapes de la construction.

3.3 Descripteurs de clés pour l'accès et la représentation

3.3.1 Codes de document existants.

Au niveau document, plusieurs identifiants peuvent être utilisés.

Numéro de document. Pour identifier un document lors des traitements, nous privilégions le numéro de décision. Cet identifiant est un texte avec un ou plusieurs numéros, un indicateur et une date, comme dans cet exemple : `Décision n° 2017-651 QPC du 31 mai 2017`

La même information apparaît de façon légèrement différente, dans le lien html :

`decision-n-2017-651-qpc-du-31-mai-2017.149036.html`

Code ECLI. Toutefois, les juridictions françaises utilisent depuis quelques temps un numéro visant à servir d'index au niveau européen dans le but de la constitution de bases de données. Ce numéro s'appelle ECLI (European Case Law Identifier[3]). Dans les décisions du Conseil constitutionnel, on le trouve souvent à la fin, en pied de page. Chez 99% des juristes, ce numéro n'est pas naturellement recherché mais il existe et finira sans doute par s'imposer ou à coexister de manière pérenne. La Figure 3 montre comment le numéro ECLI est généré pour le Conseil constitutionnel.

La double représentation ECLI et numéro de décision est en fait possible dans le prototype et nous proposons les deux comme facettes de recherche dans l'index.

[3] https://e-justice.europa.eu/content_european_case_law_identifier_ecli-175-fr.do

Identifiant France + Conseil constitutionnel	Année du jour de la décision	Numéro de la décision	Type de la décision
ECLI :FR :CC :	2017 :	2017.640.	QPC

Figure 3: Composants du Code ECLI

3.3.2 Codes choisis

Nous avons choisi le numéro de décision comme descripteur principal pour la représentation.

Au niveau source. Au niveau des données, l'unité de base est le document, dont nous considérons comme identifiant principal, le "numéro de document" (un texte avec un ou plusieurs numéros, un indicateur et une date, comme décrit ci-dessus).

Au niveau prototype. Au niveau du contexte produit (système d'information), l'unité de base est un objet désigné par cette information. Nous lui associons des propriétés le décrivant et qui semblent utiles à la recherche parmi ces objets et ces propriétés.

La clé privilégiée est ainsi la même dans le source et dans sa transformation. Les traitements s'appuient sur cet identifiant.

3.3.3 Unité d'information et balisage XML/html.

La décision entière en tant que document est l'unité d'information minimale en l'état. Mais conserver la décision entière comme seule unité d'information, impose des résultats mixtes et certaines limitations.

Il serait possible de gagner en précision en exploitant un balisage formel du document html, certes rudimentaire mais déjà plus fin, qui distingue les grands blocs, dont les fameux motifs.

Pour prendre en compte ces différents niveaux dans le système d'information, il serait possible d'ajouter des objets repérant des composants de document. il faudrait alors affiner la clé relative à ces éléments.

3.4 Choix des descripteurs de propriétés pour les facettes

3.4.1 Descripteurs temporels

Au niveau source. La date est indiquée dans le source comme dans les exemples de descripteurs de clés ci-dessus.

Au niveau prototype. Nous pouvons la représenter dans le contexte de plusieurs façons : initialement comme un type chaine ; mieux comme un type date, permettant une granularité de recherche (par année, par année et mois, etc.).

Un avantage de cette dernière représentation est aussi de pouvoir considérer les sous-contextes par année, avec leur nombre d'objets (et éventuellement d'ordonner les années dans l'index par nombre d'objets décroissant).

3.4.2 Descripteurs pour le type de document

Les décisions sont de différents types, nous en considérons quatre : QPC, DC, AN et I. Il s'agit du premier critère formel à prendre en compte. On y distingue trois groupes qui n´ont pas la même structure : 1) QPC et DC ; 2) AN ; 3) I.

Le type de décision est mentionné après le numéro de la décision.

L'extraction automatique est simple, à partir de ce numéro (figurant à plusieurs endroits, dont le nom de fichier).

Nous les représentons comme des attributs de base, ils sont regroupés en hiérarchie taxonomique (sous Categorie), par des axiomes ajoutés au contexte tels que : `axiom QPC, Categorie`

On peut traduire le groupe intermédiaire par une hiérarchie à 3 niveaux ainsi :

```
axiom QPC, QPC-DC
axiom DC, QPC-DC
axiom AN, Categorie
axiom I, Categorie
axiom QPC-DC, Categorie
```

3.4.3 Descripteurs envisagés selon le type de document

Pour des critères plus substantiels, il faudrait distinguer selon le type de décisions:

- pour celles du premier groupe (QPC-DC), on pourrait retenir comme première critère l'objet de la décision (la loi contrôlée), mais aussi les principes constitutionnels invoqués (égalité par exemple) voire le résultat de la décision (que l'on trouve au dispositif: conforme, non conforme).

- pour celles du deuxième groupe (AN), il s'agit du contrôle des élections "concrètes", les critères pourraient reposer sur l'objet de la contestation: propagande de campagne (bulletins, affiches etc.), opérations de vote et comptes de campagne.

- pour celles du dernier groupe (I), il s'agit de déterminer si la profession d'un individu est compatible avec les fonctions de député. Il existe des textes à cet égard mais ils sont souvent précisés par des décisions de justice. En l'occurrence, il y aurait un critère de sélection reposant sur les professions sur lesquelles le Conseil constitutionnel se prononce.

Les décisions de justice ont souvent une structure formelle qui facilite les choses. Cependant, elle n'est pas systématique et est souvent variable...

3.4.4 Descripteurs envisagés selon le « problème juridique»

Procéder par «problème juridique» permettrait de réduire à des résultats quasi binaires certains des résultats mixtes. Un balisage dédié pourrait être utile. Nous listons des caractéristiques à prendre en compte selon le type de décision.

Pour une décision QPC-DC : Les décisions DC et QPC sont comparables et sont les plus complexes. Pour une décision DC :

- Une décision DC peut trancher sur plusieurs « dispositions » ;

- Une disposition peut être contestée au regard de 1 ou plusieurs principes constitutionnels ;

- Pour chaque couple disposition + principe constitutionnel, le résultat peut être « non conforme », « conforme », conforme « sous réserve » (qui correspond à conforme à condition d'être appliqué ou interprété de telle ou telle manière). Les variantes « contraire »/ « non-contraire » peuvent être assimilées à conforme/non conforme. Pour une représentation peut-être plus simple, on pourrait casser ce résultat en deux résultats binaires si cela peut rendre les choses plus simples : « non-conforme »/ « conforme » puis « conforme »/ « conforme sous réserve ».

Pour une décision AN : - Une décision AN n'évoque jamais qu'une seule élection. Comme pour les décisions I, la clé unique est en fait la « personne » ou l'élection jugée. Là aussi l'information est non pertinente d'un point de vue scientifique. Pour gagner en précision, peut-être pourrions-nous descendre ici encore pour plus de précision.

Une décision AN peut évoquer plusieurs problèmes (comptes de campagnes, opération électorales, propagande par ex.)

- Un problème a un résultat binaire : « régulier » / « irrégulier » (et si dans la décision, il y a des irrégularités substantielles), il y a annulation, inéligibilité etc.

Pour une décision I : - une décision I peut théoriquement évoquer plusieurs professions, même si le plus souvent il n'y en a qu'une. La clé est la personne « jugée » mais cette information n'est pas pertinente d'un point de vue scientifique ;

- pour chaque profession, le résultat est binaire « compatible »/ « incompatible ».

3.5 Observations et difficultés

3.5.1 Format (html,..) et modèle du document

Une première difficulté de traitement automatisé tient au fait qu'un document source ne respecte pas toujours strictement le modèle attendu. Cet échec se manifeste dès le début, si on tente d'appliquer un outil reposant sur les technologies XML/html, au document html fourni.

Des outils informatiques existent pour "nettoyer" un fichier dans le but de le rendre compatible avec le format attendu. Mais nous observons lors d'un premier test, que la commande "tidy" reste silencieuse.

La solution provisoire est alors de recourir à des outils plus rudimentaires (comme "sed", procédant par motif sur chaque ligne) et ne reposant pas sur la technologie XML. Une solution plus satisfaisante serait d'appliquer et d'ajuster des méthodes de *nettoyage de données* (en anglais "data cleansing").

3.5.2 Rendu d'information sémantique

Le rendu d'une information particulière par une facette peut poser plusieurs problèmes. C'est le cas du résultat de décision de conformité que nous voulons indiquer pour les documents de type QPC et DC.

Nature du résultat. Le résultat attendu est a priori "conforme" ou "non conforme", cependant un résultat peut être :
- mixte, comme dans la décision n° 2008-573 DC du 8 janvier 2009
- avec des réserves, comme dans la décision n° 2011-628 DC du 12 avril 2011

Expression du résultat. On observe des variantes pour exprimer la conformité, par "... est conforme..." / "... n'est pas conforme ...", mais aussi :
- "... n'est pas contraire à ...", comme dans la décision n° 2010-602 DC du 18 février 2010

3.5.3 Représentation d'entités

La désignation d'une entité, d'un lieu varie selon le temps, par exemple : la désignation de département (facette *Comment_Place* dans notre prototype) est *Loire* ou *Loire Atlantique* selon la période.

Il s'agit là d'un problème non spécifique, et plus généralement la désignation varie selon la terminologie, l'expression employée (expressions synonymes, approximations).

Une grande variabilité, explicitée ou non, existe dans les données, sur le texte brut ou les codes utilisés.

3.6 Qualités attendues

Pour un système de qualité, certaines difficultés de différents ordres devront être prises en compte, nous en avons listé ci-dessus. Nous complétons avec d'autres aspects relevant de l'informatique qui concernent notre prototype et les données de l'expérience.

3.6.1 Traitement

Sur des données de nature ouverte et évolutive (avec de nouvelles décisions publiées), le traitement devrait être facilement reproductible (et aussi ouvert et adaptable).

En même temps, la complexité du système doit être "raisonnable" : en temps et place de calcul (pour une bonne interface) ; on peut aussi prendre en compte l'énergie (par exemple limiter les accès web).

Le traitement devrait produire un résultat à la fois *conforme* aux sources (et permettre d'y accéder) et *couvrant* (éviter de laisser des documents de côté).

Pour cela, un apport important du traitement automatique des langues (TAL) est à prévoir. On trouve par exemple : « n'est pas incompatible » dans la decision n^o 96-16 I du 19 décembre 1996

On trouve aussi cette écriture inhabituelle ailleurs : D É C I D E (avec des espaces).

3.6.2 Acceptabilité

Pour une meilleure aide à l'usager, il est important de s'assurer de l'intelligibilité du résultat, de choisir des structures et des nommages du contexte pertinents : pour les objets (ou unités) et pour les facettes (ou critères). Au-delà du contexte lui-même, les questions des manipulations permises (expressivité) et du temps de réponse (à une requête) avec le nombre de clics (pour une recherche) peuvent être déterminants pour l'acceptabilité du système.

4 Conclusion

Nous proposons un prototype avec d'une part des facettes générales (date, etc.) de nature objectives, factuelles et des facettes choisies par le juriste de nature subjectives, qui ont un impact sur une navigation orientée, simplifiée, sûre et la mise en évidence de caractéristiques.

La qualité/facilité de construction dépend cependant d'une certaine structuration dès le départ qui repose en partie sur un travail des juridictions elles-mêmes. Même si les décisions de justice présentent une certaine structuration formelle repérable grâce à des mots-clés, une structuration plus explicite faciliterait grandement l'exploitation de ces sources (xml notamment). En outre, des difficultés de rendu peuvent apparaître selon la nature et la représentation de l'information (permettant une classification aisée ou non, avec des données régulières ou non, une terminologie disponible ou non). C'est le cas par exemple pour le concept de fonction (activité) ... Là encore, si on ne peut demander aux juridictions de toujours employer les mêmes termes, une structuration plus explicite dès la rédaction permettrait de grandement améliorer les systèmes.

Du côté informatique, ce type d'étude permet de tester des modèles formels pour la représentation et l'extraction de l'information et l'intelligibilité des données.

Pour les travaux futurs, il faudrait prendre en compte les points suivants :

- l'évolution du style de rédaction (FR recul de l'imperatoria brevitas), perte des mots-clés typiques pour la structure ("considérant", "attendu") qui nécessiterait un appui du traitement automatique des langues (TAL).
- l'européanisation et l'internationalisation des sources du droit (droit comparé, problématique transnationale) et le référencement des décisions au niveau européen (ECLI (European Case Law Identifier))
- une terminologie ouverte et complète des termes juridiques de référence (avec leurs variantes).

Du côté droit, une telle étude permet d'envisager une amélioration des méthodes de travail des chercheurs et augure d'un possible gain de productivité. Elle oblige aussi à une certaine réflexivité sur les objets étudiés afin de permettre la construction de systèmes d'information adaptés. Dans cette perspective, l'association de chercheurs en droit et d'informaticiens est une étape intermédiaire avant que les juristes n'acquièrent un peu plus d'autonomie.

	Conseil d'état	Cour de Cassation	Conseil Constitutionnel	Legifrance [et Dalloz , Lexis Nexis]
interface dédiée	oui (Ariane)	non	oui (rech.exp.)	oui (rech.exp.)
arrêts et décisions compilés	jurisprudence interne (incomplet)	jurisprudence interne (incomplet)	jurisprudence interne (complet)	toutes juridictions (séparé pour Légifrance)
recherche transversale (types de décision)	sans objet	sans objet	oui	oui
(juridictions)	sans objet	sans objet	sans objet	non (Legifrance) oui (Dalloz, Lexis Nexis)
filtrage sur critères formels (N^o, date etc.)	oui	limité	oui	oui
"perspective" (critères de fond)	classification interne (Leb.)	classification interne (Bull.)	version "papier"	reprise des classifications internes des juridictions
recherche textuelle ("langage" de requêtes)	très limitée aucun	non	très limitée aucun	$\leq$ 5 termes (et/ou, 1 sauf) − err. sur caract. spéc. [$\leq$ 4 (et/ou/sauf) Dalloz] [$\leq$ 5 (et/ou/sauf) Lexi360]
données structurées	non	non	structuration légère	non

Figure 4: Caractéristiques et limites des systèmes actuels

A Annexe - Présentation des systèmes d'information actuellement utilisés par les chercheurs en Droit

Pour l'essentiel, les outils utilisés par les chercheurs en droit ne sont pas produits par eux ou très indirectement. En tout état de cause, il n'en ont pas la maîtrise. En effet, ceux-ci sont principalement constitués des sites officiels des juridictions, des bases de données d'éditeurs juridiques (Dalloz, Lexis Nexis, Francis Lefebvre etc.) et du site Légifrance. D'un point de vue épistémologique, ces sources présentent toutes des inconvénients :

- elles sont d'avantage orientées vers la diffusion des décisions de justice plutôt que vers l'intelligibilité de l'information juridique ;
- elles sont incomplètes ou parcellaires ;
- elles ne compilent pas des données brutes mais plutôt des textes ou des classifications internes des juridictions ;
- leurs caractéristiques techniques peuvent limiter le chercheur dans ses démarches etc.

A.1 Limites techniques des sites officiels des juridictions.

Les sites officiels des juridictions permettent l'accès aux arrêts et décisions de justice. Force est cependant de constater que pour au moins le Conseil d'Etat et la Cour de cassation, ils n'ont pas vraiment vocation à assurer cette fonction de diffusion et de publication.

Conseil d'Etat et Arianeweb. Sur le site du Conseil d'Etat, l'interface dédiée est "ArianeWeb"[4]. Le mode avancé permet principalement de filtrer selon des critères formels (numéro de décision, date etc.). La limitation principale se manifeste lors de recherches fondées sur des critères substantiels qu'elles soient transversales et/ou thématiques. Pour que le système lui retourne des résultats fondés sur des critères de fond, l'utilisateur peut utiliser la classification du Conseil d'Etat. Toutefois, n'est classée qu'une petite fraction des arrêts (ceux publiés ou mentionnés au Recueil Lebon) et la logique de la classification échappe au chercheur voire est parfois incohérente. L'autre possibilité pour l'utilisateur est de recourir à la recherche textuelle très limitée pour laquelle l'outil n'est manifestement pas conçu (pas d'équation de recherche, seulement pluriels et synonymes).

[4] http://www.conseil-etat.fr/Decisions-Avis-Publications/Decisions/ArianeWeb

ArianeWeb n'est en fait qu'une version bridée du logiciel professionnel utilisé par les personnels du Conseil d'Etat eux-mêmes. Il répond donc avant tout à leurs besoins spécifiques et s'adresse donc plutôt à des utilisateurs ayant une idée très précise de ce qu'ils cherchent ou ayant déjà consulté une autre source documentaire et qui viendraient compléter leurs informations grâce à cet outil.

Site officiel de la Cour de cassation. Sur le site de la Cour de cassation[5], il n'y a pas d'interface spécifique, simplement une page "jurisprudence". Le site classe les arrêts en fonction de critères formels (formations de jugement, dates, numéros de requête, objet normalisé du litige, résultat). Aucune recherche textuelle n'est possible. Pour une recherche thématique ou transversale, l'utilisateur doit utiliser la classification "par rubrique" effectuée par la Cour et qui est inachevée (au mieux les arrêts remontent aux années 2000). Il peut aussi naviguer à tâtons s'il a une certaine connaissance des compétences de chacune des formations et chambres de la Cour. Mais pour une recherche plus précise, il doit s'appuyer sur les publications aux bulletins de la Cour qui ne font l'objet que d'une numérisation très insuffisante (un pdf de la version papier est téléchargeable sur le site; le prétendu bulletin numérique ne fait que reprendre les arrêts cités dans la version papier selon des critères temporels et sans permettre de recherche par mots clés par exemple).

Site officiel du Conseil constitutionnel. Le site du Conseil constitutionnel fait presque exception[6]. Ayant fait l'objet d'une réfection récente, il donne accès à des fonctionnalités un peu plus avancées. Le classement se fait toujours selon des critères formels mais le mode avancé permet la recherche selective ou transversale selon le type de décision. Il permet aussi la recherche textuelle sur des blocs spécifiques des décisions (visas, motifs, dispositifs), signe de la présence d'un balisage plus fin des décisions. La recherche textuelle reste cependant limitée (pas d'équation de recherche) et la normalisation des décisions n'est pas achevée (certaines décisions du même type ne sont pas classées ensembles). De plus, si la recherche transversale sur différents types de décision est permise, il n'existe sur le site aucune classification fondée sur des critères substantiels. Pour y accéder, l'utilisateur doit télécharger les tables analytiques du Conseil constitutionnel qui n'existent qu'en pdf.

A.2 Limites techniques de Légifrance.

Le site Légifrance[7] a vocation à permettre une meilleure diffusion du droit en général. Concernant la question plus spécifique des arrêts et décisions de justice, il permet un recensement souvent plus complet que les sites officiels des juridictions car il donne aussi accès à des arrêts mineurs. Son intérêt principal vient surtout de la recherche textuelle qu'il autorise alors que les sites officiels des juridictions ne la permettent pas ou seulement de manière très limitée. Il permet aussi d'utiliser la classification interne des juridictions.

S'il est un complément utile, Légifrance présente toutefois des défauts importants pour un chercheur:

- il n'autorise pas la recherche simultanée sur les fonds de la juridiction administrative, de la juridiction judiciaire voire éventuellement du Conseil constitutionnel alors même qu'une des caractéristiques de la législation moderne est de générer des problématiques ressortissant de plusieurs branche du droit ;

- les critères de recherche sont principalement formels (date, numéro de requête, publication etc.)

- les filtres de recherche textuelle sont souvent limités dans leur efficacité; l'absence de balisage formel des arrêts et décisions empêche la limitation du champ à des blocs spécifiques des arrêts; l'équation de recherche est limitée à 5 termes (4 et/ou et 1 exclusion); certains caractères "spéciaux" génèrent des erreurs.

- puisqu'il s'adresse à un public large, Légifrance adopte une présentation sans perspective ou presque.

[5] https://www.courdecassation.fr/

[6] http://www.conseil-constitutionnel.fr/
http://recherche.conseil-constitutionnel.fr/?expert=2

[7] https://www.legifrance.gouv.fr/

A.3 Limites techniques des systèmes d'information payants.

La recherche juridique s'appuie par ailleurs sur des systèmes d'information produits par des éditeurs juridiques, les principaux étant Dalloz et Lexis Nexis. La plus-value de ces outils vient surtout du fait qu'ils permettent d'accéder en même temps aux décisions de justice et aux articles de doctrine qui peuvent en constituer le commentaire. Mais il est rare que leurs systèmes d'information ajoutent beaucoup à ceux déjà existants et en particulier à Légifrance.

Le site Dalloz[8] reprend globalement les fonctionnalités du site Légifrance en y ajoutant la recherche transversale sur toutes les juridictions mais réduit l'équation de recherche à 4 membres et les articulations et/ou/sauf.

Lexis360 permet la recherche textuelle transversale, une équation de recherche à 5 membres avec les articulations et/ou/sauf et des critères de proximité. Il superpose par ailleurs les analyses effectuées par l'éditeur lui même dans sa base de donnée propre appelée jurisdata.

A.4 Bilan

Globalement, les systèmes d'information actuellement utilisés par les juristes fonctionnent comme des banques d'arrêts et décisions. Tant que la recherche repose sur des critères formels basiques et référencés (date, numéro de décision etc.), ils sont relativement satisfaisants bien que perfectibles. Mais précisément, tous les critères formels ne sont pas retenus (par ex. composition de la juridiction, origine de l'appel, date de saisine etc.).

Reste cependant que le défaut majeur des systèmes d'information actuels vient de l'absence de critères de recherche substantiels et de l'insuffisance des moyens de compensation proposés. Les critères de classification substantiels par les juridictions posent des problèmes épistémologiques et interroge sur les rapports entre doctrine universitaire et doctrine organique des juridictions; la recherche textuelle demeure aléatoire et trop limitée.

A l'évidence, il ne peut y avoir une seule et unique classification substantielle. Toutes auront potentiellement des "défauts". La critères retenus révèlent en effet la perspective du chercheur. Mais au moins, cette perspective est explicite. Mieux, elle constitue la plus-value d'une étude qui réside justement dans un travail de classification et de modélisation qui ne se borne pas à reprendre des critères formels non pertinents ou substantiels mais imposés par d'autres.

[8] http://www.dalloz.fr/
http://www.dalloz.fr/documentation/Recherche?ctxt=0_dCRzMD1ET0NUwqdkJG5UZXh0ZTI9My8xMQ==

Gestion Pro-Active des Obligations Contractuelles

Manuel Munier
(informatique)
Univ Pau & Pays Adour / E2S UPPA
LIUPPA, EA 3000
Email : manuel.munier@univ-pau.fr

Vincent Lalanne
(informatique)
Univ Pau & Pays Adour / E2S UPPA
LIUPPA, EA 3000
Email : vincent.lalanne@univ-pau.fr

Xavier Daverat
(droit privé)
Université de Bordeaux
Email : xavier.daverat@wanadoo.fr

***Résumé*—En gestion de projet il est de plus en plus fréquent d'impliquer le client final dans le processus de conception. C'est la vision agile. Du point de vue de la contractualisation, il est nécessaire de préciser les obligations de chaque partie. Il serait néanmoins dommage que ces contrats restent des documents "statiques". Notre approche consiste à superviser la phase d'exécution de ces contrats agiles pour assurer une gestion pro-active des obligations contractuelles, favorisant ainsi une collaboration plus réactive et une meilleure adaptation aux évolutions.**

I. Introduction

Le digital impose un nouveau rythme et redéfinit les organisations, les processus, la culture, les valeurs, les pratiques et les comportements. Alors que les contrats de développement ou d'intégration classiques réservent à l'expression initiale des besoins un rôle déterminant, notre approche consiste à (re)placer la valeur au centre du processus décisionnel de l'entreprise, ce qui implique de mettre en place un modèle de contractualisation et un mode de pilotage des projets centrés sur la valeur créée, ceci dans une démarche collaborative entre client et prestataire.

Porteuses de grandes promesses, les méthodes agiles impliquent néanmoins des droits et devoirs pour les deux parties, qu'il est nécessaire de consigner dans une cadre légal adéquat. Dans cet article nous présentons une brève réflexion sur la définition des obligations contractuelles dans le cadre de tels projets et sur l'opportunité de concevoir un outil informatique dédié à la gestion pro-active de ces obligations au travers d'indicateurs et de tableaux de bord.

II. Problématique

Afin d'illustrer notre problématique nous prendrons comme cas d'étude la conception d'un système informatique. Le client fait appel à un prestataire de services pour le développement de son nouveau système numérique. Bien évidemment, cette relation entre le client et le prestataire fait l'objet d'un contrat. Notre approche consiste à impliquer le client dans toutes les phases de conception et de développement du système afin d'éviter des problèmes (malheureusement) bien connus en génie logiciel que sont le non respect des délais, une couverture que partielle des besoins exprimés, le "big bang" en fin de projet, la non adéquation aux habitudes de travail des utilisateurs finaux,...

Afin de stimuler la collaboration entre le client et le prestataire, nous proposons de mettre en place des contrats agiles (en anglais *smart contracts*) où, en quelque sorte, l'objet du contrat ne sera plus seulement le produit mais intégrera également la "validation" par le client. Pour ce faire, il faut prévoir des clauses spécifiques permettant au prestataire de consulter le client sur des versions intermédiaires, de lui poser des questions pour préciser tel ou tel point,... De son côté, le client devra (de par les-dites clauses) répondre rapidement, signaler d'éventuels bugs, informer le prestataire au plus tôt s'il demande des modifications et/ou s'il désire ajouter des fonctionnalités (par rapport à l'expression des besoins réalisée initialement). Un outillage informatique orienté gestion de projet permettra de suivre ces échanges en vu d'améliorer le processus génie logiciel de développement. Ce faisant, tant le client que le prestataire aura sa part de responsabilité et, le cas échéant, la traçabilité des échanges pourra permettre la pré-constitution de preuves en cas de litige ultérieur.

Dans un telle approche chaque partie pourra y trouver son intérêt. Du côté du prestataire, nous pouvons lister les bénéfices suivants :

- Validation au plus tôt par le client et implication du client dans les choix de conception.
- Communication avec les usagers dès le départ : on leur demande leur avis ! Ceci facilitera d'autant leur formation sur le nouveau produit que le déploiement lors de la mise en production. L'objectif est que tous les éventuels problèmes aient déjà été rencontrés lors des validations intermédiaires et qu'ils aient été solutionnés.
- Ces échanges permettront au prestataire d'avoir des retours sur la "qualité de service" au fil de l'eau, notamment du point de vue des objectifs atteints sur les fonctionnalités implémentées par rapport aux fonctionnalités envisagées sur l'échéancier (et donc le contrat).
- En termes de gestion de projet sa prise de risque est donc moindre, ce qui pourra lui permettre de diminuer ses coûts de développement et par conséquent d'être plus compétitif. La pratique du forfait à échéance des livrables peut d'ailleurs être envisagée. En cas de dysfonctionnement nous aurons un partage de responsabilité(s) avec le client, les retours de ce dernier étant devenus des obligations contractuelles.

Cette approche augmente largement la responsabilité du client. Le prestataire pourra en effet évoquer des manquements du client à ses obligations. Par exemple, s'il a des difficultés

à tenir les délais, le prestataire pourra reprocher au client sa non participation ou une participation insuffisante dans le processus, une mobilisation abusive du personnel du prestataire [1],... Cette implication dans le projet apportera néanmoins des bénéfices significatifs au client :

— Le client est impliqué dans le processus de développement, on lui demande comment il travaille, quels sont ses retours d'expérience,... Ceci est d'ailleurs un des fondements des méthodes agiles en pilotage et réalisation de projets [2].

— Certes cela génère une surcharge de travail pour le client et lui impose de prendre ses responsabilités quant aux retours qu'il fera au prestataire (réponses aux questions, réactivité, signalement de bug,...), mais le produit final correspondra mieux à ses besoins.

— En outre, vu que la prise de risque est moindre pour le prestataire, le client peut espérer que les coûts de développement facturés en seront diminués d'autant.

— Si une véritable coopération se met en place, il y est fort probable que le temps de réalisation puisse être sensiblement écourté.

Afin de pouvoir valider cette approche il nous faudra procéder en deux temps. Tout d'abord, via l'outillage juridique adéquat, nous devons exprimer sous forme de clauses ces obligations contractuelles des différentes parties. Deuxièmement, il est nécessaire de mettre en place une infrastructure numérique capable de surveiller la bonne exécution de tels contrats agiles, soit de manière complètement automatisées (ex : contrats intelligents ou *smart contracts*), soit de manière semi automatisée (ex : GED avec calculs d'indicateurs et tableaux de bord pour des opérateurs humains).

III. Obligations Contractuelles

Pour la constitution du contrat nous supposerons respectés les principes de confiance mutuelle entre client et prestataires, de loyauté réciproque, d'obligation d'information. Comme nous l'avons indiqué précédemment, une première étape dans la rédaction du contrat concernera la modélisation des besoins exprimés, par exemple lors de la négociation du contrat. Ce n'est que dans un second temps que nous pourrons rédiger les clauses permettant, par exemple, au prestataire d'envoyer au client des versions à valider ou de lui poser des questions attendant des réponses précises. Dans ce cas précis, il faudra notamment fixer les critères de validation par le client à chaque étape (livrable) jusqu'à la recette finale.

Outre le travail plus "traditionnel" de rédaction des clauses contractuelles, nous voulons dans cet article mettre l'accent sur certaines consécrations issues du nouveau droit des contrats tel que réformé en 2016. Par consécration, on entend les intégrations légales de dispositifs issus de la jurisprudence.

1. À l'opposé, il ne faut pas qu'un client "spamme" le prestataire avec des retours inutiles juste pour lui faire perdre du temps ! Toutefois, un tel comportement devrait pouvoir être détecté automatiquement via l'outil informatique (ex : une GED ou Gestion Électronique de Documents)

2. Wikipédia : https://fr.wikipedia.org/wiki/Méthode_agile

La réforme du droit des contrats est issue de l'ordonnance n° 2016-131 du 10 février 2016 portant réforme du droit des contrats, du régime général et de la preuve des obligations. Cette ordonnance est entrée en vigueur le 1er octobre 2016. Elle écarte la définition romaine du contrat pour consacrer celle retenue par une partie de la doctrine, mais en ajoutant la vocation à modifier, transmettre ou éteindre (des obligations) : *"un accord de volonté entre deux ou plusieurs personnes destiné à créer, modifier, transmettre ou éteindre des obligations"*.

Parmi les dispositions introduites par cette réforme dans le droit de contrats, en voici quelques-unes qui nous intéressent plus particulièrement dans le cadre de nos travaux :

— Dans les dispositions liminaires :
  - Le principe de la liberté contractuelle est clairement posé à propos du choix de contracter, du contractant, du contenu et de la forme contractuels (art. 1102 nouv.).
  - L'obligation de bonne foi non seulement au titre de l'exécution du contrat mais également dans les négociations précontractuelles et la formation du contrat (art. 1104 nouv.).

— Dans la formation du contrat, notamment la phase des négociations :
  - Le principe de liberté contractuelle (y compris la rupture) et de bonne foi (art. 1112 nouv.).
  - L'obligation précontractuelle d'information (art. 1112-1 nouv.). Elle se loge dans la sous-section relative aux négociations, mais la dépasse sans doute pour concerner toute la formation du contrat.
  - L'obligation de confidentialité (art. 1112-2 nouv.).

L'expression des besoins que nous avons évoquée à la section précédente pourrait faire partie de la phase de négociation du contrat. Le nouveau droit des contrats impose la loyauté et le devoir d'information. Dans notre cas nous exploiterions cette disposition "dans les 2 sens" : le prestataire de service en tant que professionnel du numérique ; le client en tant que "professionnel" de son métier et donc avec l'obligation de bien expliquer au prestataire ses pratiques, ses attentes, ses contraintes,...

À titre purement illustratif, un exemple de clauses est donné à la section V. Cet exemple n'est qu'une première ébauche et mérite bien évidemment d'être approfondi.

IV. Outillage Numérique

Afin de pouvoir contrôler la bonne exécution du contrat nous proposons de mettre en place un outil informatique (une Gestion Électronique de Documents par exemple) qui supervisera et tracera les échanges entre les parties : *"qui"* a envoyé le message, *"quoi"* (ou quel est le contenu informationnel du message), *"quand"* il a été envoyé, *"délais"* de réponse,...

À l'image des plateformes dédiées à la gestion de projets, cette GED sera le nœud central pour toutes les formes de communication entre les acteurs. Comme nous l'avons indiqué à la section II, le prestataire pourra solliciter son client pour

lui demander de valider les fonctionnalités d'une version intermédiaire par exemple, ou encore pour lui poser des questions pour préciser tel ou tel point du projet. Tous ces échanges, y compris leurs métadonnées, seront tracés et, le cas échéant, pourront servir d'éléments de preuve en cas de litige ultérieur entre les parties.

C'est donc via cet outil que le client devra valider les différentes fonctionnalités au fur et à mesure qu'elles seront "publiées" par le prestataire (cf. méthodes de conception agiles). Il pourra ainsi signaler des dysfonctionnements (*bugs*). Il nous faudra pour cela prévoir des procédures de classification et de traitements de ces remontées d'information, tant d'un point de vue technique informatique (ex : rattachement d'un bug à une fonctionnalité) que d'un point de vue juridique (ex : identification des responsabilités, respect des engagements dictés par les clauses du contrat).

Le client pourra également être acteur du processus de développement en demandant au prestataire des améliorations ou des ajouts de fonctionnalités. Comme nous l'avons déjà indiqué plus tôt dans cet article, il faudra bien distinguer une amélioration (rattachée à une fonctionnalité existante) d'un ajout. Cette étape de mise en correspondance vis-à-vis des besoins exprimés est indispensable pour éviter qu'un client mal intentionné ne fasse passer un ajout de fonctionnalité pour une correction de bug !

Concrètement, nous devrons formaliser ces communications via un modèle de langage (ex : ontologies) afin de pouvoir assurer une traçabilité efficace de ces échanges entre les parties au contrat. Dans les processus très complexes (avec de nombreux échanges prévisibles) il faudra mettre en place un système d'administration (retours et liens clients/prestataire). Mais cette traçabilité ne servira pas que d'historique car notre objectif est de mettre en œuvre une gestion pro-active des obligations, c'est-à-dire d'évaluer tout au long de la phase d'exécution du contrat que les clauses sont bien respectées et, le cas échéant, de signaler tout comportement déviant. Au fur et à mesures des échanges, il sera également possible d'activer ou de désactiver certaines clauses du contrat (contrats agiles). Étant donné que nous aurons défini une sémantique formelle pour les communications, nous pourrons prévoir des outils informatiques pour réaliser un reporting régulier. Par exemple, en fonction des itérations, il serait intéressant de mesurer ce qui reste à accomplir pour atteindre les fonctionnalités initialement prévues (une sorte de tableau de bord du projet). Dans le même ordre d'idée, nous pourrions envisager d'automatiser des tests à chaque livrable (logique smart).

Pour la conception et la mise en œuvre de cet outillage numérique nous nous appuierons tout d'abord sur l'expérience acquise quant à l'utilisation d'ontologies pour la formalisation des contrats dans les architectures orientées services [1][2][3]. Du point de vue de la traçabilité des échanges, de la collecte de métadonnées ou du calcul d'indicateurs pour la génération de tableaux de bord, nous nous baserons cette fois-ci sur nos travaux publiés dans [4] et [5] ainsi que sur nos expérimentation réalisées avec la société BackPlan, notamment dans le cadre des carnets d'avancements numériques.

V. Exemples de Clauses

À titre purement illustratif, voici un exemple de clauses. Celui-ci n'est qu'une première ébauche et mérite bien évidemment d'être approfondi.

Article XXX. - Mise en œuvre du contrat

XXX. 1 - Période préparatoire

[...]

XXX. 2 - Période de développement

La période de développement est constituée de huit Sprints de trois (3) semaines chacun. Elle s'achève par la présentation du développement du Produit au Comité de pilotage, puis sa réception par le CLIENT selon la procédure de recette visée à l'article XXX des présentes.

Il est possible, sur décision du Comité de pilotage, d'ajouter des Sprints supplémentaires en cas de révision du Product Backlog.

L'atteinte anticipée des objectifs du CLIENT permet de clore la période de développement selon la procédure prévue à l'article XXX des présentes.

[...]

Article XXX. - Gouvernance du projet

XXX. 1 - Responsables de projet

Le PRESTATAIRE désignera un « Responsable de Projet PRESTATAIRE » dans un délai maximal de XXX à compter de la signature des présentes. Le « Responsable de Projet PRESTATAIRE » assure la fourniture au CLIENT des prestations faisant l'objet des présentes, élabore et met à disposition les indicateurs visés dans le Plan Qualité de Service et co-préside le Comité de pilotage avec le « Responsable de Projet CLIENT ».

Le CLIENT désignera un « Responsable de Projet CLIENT » dans un délai maximal de XXX à compter de la signature des présentes. Le « Responsable de Projet CLIENT » assure le PRESTATAIRE de sa collaboration pleine et entière pour l'exécution des prestations faisant l'objet des présentes. Il s'engage notamment à mettre à disposition le personnel nécessaire à la réalisation des prestations et à la mise en œuvre de la gouvernance du projet. Il met également à la disposition du PRESTATAIRE les moyens et informations nécessaires à la parfaite exécution du présent contrat. Il assure le suivi de la fourniture des prestations par le PRESTATAIRE sur la base des indicateurs visés dans le Plan Qualité de Service et co-préside le Comité de pilotage avec le « Responsable de Projet PRESTATAIRE ».

XXX. 2 - Comité de pilotage

Le Comité de pilotage est constitué du « Responsable de Projet CLIENT », du « Responsable de Projet PRESTATAIRE », du responsable de l'Équipe de développement du PRESTATAIRE, de l'ingénieur responsable du service informatique du CLIENT, du directeur commercial du PRESTATAIRE ou de son représentant, du directeur financier du CLIENT ou de son représentant.

Il est expressément convenu par les Parties que le directeur commercial du PRESTATAIRE ou son représentant, et que le directeur financier du CLIENT ou son représentant disposent du pouvoir d'engager la Partie qu'il représente.

XXX. 3 - Cellule de travail

Le Comité de pilotage décide de l'intégration des équipes du CLIENT et du PRESTATAIRE nécessaire au développement du projet par constitution d'une Cellule de travail. Il désigne les participants et détermine leur fonction. La Cellule se réunit tous les deux jours ouvrables pendant trente (30) minutes.

Une réunion de fin de Sprint est organisée par la Cellule quarante-huit (48) heures au plus tard avant la fin de chaque Sprint. La Cellule dresse le bilan du développement du projet et transmet, le cas échéant, des propositions de modification du Product backlog au Comité de pilotage.

Toutes les réunions de la Cellule se tiennent dans les locaux du CLIENT. Il appartient au PRESTATAIRE de prévoir les autorisations nécessaires aux déplacements de ses salariés hors de ses locaux.

Article XXX. - Plan Qualité de Service

XXX. 1 - Adoption du PQS

Le Plan Qualité de Service est élaboré pendant la période de lancement, et au plus tard dans les trois (3) mois à compter de la signature des présentes. Il est arrêté par le Comité de pilotage.

XXX. 2 - Contenu du PQS

Le Plan Qualité de Service détermine le niveau de service fourni par le PRESTATAIRE au CLIENT.

Le Plan Qualité de Service fixe l'objet des Sprints. Il précise le niveau de service requis à chaque étape du développement du projet.

Il contient notamment :

- Les missions et les responsabilités des Parties, telles qu'elles sont réparties dans les Sprints.
- La liste des Indicateurs avec toutes les caractéristiques qui y sont associées ; les Indicateurs ne seront pas modifiés en cours d'exécution des prestations prévues au présent contrat afin de conserver une mesure de la qualité de service selon une forme pérenne.
- La nature des manquements aux niveaux de service requis, le mode de calcul et le régime des pénalités qui y sont afférentes ; les pénalités seront reprises dans un procès-verbal établi lors de la réunion utile du Comité de pilotage avec indication, le cas échéant, des pénalités qui ne seraient pas appliquées. La procédure de paiement des pénalités est détaillée dans le PQS.

VI. Conclusion

Les méthodes agiles (ex : Scrum [3]) ont libéré les contrats de développement et d'intégration de leur carcan trop rigide et formaliste qui repose sur la connaissance a priori du périmètre du projet et de sa stabilité dans le temps. Toutefois, au rythme des changements auxquels sont soumises les organisations, cette stabilité du périmètre est devenue un leurre qui fait dériver les projets ou mène au contentieux.

Dans ce court article nous avons présenté vision sur la base d'un contrat agile où nous formaliserions la sémantique des clauses contractuelles à l'aide d'ontologies afin de pouvoir concevoir une plateforme de gestion de projet dédiée à la supervision de la phase d'exécution du contrat. En s'appuyant, entre autres, sur les informations de traçabilité des échanges entre les parties, cette plateforme permettra une gestion proactive des obligations contractuelles. En effet, le contrat ne doit pas constituer un frein à l'agilité mais être flexible et s'adapter à ce type de projets.

À l'issue de cette réflexion, deux questions importantes à nos yeux n'ont pas été traitées. La première concerne la rédaction, en amont, d'un contrat-cadre pour définir les clauses contractuelles générales, la collaboration, la gouvernance et les mécanismes d'arbitrage et de gestion du changement, que ces changements soient à l'initiative du client (évolution de périmètre) ou du prestataire (évolution de sa performance). La seconde est relative à la titularité des droits de propriété intellectuelle sur les créations réalisées dans un contexte proactif, étant donné que le client se retrouve impliqué dans le processus de conception.

Remerciements

Nous tenons à remercier les organisateurs de ces "Convergences du Droit et du Numérique", édition Bordeaux 2017 [4]. Cet événement, tant dans son contenu que dans sa forme, est un lieu idéal pour établir des collaborations entre les professionnels du droit et ceux du numérique.

Références

[1] G. E. Jaramillo, M. Munier, and P. Aniorté, "Du contrôle de la collaboration humaine vers des contrats de service sémantiques pour la sécurité de l'information," *Ingénierie des Systèmes d'Information*, vol. 22, no. 1, pp. 43–64, 2017. [Online]. Available : https://doi.org/10.3166/isi.22.1.43-64

[2] E. Jaramillo, "Un modèle sémantique de contrat et un processus basé sur la connaissance pour garantir la contrôlabilité dans les approches orientées services," Ph.D. dissertation, UPPA - ED211 Sciences Exactes et leurs Applications, 12 2016.

[3] E. Jaramillo, P. Aniorté, and M. Munier, "Service Contracts : Beyond Trust in Service Oriented Architectures," in *34ème Congrès INFORSID, Atelier "Sécurité des systèmes d'information : technologies et personnes", SSI'16*, 2016.

[4] V. Lalanne, "Gestion des risques dans les architectures orientées services," Ph.D. dissertation, UPPA - ED211 Sciences Exactes et leurs Applications, 12 2013.

[5] V. Lalanne, M. Munier, and A. Gabillon, "Information security risk management in a world of services," in *PASSAT*, 2013.

3. Wikipédia : https://fr.wikipedia.org/wiki/Scrum_(développement)

4. CDN'2017 : http://cdn.u-bordeaux.fr/

Quels droits sur les données numériques ?

Rose-Marie Borges, Maître de conférences en droit privé, Université Clermont-Auvergne

Manuel Munier, Maître de conférences en informatique, Université de Pau et Pays de l'Adour

Introduction

La transformation digitale du monde à laquelle nous assistons se traduit par la mise en données de la réalité, conduisant à l'apparition d'une nouvelle matière première : la donnée. Celle-ci devient la nouvelle unité de base d'un capital que les entreprises se constituent et valorisent en organisant sa circulation[1]. Toutefois, la simple collecte de données brutes ne suffit pas à donner une valeur à celles-ci. La valeur principale des données réside dans le traitement de celles-ci et dans les métadonnées qui leur sont associées. Le modèle économique de la majorité des réseaux sociaux et des moteurs de recherche repose sur la mise à disposition d'un service gratuit en contrepartie d'une collecte de données de l'utilisateur pour une utilisation commerciale : il y a alors échange de valeur entre les données fournies par l'utilisateur et l'accès au réseau[2]. Cependant, la valeur reçue par le citoyen en échange de ses données est bien moindre que la valeur qu'il concède à l'opérateur. Facebook, pour ne citer que lui, dégage des millions de dollars de profits en vendant les données qu'il collecte ce qui ne correspond en rien au coût du service fourni aux utilisateurs de ce réseau social. Chez les data brokers[1] aux Etats-Unis, chaque donnée personnelle fait l'objet d'une évaluation monétaire[2].

Cette contribution est une ébauche de réflexion sur une problématique : comment partager les fruits résultant de l'exploitation de données fournies gratuitement ? Nous envisagerons d'abord les différents types de données numériques (I) avant de nous interroger sur la nature des droits que l'on peut appliquer à ces données (II).

1 Les data brokers ou courtiers de données, sont des entreprises dont l'activité commerciale consiste à revendre des données à des annonceurs ou des prestataires marketing.

2 https://www.challenges.fr/high-tech/vos-donnees-personnelles-sur-internet-peuvent-valoir-de-l-or_57690 ; https://www.challenges.fr/finance-et-marche/facebook-rompt-avec-certains-courtiers-de-donnees_577178

I. Les différents types de données numériques

1.1 Les données numériques

Avant toute chose, il convient de s'entendre sur la notion de donnée numérique.

L'enrichissement du vocabulaire informatique[3] définit la donnée comme la « représentation d'une information sous une forme conventionnelle destinée à faciliter son traitement (en anglais : *data*) ».

Au-delà de cette définition datant de 1981, nous pouvons également nous référer à celle de l'administrateur général des données[4] (AGD) datant de 2014, équivalent d'un *Chief Data Officer* au sein de l'État, pour qui « une donnée numérique est la description élémentaire de nature numérique, représentée sous forme codée, d'une réalité (chose, événement, mesure, transaction, etc.) en vue d'être :

- Collectée, enregistrée
- Traitée, manipulée, transformée
- Conservée, archivée
- Échangée, diffusée, communiquée

Selon leur destination, les données peuvent être « fermées », c'est-à-dire réservées à quelques personnes ou à des organisations, « partagées » ou « ouvertes » à tous utilisateurs. Lorsqu'elles sont partagées ou ouvertes, les conditions d'utilisation des données doivent être contractuellement claires (licences, conditions générales d'utilisation…).

Il convient dès lors de distinguer la donnée de l'information, avant d'aborder les différentes catégories de données.

1.1.1 Donnée et information

---

3 Arrêté du 22 décembre 1981 relatif à l'enrichissement du vocabulaire informatique : http://bbf.enssib.fr/consulter/bbf-1982-06-0355-009

4 Décret n°2014-1050 du 16 septembre 2014 : « L'administrateur général des données (AGD) coordonne l'action des administrations en matière d'inventaire, de gouvernance, de production, de circulation et d'exploitation des données par les administrations » : http://agd.data.gouv.fr/la-fonction/

Une information est un ensemble de données agrégées en vue d'une utilisation par l'homme. Afin de bien distinguer ces deux notions, prenons l'exemple de l'information « il fait beau ». Celle-ci est construite à partir de plusieurs données telles que la température, l'ensoleillement, le taux d'humidité, la force du vent, etc. Toutefois, en fonction du contexte (par exemple, géolocalisation des données en Aquitaine, à Tahiti et au Mont Blanc), le même jeu de données n'aboutira pas à la même information.

Une autre confusion usuelle est faite entre information numérisée et donnée.

Une information numérisée peut être copiée et collée informatiquement ; c'est le cas pour un texte dans une page internet ou l'extrait d'un document au format PDF. Cependant, sous cette forme, les informations ne peuvent pas directement être traitées dans un logiciel pour obtenir de nouveaux résultats (calculs, représentations graphiques...). Les informations sous forme numérisée ne sont donc pas nécessairement des données.

Une donnée, quant à elle, peut faire l'objet d'un traitement automatisé. Elle peut notamment être classée et stockée dans un tableau ou une base de données. Ces données sont (souvent) structurées. Dans le cas d'un tableur, la première ligne est dédiée aux en-têtes de colonnes et les lignes successives contiennent des informations. Chaque cellule contient une donnée. Dans le cas d'un graphe, les sommets contiennent les données « élémentaires » et les arcs représentent les liens « sémantiques » entre ces données. Ces arcs peuvent être étiquetés avec des métadonnées.

Lorsque l'information est organisée en donnée, son interprétation peut prendre plusieurs formes : des points sur une carte, un affichage de chiffres ou de lettres dans un texte, des relations entre personnes sur un réseau social, etc.

1.1.2 Catégories de données

Les données peuvent être catégorisées selon plusieurs critères : leur contenu, leur mode de collecte, leur architecture…

1.1.2.1 Donnée brute, donnée primaire, donnée enrichie

Les données brutes, aussi appelées données primaires, sont « les données non interprétées émanant d'une source primaire, ayant des caractéristiques liées à celle-ci et qui n'ont été soumises à aucun traitement ou tout autre manipulation »[5].

---

5 https://fr.wikipedia.org/wiki/Données_brutes

À l'inverse, une donnée enrichie est une donnée dont la pertinence s'est vue améliorée par un traitement :

- Amélioration du jeu de données par l'ajout d'un autre jeu de données ;
- Retraitement de la donnée suite à expertises ;
- Croisement de différentes données...

« Pour autant, parce qu'elle est construite par l'homme, une donnée est forcément le résultat d'un traitement, d'une action. Il est donc toujours important d'interroger le mode de production de chaque donnée.»[6]

1.1.2.2 Donnée contextualisée, métadonnée

Les données sont parfois enrichies par des informations de contexte qui peuvent être mentionnées dans un fichier texte : les critères de construction du jeu de données sont clairement expliqués, les différents intitulés de colonnes sont expliqués, les codes utilisés sont traduits pour être compréhensibles par un humain...

Ce travail sera d'autant plus exploitable s'il est traduit en métadonnées, considérées comme « des données sur les données ». Les métadonnées fournissent des informations permettant de comprendre les données (documents, images, bases de données), des concepts (systèmes de classification) et des entités du monde réel (les personnes, les organisations, les lieux, des produits...). Ces métadonnées peuvent prendre plusieurs formes :

- **métadonnées descriptives**, décrivant une ressource à des fins de découverte et d'identification ;
- **métadonnées structurelles**, décrivant par exemple des modèles de données et/ou des données de référence (métadonnées officielles, normalisées) ;
- **métadonnées administratives**, permettant la gestion, la validation et l'archivage d'une ressource.

1.1.2.3 Donnée liée

---

6 Lexique « les mots de l'infolab » : http://infolabs.io/mots-infolab

Au-delà du stockage des données, le principal intérêt de l'informatique et des traitements automatisés est de pouvoir corréler ces données en exploitant les liens sémantiques. On parle alors de données liées. De façon très synthétique, il s'agit de publier et de connecter des données structurées sur le web en utilisant des technologies web standard pour créer des connexions lisibles par ordinateur, permettant aux données provenant de différentes sources d'être connectées et interrogées, via les métadonnées notamment, contribuant ainsi à une meilleure interprétation et analyse. Le web sémantique et le big data sont deux technologies basées sur les données liées.

1.2 Architecture des données

D'une manière très générale, nous pouvons distinguer trois catégories d'acteurs dans le partage de données : les producteurs, qui fournissent les données « primaires », les centres de collecte et de stockage, et les consommateurs qui utilisent différentes sources de données pour alimenter leurs traitements automatisés. En pratique cependant, les rôles ne sont pas toujours si distincts :

- Un acteur peut consommer des données, effectuer des traitements, et publier leurs résultats, devenant à son tour un producteur. Le fait de pouvoir ainsi réinjecter des données dans l'architecture relativise la notion de « donnée primaire » ;

- Un acteur peut consommer des données à partir de différents centres de stockage, les agréger, les lier et mettre ces « données enrichies » à disposition en tant que centre de stockage lui-même, à l'image d'un méta-moteur de recherche sur le web ou d'un comparateur de prix. Certains acteurs peuvent même facturer cette valeur ajoutée à leurs clients.

Dans le cas des méga-données (big data), apparaît un problème supplémentaire lié au caractère transitoire des données. Au vu des quantités de données traitées, il n'est techniquement pas possible de les stocker indéfiniment. Elles sont donc purgées régulièrement. Pour autant, les inférences construites à partir de ces données d'entrée vont persister et seront à leur tour intégrées aux données qui seront traitées par la suite. Elles pourront ainsi, par effet de cascade, influencer de futures déductions. L'intérêt des méga-données est double : travailler sur de gros volumes de données et disposer de données récentes. Néanmoins, il est du coup souvent impossible de retracer le processus de calcul jusqu'à son origine, ni même de retrouver les données « primaires ». Au-delà du fait qu'une erreur puisse perdurer bien après la disparition de la cause (calcul d'une donnée erronée par exemple), c'est surtout sous l'angle du rattachement des données que cette purge peut devenir problématique, en cassant potentiellement les liens entre les données primaires et les données enrichies.

D’un point de vue sécurité en vue de tracer l’utilisation des données, les technologies actuelles ne sont pas adaptées à notre problématique. D’un côté, les entrepôts de données et leurs politiques de contrôle d’accès ne permettent pas de garantir qu’un acteur ayant eu connaissance de certaines données ne les réplique pour les diffuser à son tour. Les solutions de gestion des droits numériques (ou DRM pour *Digital Rights Management*) ne sont pas plus adaptées car, pour garantir le « bon usage » des données, celles-ci nécessitent des outils dédiés pour pouvoir décrypter les contenus des données en fonction des droits octroyés aux acteurs. Une telle contrainte sur l’architecture est inévitablement un frein au partage d’information, comme les DRM l’ont été dans le domaine du multimédia. De telles solutions à base de DRM ne pourront donc cibler que des domaines d’application bien particuliers. D’autres techniques, pour l’instant moins répandues, telles que l’anonymisation, le chiffrement homomorphe[7] ou le tatouage numérique[8] pourraient apporter des solutions, mais uniquement sur des cas bien précis.

Dans un tel écosystème numérique, comment partager les fruits résultant de l’exploitation de données fournies gratuitement ? Quels sont les droits sur les données aux différents stades du traitement ?

II. Droits sur les données numériques

La question de la titularité et de l'étendue des droits sur les données provient du fait que, dans l'écosystème numérique, la valeur des données entrantes (données primaires) est différente de la valeur des données sortantes (données enrichies ou contextualisées). En effet, la donnée brute n'a pas de valeur intrinsèque. La valeur est essentiellement créée par la nature des traitements qui lui sont appliqués[9].

Plusieurs droits peuvent être mobilisés selon que l’on envisage ou non de patrimonialiser les données[10]. On peut en effet faire appel au droit commun de la propriété, au droit de la propriété intellectuelle ou à la notion de biens communs.

---

7 https://fr.wikipedia.org/wiki/Chiffrement_homomorphe

8 https://fr.wikipedia.org/wiki/Tatouage_numérique

9 H. Isaac, Valeur des données à l'ère des data-driven business models : http://chaireieso.fondation.dauphine.fr/blog/detail-dun-billet/article/conference-du-cercle-geopolitique-lamerique-latine-un-concept-de-stabilite-2/

10 Cette question est particulièrement importante en ce qui concerne les données personnelles qui constituent l’essentiel de la valeur pour les entreprises qui les traitent

2.1 Droit de propriété

Depuis longtemps, le statut juridique de la donnée et de l'information, si tant est que ces termes soient juridiquement équivalents, est discuté[11].

En principe, la collecte et la diffusion des données sont libres, sous réserve du droit des personnes. Les données personnelles relèvent de l'activité d'une personne déterminée, identifiable ou identifiée. Ces données constituent le matériau de base permettant notamment d'établir des profils comportementaux qui peuvent ensuite être exploités par les opérateurs du marché. La reconnaissance d'un droit de propriété sur les données peut tout à fait se concevoir si les divers éléments du droit de propriété sont réunis.

La propriété peut s'acquérir par convention (vente, donation, échange…) ou au terme d'une situation de fait (occupation, possession…). La possession est le pouvoir de fait exercé sur une chose avec l'intention de s'en affirmer le maître[12]. Les données peuvent être considérées comme un meuble incorporel sur lequel la possession a vocation à s'exercer dès lors que sont présents le corpus et l'animus constituant la possession.

Le corpus traduit une emprise matérielle de l'homme sur la chose, au travers de différents actes. Ces actes (utilisation, contrôle) peuvent s'exercer sur les données, au travers notamment de la volonté de les garder secrètes ou de les divulguer.

L'animus exprime quant à lui la volonté de l'individu de se comporter comme le maître de la chose. En matière de données, cette volonté peut notamment se traduire par l'interdiction faite à autrui d'en faire usage.

Lorsque le corpus et l'animus s'exerce sur les données, le possesseur de celles-ci peut être considéré comme leur propriétaire, en application de l'article 2276 du code civil[13]. Si l'application de cet article aux meubles incorporels est loin de faire l'unanimité parmi les juristes, elle ne peut cependant être écartée[14]. C'est « l'emprise possessoire conférée par le secret ou le maintien dans l'intimité de la personne - le contrôle de la donnée par son émetteur initial - qui permet de déduire qu'un droit de propriété s'exerce »[15].

---

11 P. Catala, Ebauche d'une théorie juridique de l'information, Rev. Droit prospectif 1983, n° 1, p. 185 ; P. Catala, La propriété de l'information, Mélanges Raynaud, Dalloz Sirey 1985, p. 97 ; E. Daragon, Essai sur le statut juridique de l'information, D. 1998, chr. p. 64 ; J-C Galloux, Ebauche d'une définition juridique de l'information, D. 1994, chr. p. 229 ; R. Hilty, La privatisation de l'information par la propriété intellectuelle : problèmes et perspectives, Rev. Intern. Droit éco. ; J. Passa, La propriété de l'information : un malentendu ?, Droit et patrimoine, mars 2001, n° 91, p. 65

12 F. Terre et P. Simler, Droit des biens, 9ème éd., Dalloz 2014, n° 138

13 Article 2276 c.civ. al. 1er : « En fait de meubles, la possession vaut titre »

14 V. notamment W. Dross, Droit des biens, 3ème éd., LGDJ 2017 ; A. Pélissier, Possession et meubles incorporels, Dalloz 2001

15 N. Binctin, Créer une patrimonialité des données à droit constant in « Mes data sont à moi », Rapport Génération Libre, janvier 2018, disponible sur https://www.generationlibre.eu/wp-content/uploads/2018/01/2018-01-generationlibre-patrimonialite-des-donnees.pdf

2.2 Droit de propriété intellectuelle

Certains prônent la création d'un nouveau droit de propriété intellectuelle qui protègerait les données personnelles contre l'exploitation qi en est faite par les plateformes sur Internet.

La propriété intellectuelle désigne un ensemble de droits sur les créations intellectuelles, pouvant relever soit de la propriété littéraire et artistique (droit d'auteur, droits voisins), soit de la propriété industrielle (brevets, marques, dessins et modèles…). Ces droits confèrent une exclusivité d'exploitation à leur titulaire mais ne peuvent trouver à s'appliquer que tant qu'il existe une création intellectuelle. Or, les données brutes, qu'elles soient personnelles ou non, ne constituent pas une création au sens du droit de la propriété intellectuelle. Des données générées automatiquement par des appareils connectés ne sauraient être considérées comme des créations.

Certaines données cependant peuvent faire l'objet d'une protection par le droit de la propriété intellectuelle. C'est notamment le cas des bases de données. L'article L 341-1 CPI dispose notamment que « le producteur d'une base de données, entendu comme la personne qui prend l'initiative et le risque des investissements correspondants, bénéficie d'une protection du contenu de la base lorsque la constitution, la vérification ou la présentation de celui-ci atteste d'un investissement financier, matériel ou humain substantiel ». La protection est toutefois subordonnée à un traitement des données et non à la simple compilation de celles-ci.

Les données brutes ne pourraient dès lors faire l'objet d'aucun droit de propriété intellectuelle.

2.3 Biens communs

Les données pourraient également être considérées comme des biens communs, ce que défendent certaines personnes. Les biens communs se caractérisent par deux éléments : la non exclusion et la rivalité[16]. La non exclusion signifie que l'on ne peut exclure personne de l'usage du bien. La rivalité signifie que la consommation par un individu diminue la quantité de bien disponible pour les autres individus car la ressource est consommée au moins en partie.

---

16 **B.** Boidin**, D.** Hiez **et S.** Rousseau, « Biens communs, biens publics mondiaux et propriété. Introduction au dossier. », *Développement durable et territoires* [En ligne], Dossier 10 | 2008, mis en ligne le 07 mars 2008 ; M. Cornu, F. Orsi et J. Rochfeld, Dictionnaire des biens communs, Quadrige PUF 2017

Lors des discussions de la loi pour une République numérique[17], certains députés avaient déposé un amendement proposant de considérer les données personnelles comme un bien commun[18].

Or, les données ne répondent pas aux critères de qualification des biens communs. Elles ne sont pas rivales puisque leur usage par un individu n'empêche pas ce même usage ou un usage différent par quelqu'un d'autre.

Elles peuvent également être exclusives. Tel est le cas notamment des données personnelles puisque la personne concernée peut en interdire l'usage à d'autres individus.

Les données numériques sont multiples et le droit peine encore à élaborer un régime uniforme face à cette multiplicité.

17 LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, JO 8 octobre 2016

18 L'amendement porté par les députés PS Delphine Batho et Laurent Grandguillaume et finalement rejeté, était ainsi libellé : « Les données à caractère personnel, lorsqu'elles forment un réseau indivisible de données liées qui concernent plusieurs personnes physiques, constituent un bien commun qui n'appartient à personne et dont l'usage est commun à tous, dont la protection et l'utilisation sont régies par la présente loi. »

LES AVOCATS ET LA REVOLUTION NUMERIQUE

Thierry Wickers
t.wickers@exeme-avocats.com

La révolution numérique n'épargne pas les avocats, qui sont concernés à la fois en tant qu'individus, mais aussi en tant que professionnels.

Les technologies de l'information vont profondément transformer l'activité des avocats, qui vont bientôt pouvoir utiliser l'intelligence artificielle ou la justice prédictive.

Le marché du droit lui-même est susceptible de connaître d'importants développements, car le modèle actuel ne donne pas, contrairement à ce que peuvent croire les avocats, entièrement satisfaction.

Pour le moment, c'est néanmoins des acteurs nouveaux, issus de la Legaltech, qui ont mieux su que les avocats s'implanter sur le marché, inventant des services nouveaux.

Ils ont compris, avant les avocats, que désormais, c'était d'abord sur le réseau que l'on recherchait la solution à ces problèmes. En répondant aux questions des internautes, ils peuvent contrôler l'accès aux avocats. Il faut donc que les avocats, à leur tour, s'emparent de ces nouveaux outils, pour modifier leur offre et répondent aux demandes du vaste marché latent du droit.

Il n'y a pas de raison qu'ils n'y parviennent pas, mais ils ne doivent plus perdre de temps.

Le paradoxe de l'accès au droit

Les avocats ont développé un modèle très stable de fourniture des services juridiques. Il est à peu près le même partout, ce qui explique d'ailleurs qu'en dépit d'environnements législatifs pouvant être différents, la profession d'avocat présente des traits communs, notamment en ce qui concerne les principes éthiques (indépendance, secret professionnel, qualification professionnelle, réglementation des conflits d'intérêts), et qu'elle dispose normalement d'activités réservées.

L'offre des avocats présente en général quatre caractéristiques majeures :

- Les avocats proposent à leurs clients des solutions individuelles, fabriquées « sur-mesure ». Il s'agit de fournir une réponse personnalisée, à des problèmes individualisés, en général dans le cadre d'un tête-à-tête entre l'avocat et son client

- Les avocats ne proposent pas une « gamme » de produits, en procédant à des arbitrages prix/qualité. Au sein d'un même cabinet, toutes les prestations juridiques sont produites selon les mêmes méthodes, pour arriver à des résultats identiques. La qualité des prestations est constante, elle est assurée à tous les clients (elle ne dépend pas de la situation de fortune du client) de manière uniforme.

- Les avocats sont à la recherche de la plus haute qualité possible de prestation. Ils cherchent à fournir à leurs clients, en toutes circonstances, le meilleur service dont ils sont capables.

- Les avocats produisent de manière artisanale des prestations individualisées, en recherchant la plus haute qualité possible. Donc les services juridiques sont d'un prix très élevé.

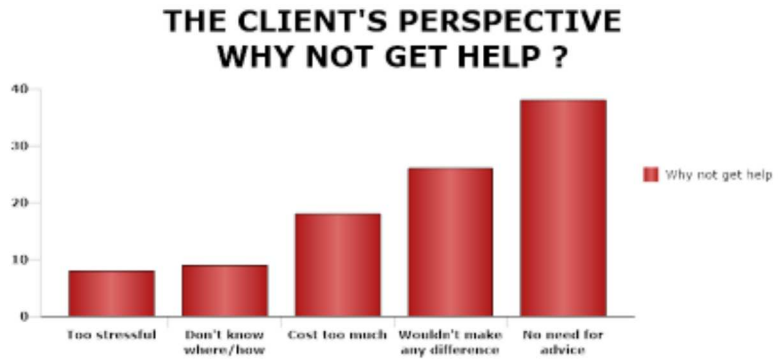
Quand il s'agit d'apprécier la pertinence de ce système et de vérifier sa capacité à garantir l'accès au droit, on s'est longtemps appuyé sur des enquêtes quantitatives, mesurant l'activité du système judiciaire, le nombre d'affaires traitées, la vitesse de déroulement des procédures…

Des études récentes ont adopté un point de vue différent, essayant de comprendre comment les personnes confrontées à un problème juridique réagissaient et si l'offre de droit assurait efficacement sa prise en charge.

Ces études, dont les premières remontent à la fin des années 80 et ont été réalisées aux USA ont montré l'existence d'une « justice gap », une part grandissante des particuliers ou des entreprises ne recourant pas aux services des avocats.

9% of respondents apprehend the legal nature of the problem they face

9% des personnes interrogées appréhendent la nature juridique du problème auquel elles sont confrontées



Il existe de multiples causes à cette situation, mais ce qui est frappant c'est de constater que les principales raisons mises en avant par les personnes interrogées révèlent une totale méconnaissance de l'utilité des avocats pour régler les problèmes juridiques et plus encore, une véritable incapacité à en appréhender la nature.

Le paradoxe de l'accès au droit, tel qu'il est organisé par les avocats, est que leur offre ne concerne que les personnes qui ont été capables de diagnostiquer la nature juridique de leur problème, de comprendre qu'il s'agissait d'un problème juridique et d'identifier un professionnel du droit susceptible de les prendre en charge.

Il s'agit en définitive d'une faible minorité de personnes, de sorte qu'il existerait une importante demande latente de droit, qui ne demanderait qu'à se manifester.

L'existence de ce marché latent est donc une formidable opportunité pour les avocats, puisque s'ils parviennent à capter cette demande, ils pourront considérablement améliorer leur situation économique et que leur profession pourra connaître un important développement.

L’apparition de la « legaltech »

La théorie de l’innovation disruptive apprend que lorsqu’un marché est mal desservi par les entreprises en place, de nouveaux acteurs peuvent être tentés d’entrer sur ce marché.

Ces nouveaux acteurs s’intéressent d’abord à la partie la moins solvable du marché, celle que négligent, pour cette raison les entreprises en place.

Ils proposent des services plus simples, moins coûteux, plus faciles d’accès, plus lisibles.

Depuis quelques années, d’abord sur le marché américain, mais maintenant également en Europe, apparaissent des non-avocats, qui proposent des services juridiques sur Internet.

Si on tente une typologie des sites de la Legaltech, on peut essentiellement identifier :

- des fournisseurs d’informations juridiques. Ils font preuve d’une manifeste volonté de rendre le droit beaucoup plus accessible, en mettant en ligne une information conçue dans le langage de ceux-là mêmes qui expriment leurs besoins sur les réseaux.

- les services d’assistance judiciaire. Ils proposent d’accompagner leurs clients dans une démarche contentieuse, qu’il s’agisse de la formulation d’une réclamation (rédaction d’une mise en demeure) ou de l’engagement d’une procédure. L’offre peut s’étendre à la saisine de la juridiction ou à l’accomplissement des formalités liées au respect du principe du contradictoire. Une variante, également présente, se propose d’agréger les personnes concernées par le même problème juridique, pour leur permettre d’agir de concert en mutualisant les coûts.

- les sites de rédaction d’actes juridiques en ligne. Ils commencent à utiliser « l’intelligence artificielle ». Des questions ciblées permettent la rédaction automatisée d’actes relativement simples, mais comprenant néanmoins une certaine dose de personnalisation.

- les plateformes de règlement des litiges en ligne. Elles proposent une alternative à l’action judiciaire et elles se passent parfois de toute intervention humaine.

- les sites de justice prédictive. Ils exploitent les possibilités du « big data », pour prédire les chances de succès d’une procédure ou chiffrer les résultats d’une action judiciaire.

- les services de mise en relation. Même les sites dont ce n’est pas la vocation première, proposent souvent d’assurer la mise en relation avec un avocat. Mais à la différence d’un simple annuaire, ils procèdent à une sélection préalable des professionnels. Parfois, ils permettent aussi au client de porter une appréciation sur la prestation juridique fournie par l’avocat recommandé.

- les services de financement du procès. Il peut s’agir de financer un litige d’intérêt général, comme un litige environnemental, ou des litiges privés.

Le nouveau parcours des clients

Ce que les entreprises de la legaltech ont compris avant les avocats, c'est que les habitudes des clients avaient changé et que désormais, le parcours d'un client potentiel commençait sur Internet.

Toute personne, confrontée à un problème de toute nature, a désormais pour premier réflexe de se connecter (de plus en plus via un smartphone) à Internet, pour chercher une réponse ou une aide.

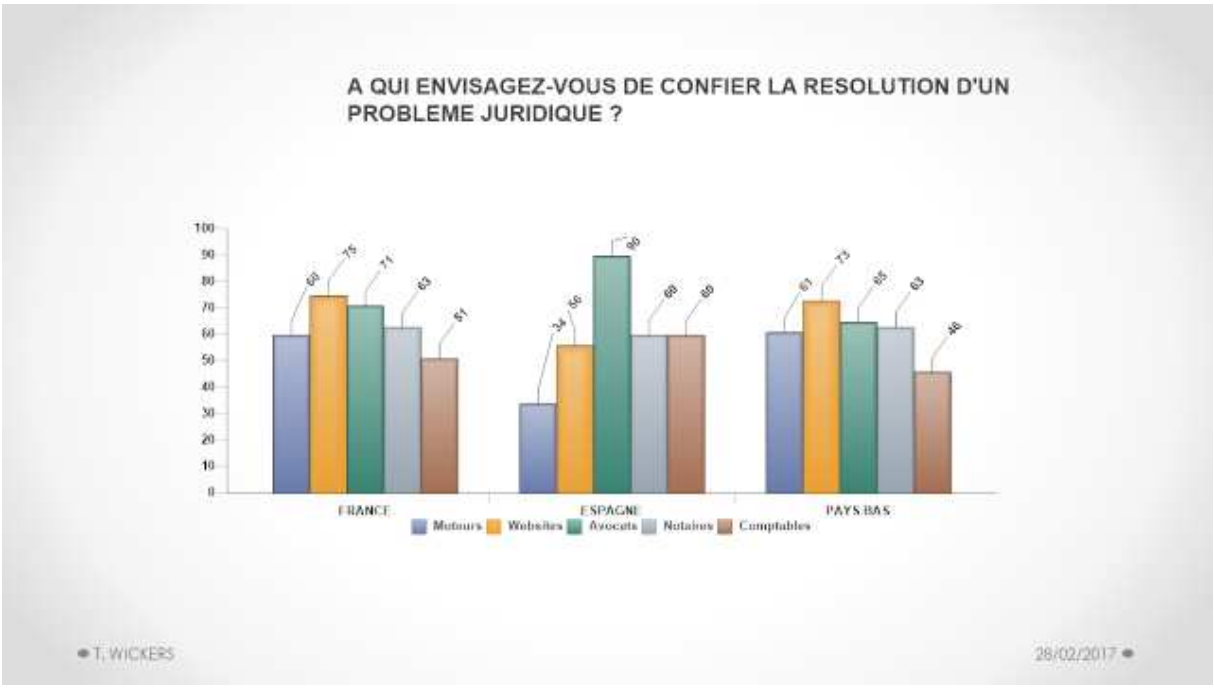
Cette tendance ne fera que s'accentuer, la génération des « millenials », comme les générations suivantes, sont nées avec Internet. C'est avant tout vers le réseau, par le biais d'un smartphone, qu'elles se tournent.

À la différence de ce qui se passait jusqu'à présent, il n'existe plus potentiellement de question qui puisse rester sans réponse. En se déplaçant sur Internet, le « marché latent » peut trouver l'occasion de s'exprimer.

Pour le moment, ce sont les entreprises de la legaltech qui s'imposent sur ce nouvel espace.

Les conséquences pour les avocats peuvent être importantes.

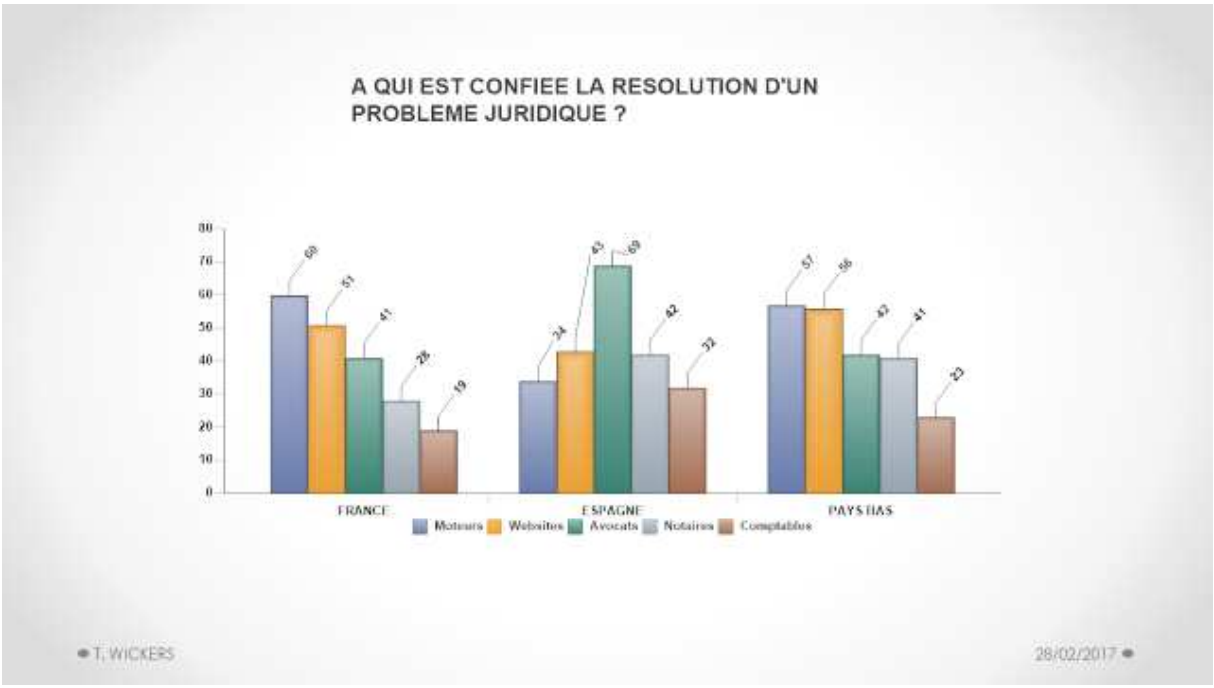
L'étude de marché réalisée par Rocket Lawyer dans les trois pays (France, Espagne, Pays-Bas) dans lesquels la société a décidé de proposer une offre juridique, apporte une réponse significative.



Certes, les personnes interrogées continuent de placer les avocats à la première place, lorsqu'on leur demande à qui ils confieraient un problème juridique, s'ils en rencontraient un. Mais (notamment en France ou au Pays-Bas), ils sont désormais nettement concurrencés par les moteurs de recherche ou les sites spécialisés, dans l'esprit des internautes.

Or il ne faut pas oublier qu'il y a deux ou trois ans à peine, ces services n'existaient pas. Ils se sont donc imposés à une vitesse stupéfiante et rien ne dit que l'évolution est achevée et qu'ils ne vont pas continuer à progresser.

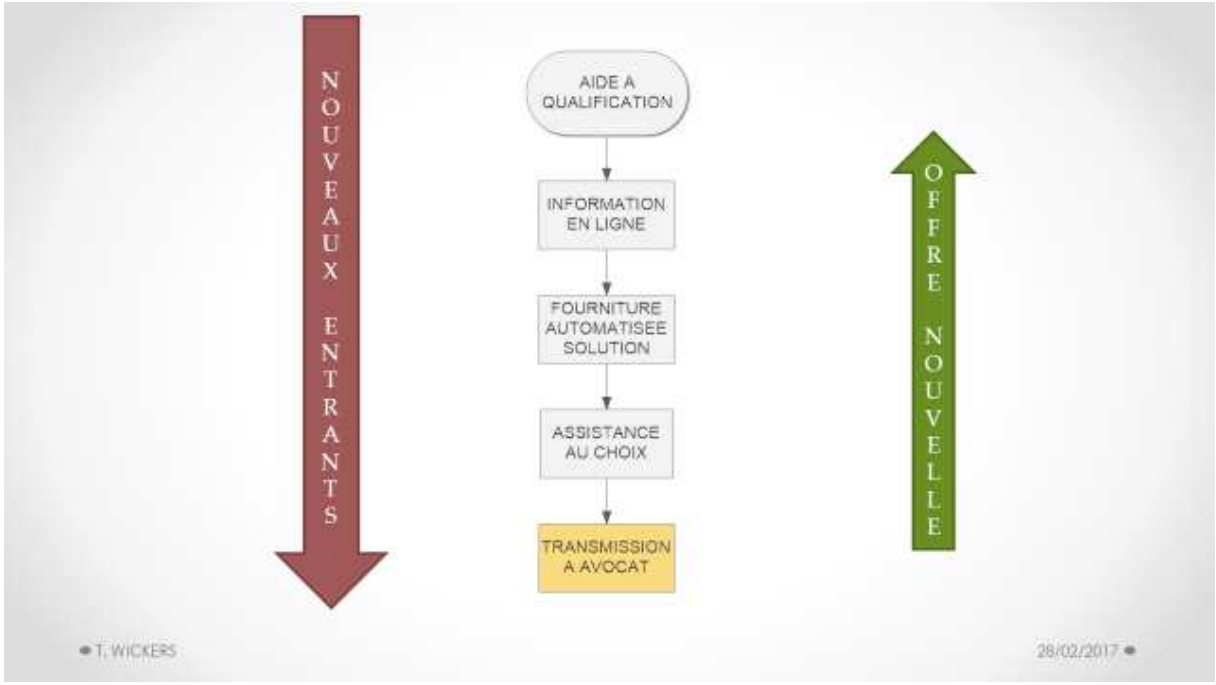
En outre, on voit que, lorsqu'on s'intéresse non plus aux simples intentions déclarées, mais au comportement effectif des personnes interrogées, ces derniers donnent déjà la priorité, dans deux des trois pays concernés, aux sites ou aux moteurs de recherche, sur les avocats !

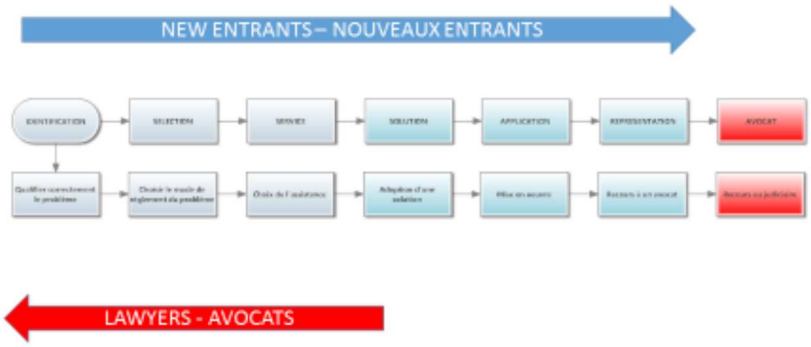


Celui qui, le premier, fournit à l'internaute à la recherche d'une réponse à son problème, les renseignements qui vont lui permettre de s'orienter, est susceptible par la suite, de devenir un prestataire de services et un prescripteur.

Naturellement, ces nouveaux développements ne feront pas disparaître les circuits classiques. Il restera toujours des clients qui préféreront effectuer personnellement le choix d'un avocat, sans s'en remettre à des sites pour les orienter.

D'un autre côté, ce que permettent les technologies de l'information, c'est l'extension du « marché du droit », qui pourrait donc s'accroître considérablement, également au profit des avocats.





Quelles stratégies pour les avocats

Les moyens et les technologies mises en œuvre par les acteurs de la legaltech sont également à la disposition des avocats.

D’ailleurs, bien souvent, on retrouve d’anciens avocats (ou même magistrats) à la tête des startups du droit.

Rien n’empêche les avocats d’imaginer des stratégies nouvelles, pour capter à leur profit, les tendances qui se dessinent.

Tous les avocats qui souhaitent profiter des développements nouveaux du marché du droit doivent aujourd’hui disposer d’une stratégie numérique. Celle-ci n’a pas de raison de rester l’apanage des sociétés de la legaltech.

Thème D

« Droit des données à caractère personnel »

La « de-anonymization » : une atteinte à la vie privée ? Quelle protection pour les utilisateurs de la nouvelle technologie

Convergences droit et numérique, 11,12, 13 septembre 2017, pôle juridique et judiciaire de Bordeaux.

Lamia El Bouchtioui, doctorante en droit public, CRDEI, Université de Bordeaux
lamia.el-bouchtioui@u-bordeaux.fr
Valeria Loscri, chercheuse, INRIA, Lille-Nord Europe
valeria.loscri@inria.fr

Actuellement, le numérique règne dans nos sociétés contemporaines et produit de plus en plus d'effets sur le fonctionnement des démocraties et sur leurs droits fondamentaux. Force est de constater que le droit est entièrement touché par le numérique, notamment les droits fondamentaux, comme le droit au respect de la vie privée ou le droit des données personnelles. Dans le cadre de la généralisation numérique deux thèses s'opposent. Des études et des rapports se sont multipliés ces dernières années pour, d'une part défendre l'avènement du numérique, et d'autres part, dénoncer l'atteinte à des libertés fondamentales.

La société du numérique s'inscrit dans un paradoxe. Il est vrai que le numérique apporte un accroissement des libertés individuelles mais, en même temps les libertés individuelles se trouvent menacées. En effet, une utilisation abusive peut se révéler contraire aux droits de l'homme et aux libertés fondamentales. Le développement du big data s'inscrit dans ce débat ventant le mérite de l'avènement du numérique et dénonçant l'atteinte aux libertés fondamentales. Il est certainement source d'importantes transformations économiques et sociales. « *La constitution et l'exploitation de ces masses de données par les entreprises afin d'améliorer leurs performances dans la production et la distribution de leurs produits et services sont riches de perspectives de croissance dans de multiples secteurs. Mais elles imposent dans le même temps de repenser les problèmes de privacy compte tenu des dangers qu'elles recèlent pour les utilisateurs* [1]». Ainsi, les données apparaissent comme le nouvel or noir de notre société.

La généralisation du numérique permet d'observer l'évolution des nouvelles technologies. De nos jours, les technologies permettent d'identifier un individu et, de presque tout savoir sur cet individu. Or, un citoyen dans une démocratie a droit à la protection de son identité et de son intimité : ce que l'on qualifie de *privacy*. Le risque de ces technologies est la collecte, l'utilisation des données dans

1 A. BENSAMOUN, C. ZOLYNSKI, Big data et privacy : comment concilier nouveaux modèles d'affaires et droits des utilisateurs ? , Petites Affiches, n°164, p 8, 18/08/2014.

un but qui sort du cadre légal. Pour ce faire, la « de-anonymization » est une méthode permettant d'identifier un individu sans son consentement. Il est certain que cette méthode porte atteinte à la vie privée, et au droit des données personnelles. Les aspects techniques de la « de-anonymization » (I) permettront de constater l'ampleur des risques d'atteinte à la vie privée. Du point de vue juridique la « de-anonymization » (II) reste un problème d'envergure pour lequel aucune réponse spécifique n'a été apportée.

I- Les aspects techniques de la « de-anonymization »

La « de-anonymization » est un processus qui doit nécessairement être mis en parallèle avec les grands débats sur « l'anonymisation ». Force est de constater que l'anonymat absolu semble être un mythe. En effet, en étudiant le processus de « de-anonymization » (A) et en observant le cas d'étude des smartphones (B) il sera mis en évidence que l'anonymat est loin d'être garanti.

A- Qu'est-ce que la « de-anonymization » ?

1- Définition

La « de-anonymization » est un terme anglais qui a pour signification : la technique de données inversées qui permet de ré-identifier des informations cryptées ou généralisées. Comme son nom l'indique la « de-anonymization » est le contraire de l'anonymisation.
Elle peut être identifiée en français comme la ré-identification. Un processus de recoupement de données anonymes, qui, peut conduire à retrouver l'identité d'une personne. De ce fait, les données « rendues anonymes » peuvent facilement être ré-identifiées sans avoir en possession les codes de sécurité du processus d'anonymisation.

2- Méthode

Pendant longtemps, on a largement cru que tant que les ensembles de données ont été "rendus anonymes", il n'existerait aucun risque concernant la vie privée d'une personne. En effet, si les ensembles de données ont été rendus anonymes, ils n'ont pas révélé l'identité d'individus connectés aux données. Malheureusement, la notion d'anonymisation parfaite a été exposée comme un mythe. Au cours des vingt dernières années, les chercheurs ont montré que les individus peuvent être identifiés dans beaucoup d'ensembles de données différents alors même que l'on pensait que ces

données étaient « rendues anonymes »[2].

À l'heure de la nouvelle technologie, du big data et du *quantified self*[3], la protection des données semble difficile. Le risque de ré-identification est avéré. De nombreuses affaires démontrent les failles de la protection actuelle des données à caractère personnel.

Grosso modo, il existe plusieurs techniques d'anonymisation[4]. Selon la (Commission Nationale de l'Informatique et des Libertés) CNIL, trois méthodes peuvent être utilisées : la pseudonymisation, le masquage et l'agrégation. N'étant pas l'objet de notre étude, les méthodes d'anonymisation ne seront pas détaillées. Malgré des techniques sophistiquées, la ré-identification n'est pas impossible. Force est de croire qu'aucune technique n'est infaillible.

La principale faille de l'ensemble des méthodes d'anonymisation tient aux données mêmes auxquelles elles sont appliquées. Pour ce faire, les liens établis entre elles sont parfois aussi identifiants que chacune de ces données prises isolément. Une étude américaine a montré que 97% des électeurs de la ville de Cambridge dans le Massachussets pouvaient être identifiés par le seul croisement de leur date de naissance et des neuf chiffres du code postal correspondant à leur adresse[5].

Mais encore, deux illustrations peuvent montrer les failles de l'anonymisation. Notamment, l'un des exemples le plus connu est l'affaire AOL concernant le big data. L'entreprise américaine AOL est fournisseur d'accès internet. En 2006, elle a publié en ligne une base de données rassemblant 20 millions de recherches effectuées sur son site par 650 000 utilisateurs. La base avait été anonymisée selon un procédé de pseudonymisation : chaque identifiant avait été remplacé par un nombre choisi aléatoirement. Toutefois, AOL a négligé l'historique de recherche d'un individu qui permet quant à lui à identifier un individu. Par recoupement, il est possible de rapprocher l'adresse des utilisateurs. Les choix de recherches peuvent aussi fournir des informations sur l'âge, la profession, les goûts d'une personne. Après la publication de cette base de données, le New-York Times a pu révéler l'identité d'une cliente de AOL : Thelma Arnold, 62 ans[6].

2 I. RUBINSTEIN, W.HARTZOG, Anonymization and Risk, Washington Law Review, Vol.91, N°2, 2016, NYU School of Law, Public Law Research Paper N°15-36.

3 Le quantified self, expression de l'internet des objets, correspond à la numérisation et à la quantification des activités humaines Ce mouvement, lancé en 2007, vient redéfinir la relation que la personne entretient avec son corps, sachant que celle-ci devient créatrice de données en s'autoévaluant, lui permettant ainsi de déterminer des comportements à adopter pour elle-même.

4 L'anonymisation peut être définie comme l'opération de suppression de l'ensemble des informations permettant d'identifier directement ou indirectement un individu, contenues dans un document ou une base de données.

5 L. SWEENEY, « Weaving technology and policy together to maintain confidentiality », Journal of Law, Medicine and Ethics, 25, p. 98-110.

6 M. BARBARO, T. ZELLER, A face Is Exposed for AOL Searcher N° 4417749, N.Y. TIMES, Aug. 9, 2006.

Il suffit de peu de données et dans la majorité des cas de données anodines pour permettre d'identifier une personne parmi d'autres, et ce en dépit de son anonymisation. Cette observation a un impact sérieux sur les données de santé. Le rapport de Pierre-Louis et André Loth sur la gouvernance et l'utilisation des données de santé, indique qu'il est possible d'identifier les patients. Ainsi, 89% des patients ayant connu un séjour à l'hôpital en 2008 sont identifiables avec pour seules informations : l'hôpital d'accueil, le code postal du domicile, le mois et l'année de naissance, le sexe, le mois de sortie et la durée du séjour. Des données anodines en apparence, mais, qui peuvent permettre d'identifier une personne et d'avoir accès à leur dossier médical.

« La ré-identification par croisement d'informations qui subsistent dans la base, après anonymisation, se trouve d'ailleurs grandement facilitée par le recours à d'autres jeux de données publiés, qui peuvent rendre identifiant des liens entre plusieurs données qui jusqu'alors ne paraissaient pas permettre de caractériser une personne »[7]. Dans l'affaire Netflix, ce procédé a permis l'identification de nombreux clients. L'entreprise Netflix offre un service de diffusion de film en ligne et permet à ses utilisateurs de noter et recommander les films. Les appréciations permettent à l'entreprise de cerner les goûts de ses clients et de leurs proposer des films susceptibles de plaire. Pour affiner ses programmes d'analyse de préférences de ses utilisateurs, Netflix met en jeu un prix d'un million de dollars[8]. Elle publie en ligne les recommandations de 500 000 utilisateurs, afin que des programmes indépendants développent des algorithmes plus performants que ceux utilisés par l'entreprise pour proposer à ses clients des films conformes à leurs goûts. Les données ont été anonymisées : les identifiants anonymisés ainsi que les autres données annexes. Arvind Narayanan et Vitaly Shmatikov sont parvenus à ré-identifier plusieurs profils d'utilisateurs[9]. Les informaticiens ont pu avec l'information donnée par le croisement entre l'appréciation portée sur trois films, et la date à laquelle ils ont été loués retrouver l'auteur de ces appréciations.

B- Un cas concret : le smartphone et l'interface WIFI

1- Objectifs

Grace à notre smartphone, aujourd'hui nous pouvons bien profiter de nouveaux services très

7 G. GORCE, F. PILLET, La protection des données personnelles dans l'open data : une exigence et une opportunité, Rapport d'information n°469 (2013-2014), 16 avril 2014.

8 The Netflix Price Rules, Netflix (2006), voir http://www.netflixprize.com/assets/rules.pdf [https://perma.cc/8XUU-G4GK].

9 A. NARAYANAN, V. SHMATIKOV, « Robust de-anonymisation of large sparse datasets », Proceedings of the 2008 IEEE Symposium on Security and Privacy, p. 11-125, cité par Kieron O'Hara, préc., p. 40-42.

innovants qui rentrent dans différentes catégories comme *l'entertainment*, *l'online banking*, etc,... Certainement, ces dispositifs ont beaucoup changé la façon d'appréhender et de voir la technologie. En effet, la technologie va faciliter la vie quotidienne, en permettant de se repérer plus facilement dans des lieux inconnus ou de trouver des points d'intérêts, simplement en se connectant au WIFI. De ce fait, les objectifs les plus importants du WIFI dans les smartphones peuvent être synthétisés comme la possibilité de générer des services innovants qui facilitent la vie de l'utilisateur.

2- Méthode

Pour pouvoir permettre aux utilisateurs d'être connectés « anytime» et « everywhere», il y a des protocoles qui réalisent la connexion à internet d'une façon presque transparente à l'utilisateur. Ça signifie qu'il y a des actions automatiques qui sont exécutées. Par exemple, la génération de paquets « probes requests» est une action automatique permettant d'activer une connexion. En particulier la « méthode» standard se base sur le protocole de communication 802.11 et consiste en plusieurs étapes qui, sont exécutées automatiquement, si l'interface wireless est active. Les dispositifs envoient automatiquement des requêtes sous forme de « WIFI probe request» pour savoir s'il y a des réseaux disponibles.

Le dispositif est identifié d'une façon unique avec un identifient MAC (l'adresse MAC du dispositif). Les paquets « probes» peuvent être de deux types : *broadcast* (diffusion) ou direct. Les paquets *broadcast* cherchent pour tous les points d'accès (Acces Points - AP) alors que les directs contiennent une adresse spécifique d'un AP. Les systèmes opératifs (Operating Systems - OS) des portables sauvegardent les réseaux auxquels l'utilisateur a été connectés. En particulier cette information est sauvegardée comme Preferred Network List (PNL). Les PNL ont été conçus pour se connecter plus vite et épargner la consommation d'énergies. Normalement, les dispositifs utilisent les PNL pour essayer de se connecter à un des réseaux disponibles. La tentative de connexion est toujours réalisée à travers les paquets probes qui sont envoyés en « clair» (c'est à dire sans être cryptés). Ça signifie qu'il est extrêmement facile de capter les données et d'acquérir des informations sur les dispositifs qui se connectent.

- « La de-anonymization » des évènements

Le travail auquel on fait référence s'est fondé sur les données du datasets du WIFI qui sont accessibles par tout le monde[10]. En particulier, ces données concernent les élections nationales de 2013 qui se sont déroulées en Italie, pendant lesquels il y a eu deux importantes rencontres. La

10 M. V. Barbera, A. Epasto, A. Mei, V. C. Perta, and J. Stefa, "Signals from the crowd: uncovering social relationships through smartphone probes,"in Proc. of ACM IMC, 2013.

première rencontre était la clôture de la campagne électorale du Mouvement Cinq Étoiles (M5S). La seconde rencontre est intervenue après les élections et concernait le PDL (dont le leader était Silvio Berlusconi). Les auteurs ont montré qu'à partir des données que les smartphones utilisent pour effectuer la connexion à internet, et, en faisant des « suppositions» on peut bien arriver à élargir la connaissance et découvrir des informations. Il s'agit de la première étape de la « de-anonymisation ».

Une première intuition/supposition des auteurs est de dire que les personnes se connectent aux réseaux des lieux où ils passent beaucoup de temps (maison, boulot, école, etc.). Seconde supposition, en principe, tous ces lieux se trouvent dans la même région ou ville. Donc, pour « de-anonymiser » on peut utiliser les coordonnées GPS des AP qui se trouvent dans la PNL. Les paquets de probes contiennent les identifiants du réseau (SSID), mais on peut enrichir cette information avec les coordonnées géographiques en utilisant par exemple le Wigle.net (Wireless Geographic Logging Engine). Ainsi cette information peut être utilisée pour la de-anonymisation. Le résultat de cette « de-anonymization », est de pouvoir définir le lieu de vie de chaque participant. Grâce à la liste PNL, ainsi que les données géographiques, il a pu être défini le lieu de provenance des participants, sans savoir accès aux listes détenues par les autorités locales.

3- Aller plus loin

Les adresses Wifi enregistrées sur un smartphone peuvent révéler des choses bien plus qu'on imagine. Certaines adresses contiennent le nom des utilisateurs, par exemple, par défaut lors de la création d’un hotspot personnel avec un terminal Apple. D’autres adresses Wifi contiennent des informations sur les adresses postales, sur les noms des entreprises, sur les aéroports fréquentés, les gares, les hôtels, les restaurants...

Finalement, les adresses dévoilent tout le quotidien de l'utilisateur. Dans d'autres recherche similaires, les informaticiens ont voulu savoir ce qu’on pouvait en déduire notamment, en termes de liens entre individus. L'hypothèse émise était que l’on devait pouvoir mesurer un indice de relation en tirant parti des noms de réseaux les plus spécifiques comme ceux des box à domicile, exploitées par un nombre limité d’utilisateurs, en général la famille et des amis de passage, ou les noms de réseaux professionnels partagés entre collègues.

La méthode utilisée a permis de déterminer des liens entre les individus. Finalement, une simple adresse Wifi permet de déterminer la nature du lien que peut entretenir plusieurs individus. Les

données en apparence anodine, permettent d'empiéter non seulement sur la vie privée de l'utilisateur du smartphone mais aussi sur les personnes environnantes. L'identité des personnes est soulevée mettant de côté l'anonymat. Qu'en est-il des aspects juridiques en la matière.

II- Les aspects juridiques de la « de-anonymization »

Il n'existe à ce jour aucune définition juridique de la « de-anonymization ». Toutefois, l'intérêt de ce travail de recherche n'est certainement pas de trouver une définition à une méthode qui revient à mettre en cause la législation sur les données personnelles. La « de-anonymization » permet d'observer les failles de la loi informatique et libertés, qui semble dépassée par l'usage actuel du numérique. À l'heure du tous connectés, il est difficile de pouvoir garder l'anonymat absolu et une protection effective de notre vie privée. La réglementation est confrontée aux transformations des usages numériques des données personnelles (A). Ainsi, il est nécessaire de rechercher une meilleure protection des données face à ces transformations (B).

A- Une réglementation confrontée aux transformations des usages numériques des données personnelles

1- La frontière poreuse entre la sphère publique et privée

Lors de la Commission du 26 novembre 2014, la présidente de la CNIL, Mme Isabelle Falque-Pierrotin a indiqué que

> « *notre époque se caractérise par une imprégnation des données personnelles dans toutes les activités publiques, professionnelles ou privées. L'individu est de plus en plus pris dans un maillage extrêmement fin d'informations personnelles relayées par des objets de plus en plus communicants : téléphone portable, bracelets électroniques divers, dispositifs électriques, équipements de vidéosurveillance, etc. Cette " datification " (...) du monde (...) illustre (...) l'entrée dans un numérique ambiant* » dans lequel « *la dichotomie qui existait encore il y a quelques années entre les univers physique et virtuel (...) a disparu* ». Cette imprégnation « *change le rapport qui existait entre vie privée et données personnelles. Jusqu'à une période récente, les protections de ces deux sphères se superposaient. Sous l'effet des nouveaux comportements et usages, la frontière entre la vie privée et la vie publique commence à se détendre pour donner naissance à une zone un peu grise dans laquelle les personnes veulent exposer leur vie privée et se servent des données personnelles pour avoir une vie publique* » et, tout en demandant une protection, « *recherchent avant tout une maîtrise* »[11].

11 Assemblée Nationale, Rapport Numérique et libertés : un nouvel âge démocratique n° 3119 déposé le 9 octobre 2015. p103

Pour certains chercheurs les profonds bouleversements que connaît notre société amène à une certaine forme de paradoxe : la privacy paradox. Les individus exprimeraient une inquiétude vis à vis des nouvelles technologies, par exemple les réseaux sociaux, face aux risques liés aux données personnelles, dans le même temps, divulgueraient avec légèreté des données personnelles, même sensibles, sans aucune garantie ou contrôle. Une incohérence se dessinerait autour de la question des données personnelles. L'individu s'expose davantage et divulgue des données personnelles parfois sensibles. À titre d'exemple, à travers les réseaux sociaux, il est difficile d'observer la frontière entre la sphère public et privée. Ou bien, il semblerait que les personnes cherchent à maîtriser la gestion de leurs données personnelles.

2- Le nouveau paysage pour les données personnelles

- La conception évolutive des données personnelles

La définition de la notion données personnelles s'est révélée être une tâche complexe, voire impossible. En effet, pouvoir donner une définition simple et efficace semble être un travail laborieux. En réalité, la notion de données personnelles est mobile, évolutive et surtout subjective.
À titre d'exemple, Christine Balagué, titulaire de la Chaire réseaux sociaux à l'Institut Mines-Telecom, considère que les données personnelles intéressantes en matière de marketing changent constamment de nature selon le contexte. Il peut s'agir de données de la vie quotidienne dévoilées sur les réseaux sociaux et non plus uniquement de données de qualification tels que le genre, l'âge, l'adresse. Aujourd'hui se pose la question de nouveaux types de données engendrées par les objets connectés et l'Internet des objets. Des données insignifiantes mais susceptibles de contribuer à un profilage précis des individus qui permettra de connaître leurs croyances religieuses, leur orientation sexuelle, etc,...

Finalement à côté des données classiques, s'ajoutent de nouvelles données en apparence insignifiantes. Par exemple les adresses IP sont des identités anonymes mais au final définissent bien une personne. Le débat sur le caractère identifiant de l'adresse IP paraît être tranché. La caractéristique des adresses IP d'un particulier n'est plus dynamique mais fixe dans la majeure partie des cas. La Cour de Justice de l'Union européenne a estimé en 2011 que l'adresse IP était une donnée personnelle dans l'affaire *Scarlet Extended vs SABAM*. De nos jours, cette décision peut être étendue à d'autres numéros identifiants uniques comme les adresses MAC ou l'UDID des iPhones qui permettent de suivre des utilisateurs. Le cas du WIFI des smartphones permet de mettre en avant de nouvelles données qui semblent insignifiantes mais ont un potentiel de profilage non

négligeable. Pouvoir avoir la liste des adresses WIFI enregistrées sur un smartphone retracera le parcours d'un individu, les lieux qu'il fréquente, les personnes qu'il fréquente, et d'observer les relations entre les personnes.

Avec l'évolution du numérique et des nouvelles technologies la notion de données personnelles risque d'évoluer au fur et à mesure rendant impossible la mission de la définir. Le nouveau paysage des données personnelles ne s'arrête pas à l'unique définition des termes mais aussi à la conception d'anonymat.

- Vers la fin de l'anonymat ?

La « de-anonymization » pose la question de la fin de l'anonymat[12]. Paul Ohm a indiqué dans un article que la confiance dans le pouvoir protecteur des techniques d'anonymisation a été surévaluée[13]. Il est ainsi légitime d'envisager la fin de l'anonymat.

Au delà des techniques de « de-anonymization », la tendance actuelle est d'inciter les utilisateurs à utiliser leur véritable identité. Facebook et Google encouragent tous leurs utilisateurs à n'utiliser qu'un seul identifiant qui se rapproche de leur véritable identité. Pour ce faire, Facebook utilise la méthode du *crowdsourcing,* en proposant à ses utilisateurs d'identifier au sein de leur contact ceux qui recourent à des pseudonymes. Yann Leroux, docteur en psychologie, apporte quelques conclusions sur l'anonymat en indiquant qu'à l'avenir le droit à l'anonymat sera de plus en plus contesté. Une crainte qui n'est pas toujours partagée par tout le monde. En tous les cas, il est indéniable que le paysage des données personnelles a changé, ce qui rend la législation actuelle difficilement applicable.

3- Une législation actuelle inadaptée

La législation française en matière de données personnelles a su donner des principes généraux et une définition large et variable des données à caractère personnel. Cette caractéristique peut être vu comme un avantage, en effet, une donnée personnelle actuelle peut entrer dans le champ d'application de la loi informatique et libertés de 1978. Ce point concerne l'état d'esprit de la loi ainsi que la définition de données personnelles. Il est certain que son article premier reste toujours pertinent et dispose que « *L'informatique doit être au service de chaque citoyen. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques* ».

12 Voir la méthode p.2
13 P. OHM, « Broken promises of privacy : responding to the surprising failure of anonymization », 2010.

L'article 2 quant à lui apporte une définition générale d'une donnée personnelle permettant ainsi une grande marge de manœuvre quant à la qualification d'une donnée de « personnelle » ou non personnelle. Il dispose que « *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 ; Constitue un traitement de données à caractère personnel, toute opération ou tout ensemble d'opérations portant sur de telles données, quel que soit le procédé utilisé, et notamment la collecte, l'enregistrement, l'organisation, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, ainsi que le verrouillage, l'effacement ou la destruction ; Constitue un fichier de données à caractère personnel, tout ensemble structuré et stable de données à caractère personnel accessibles selon des critères déterminés. »*.

Par la définition large qu'elle prévoit, la loi a pu s'adapter à l'évolution de l'informatique ou être adaptée par la conception de la CNIL qui a su adapter la loi au fur et à mesure des évolutions successives du numérique : carte mémoire, l'e-santé, géolocalisation etc,...

Même si la loi a des points positifs, il est évident qu'elle n'arrive en aucun cas à faire face à ces nouvelles transformations des usages numériques. On assiste à une diversification de l'origine et de la nature des données personnelles collectées. Les outils actuels qu'un individu utilise, téléphone portable, objets connectés, caméras, vidéoprotection, supposent des interconnexions et souvent des géolocalisation indispensables à leur fonctionnement. En les utilisants, la conséquence première est de laisser des traces numériques, donc des données personnelles sans que l'utilisateur s'en rende compte ou en mesure les conséquences. Le cas d'étude du WIFI des smartphones est un bel exemple d'utilisation d'un outil qui transmet des données personnelles sans que l'utilisateur s'en rende compte.

Or la collecte de données suppose le consentement de l'utilisateur. Dans le cadre du big data ou bien du smartphone, les données sont collectées de toute évidence sans le consentement de l'utilisateur, ou n'a pas connaissance des méthodes adaptées dans l'usage du numérique. Un consentement volé ou un consentement donné sans connaissance de cause, les principes de loyauté et de licéité

invoqués par la loi informatique et libertés ne semblent pas respectés. En effet, le consentement préalable à toutes collectes de données soulève également des questions.

De manière détaillée, de nombreux principes sont difficilement applicables aux usages actuels du numérique. Les données résultant de l'utilisation d'un smartphone (Internet et connexion WIFI) sont soumises à un traitement automatisé et entrent dans le champ de l'article 2. En théorie, un régime protecteur doit être appliqué. Les principes de l'article 6 doivent s'appliquer. Ainsi, le principe de finalité, le principe de proportionnalité et le principe d'exactitude des données récoltées doivent s'appliquer.

Or, ces principes sont difficilement conciliables avec l'usage du numérique actuel (smartphone, Internet, objets connectés). Le big data par principe récolte en masse et en continue toutes les données. Cette pratique est incompatible avec le principe de finalité déterminée. Quant au principe de proportionnalité, il est difficile de l'accomplir en matière de big data puisque son principe même est l'imprévisibilité des résultats recherchés[14]. Les données collectées par le biais d'algorithme peut s'avérer fausses ainsi le principe de l'exactitude des données peut s'avérer impossible à atteindre. La loi ne semble plus en phase avec les évolutions technologiques. Il faut repenser la loi et rechercher une meilleure protection des données personnelles.

B- La recherche d'une meilleure protection des données personnelles face aux transformations des usages numériques

1- Le renforcement de la sécurité

Le renforcement de la sécurité a été un point mis en avant lors des nombreuses commissions sur la question des données à caractère personnel. Il a été considéré qu'il fallait recourir à de nouveaux instruments de protection de vie privée et des données. Dans ce cadre, le recours à des technologies protectrices de la vie privée a été envisagé en favorisant le développement de la *Privacy by design*[15] et de la *Privacy by default*.

- *Privacy by design*

« *La Commission propose d'encourager la conception et l'utilisation de technologies donnant à*

14 M. LANNA, « Le quantified self, un nouveau moteur du big data et menace pour la vie privée », Petites affiches, n°095, 12 mai 2016.

15 Assemblée Nationale, Rapport Numérique et libertés : un nouvel âge démocratique n° 3119 déposé le 9 octobre 2015, p 120.

tout internaute une réelle maîtrise sur l'utilisation de ses données ainsi que leur certification par des tiers indépendants, condition nécessaire à l'instauration d'une confiance et d'un différenciateur commercial pour ces technologies. »[16]. La recommandation n°50 de la Commission mise en place par l'Assemblée Nationale est d'encourager le développement de la *privacy by design* permettant de rendre effectif le principe de minimisation de la collecte de données personnelles.

L'article 25 du règlement général sur la protection des données, qui sera d'application directe au 25 mai 2018 introduit le concept de *privacy by design*. Le concept est de protéger la vie privée dès la conception. Les nouvelles technologies traitant des données personnelles doit garantir dès sa conception et lors de chaque utilisation, le plus haut niveau possible de protection des données. La *privacy by design* permet de répondre à la multiplication des traitements de données personnelles par des objets et technologies qui récoltent toujours plus de données personnelles. Cette mesure permet de répondre aux stratégies de collecte et d'utilisation des données personnelles de certaines entreprises. Il s'agit d'une mesure préventive.

- *Privacy by default*

L'article 25 du règlement général sur la protection des données, introduit aussi le concept de *privacy by default*. Il s'agit d'une protection des données par défaut qui consiste à ne collecter et traiter par défaut exclusivement les données à caractère personnel strictement nécessaires à la finalité poursuivie par le traitement. C'est une limitation à la collecte massive et l'étendue du traitement. Le consentement a du sens dans ce concept puisque l'utilisateur pourra changer les options de sécurité et accepter des collectes « massives » de données à caractère personnel.

- Le cadre légal

Dans le règlement il est prévu que le responsable de traitement devra mettre en œuvre des mesures techniques et organisationnelles pour assurer le respect des principes définis à l'article 5 : les principes de licéité, loyauté et transparence dans la collecte des données, limitation des finalités, exactitude des données traitées, limitation de la durée de conservation des données et intégrité et confidentialité des données. Le règlement mentionne les mesures de minimisation et de pseudonymisation des données à caractère personnel. Les mesures de minimisation des données personnelles consistent à ne collecter que des données adéquates, pertinentes et strictement limitées à ce qui est nécessaire pour réaliser les finalités pour lesquelles elles ont été collectées. Les mesures de pseudonymisation consistent à ce que les données soient traitées de telle façon qu'elles ne

16 Assemblée Nationale, Rapport Numérique et libertés : un nouvel âge démocratique n° 3119 déposé le 9 octobre 2015, p 120.

puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires conservées séparément[17].

En pratique, l'application de la *privacy by design* et de la *privacy by default* nécessite des process, des outils, des politiques internes et des procédures. Pour ce faire, il est nécessaire de mentionner ces principes dans le cahier des charges défini à l'occasion de la conclusion du contrat. D'ailleurs, ces principes devront être stipulés dans les contrats avec les sous-traitants.

Le renforcement de la sécurité est un point nécessaire pour renforcer la protection des données à caractère personnel, mais, cela est loin d'être suffisant. Le renforcement des droits et libertés doit aussi être recherché.

17 M. Griguer, J. Schwartz, « Privacy by Design/Privacy by Default : une obligation de conformité », cahiers de droit de l'entreprise, n°3, Mai 2017.

Télémédecine et sécurité des données de santé

Pauline Nicolas[1] **et Sébastien Cossin**[2]

[1]Doctorante en droit public, ATER, Université de Bordeaux
[2]Docteur en médecine, CHU de Bordeaux, Univeristé de Bordeaux

Résumé

La télémédecine est une forme de pratique médicale à distance qui pourrait permettre de lutter contre les déserts médicaux. Le secret médical est un droit fondamental du patient et s'impose à tous les médecins. Comment garantir ce secret lorsque les informations sont échangées à distance ? Dans cet article, nous nous sommes intéressés aux volets juridique et informatique encadrant les actes de télémédecine : que dit la loi sur la sécurité des données de santé échangées et comment est-elle traduite sur le plan informatique ?

Introduction

Contexte

Depuis plusieurs années, notre société est marquée par l'utilisation croissante des technologies de l'information et de la communication au sein de nombreux domaines, dont la santé. Ce phénomène est désigné par un terme générique: la «E-santé». Cette dernière englobe de nombreuses réalités. En effet, la «E-santé» recouvre la santé mobile, les objets connectés ou encore la télémédecine sur laquelle nous concentrerons cette étude.

Selon l'article L. 6316-1 du Code de la santé publique: «La télémédecine est une forme de pratique médicale à distance utilisant les technologies de l'information et de la communication. Elle met en rapport, entre eux ou avec un patient, un ou plusieurs professionnels de santé, parmi lesquels figure nécessairement un professionnel médical et, le cas échéant, d'autres professionnels apportant leurs soins au patient». Cette définition légale de la télémédecine appelle plusieurs remarques. En effet, l'expression «pratique médicale»[1] choisie par le législateur conduit à postuler que l'acte médical et la relation de soins sont dématérialisés. En corollaire, cette dématérialisation s'appuie sur «les technologies de l'information et de la communication», et non plus sur «des moyens de communications appropriés»[2]. En ce sens, la définition de la télémédecine a évolué concomitamment à l'évolution des techniques.

En outre, et sur le fondement de l'article L. 6316-1 du Code de la santé publique, le décret du 19 octobre 2010[3] définit les cinq actes de télémédecine (téléconsultation, téléexpertise, télésurveillance médicale, téléassistance médicale et la réponse médicale apportée dans le cadre de la régulation de l'aide médicale urgente ou de la permanence des soins), leurs conditions de mise en œuvre et de prise en charge financière.

[1]LE GOFFIC Caroline, «Consentement et confidentialité à l'épreuve de la télémédecine», Revue de droit sanitaire et social [en ligne], 2011, №6, p. 988. Disponible sur: www.dalloz.fr.
[2]Art. 32, loi №2004-810 du 13 août 2004 relative à l'assurance maladie, JORF №0190 du 17 août 2004.
[3]Décret №2010-1229 du 19 octobre 2010 relatif à la télémédecine, JORF №0245 du 21 octobre 2010.

Jusitification du sujet

La dématérialisation de la pratique médicale emporte, conséquemment, une dématérialisation des échanges de données. Ces dernières sont juridiquement encadrées. En effet, dès lors que des «données à caractère personnel» font l'objet «d'un traitement», les dispositions de la loi du 6 janvier 1978 révisée dite «Loi informatique et libertés»[4] doivent s'appliquer. En l'espèce, ces deux conditions[5], qui tiennent à la nature des données et à leurs modalités d'échanges, sont réunies dans le cadre des actes de télémédecine.

Les données échangées lors d'un acte de télémédecine doivent être qualifiées de données à caractère personnel relatives à la santé d'un individu. A ce titre, et pour la première fois, le règlement général européen sur la protection des données[6] (ci-dénommé RGPD), donne une définition légale des données de santé. Ainsi, et selon l'article 4, alinéa 15 du RGPD, les données concernant la santé sont: « les données à caractère personnel relatives à la santé physique ou mentale d'une personne physique, y compris la prestation de services de soins de santé, qui relèvent des informations sur l'état de santé de cette personne».

En outre, ces données de santé font l'objet d'un «traitement» c'est-à-dire, selon l'article 2, alinéa 3 de la loi «Informatique et libertés»: « toute opération ou tout ensemble d'opérations portant sur de telles données, quel que soit le procédé utilisé, et notamment la collecte, l'enregistrement, l'organisation, la conservation, [...], la communication par transmission, [...], l'effacement ou la destruction».

Traitées dans le cadre d'un régime juridique très protecteur, les données de santé constituent toujours des «données sensibles». Aussi, leur sécurité constitue un enjeu essentiel. En effet, la perte, le vol, le piratage de ces données peuvent être dommageables au regard de la vie privée du patient. Par exemple, la cyber-attaque mondiale, qui s'est produite le 14 mai 2017, a touché non seulement des entreprises privées mais également le système de santé britannique.

Méthodes

Dès nos premières recherches, nous nous sommes aperçus que les notions employées par les juristes (authentification, confidentialité) étaient mises en œuvre, d'un point de vue pratique, par les informaticiens. Afin d'étayer cette thèse, nous avons adopté une démarche inductive. En effet, il a fallu partir de l'acte de télémédecine afin de démontrer que la sécurité des données de santé joue à un triple-niveau, c'est-à-dire avant, pendant et après l'acte de télémédecine.

Cette démarche inductive permet de démontrer ce point de convergence: les notions avancées par les juristes en matière de sécurité sont traduites, sur un plan pratique, par les informaticiens grâce à de nombreux procédés, notamment des fonctions de cryptographie.

[4] Loi №78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

[5] MORLET-HAÏDIRA Lydia et RAHAL-LÖFSKOG Délia, «La télémédecine et la protection des données de santé par la loi informatique et libertés», Revue générale de droit médical [en ligne], 2012, №44, p. 336. Disponible sur: www.bnds.fr

[6] Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, abrogeant la directive 95/46/CE (règlement général sur la protection des données), JOUE L119/1 du 4 avril 2016.

Résultats

Sécurité des données de santé avant l'acte

Authentification des professionnels de santé

L'authentification des professionnels de santé intervenant dans l'acte constitue une condition préalable à la réalisation des actes de télémédecine. En effet, et selon l'article R. 6316-1, alinéa 1-a du Code de la santé publique: ≪Chaque acte de télémédecine doit être réalisé dans des conditions garantissant: l'authentification des professionnels de santé intervenant dans l'acte≫.

L'authentification a pour but de vérifier l'identité dont se réclame une personne ou une machine. Le Répertoire Partagé des Professionnels intervenant dans le système de Santé (RPPS) est le répertoire unique de référence qui rassemble et publie des informations permettant d'identifier les professionnels de santé, sur la base d'un « numéro RPPS » attribué au professionnel toute sa vie[7]. Le professionnel de santé s'identifie via son numéro RPPS et s'authentifie avec une carte à microcircuit contenant une bi-clé d'authentification (clé privée, clé publique) dédiée à cet usage, en provenance d'une IGC (infrastructures de gestion de clés) agréée par le groupement identifié dans l'article R.161-54 du code de la sécurité sociale[8]. Cette carte permettant l'authentification des professionnels de santé est appelée carte CPS (Carte de Professionnel de Santé). La carte CPS dispose d'un couple de clés, appelé bi-clé, utilisé par des algorithmes de cryptographie pour chiffrer et déchiffrer les échanges de messages. Pour transmettre un message, l'expéditeur utilise la clef publique du destinataire pour le chiffrer. Grâce à des propriétés mathématiques étonnantes basées sur les nombres premiers, seule la clef privée du destinaire est capable de déchiffrer le message. A la différence de la clé publique, la clé privée est par définition secrète et présente uniquement dans la carte CPS. En déchiffrant un message avec la carte CPS, le professionnel de santé peut apporter la preuve de son identité. Pour être utilisable, la carte CPS doit être dévérouillée par un mot de passe connu uniquement par le professionnel de santé détenteur de celle-ci. Il s'agit d'un mécanisme d'authentification fort reposant sur ce que la personne possède (la carte) et sur ce que la personne sait (le mot de passe). Pour usurper l'identité d'un professionnel de santé, il faudrait donc lui dérober sa carte CPS et son mot de passe. Les cartes CPS sont délivrées par l'Agence nationale des systèmes d'information partagés de santé (ASIP santé).

Identification du patient

L'identification du patient constitue également une condition essentielle pour réaliser un acte de télémédecine. A ce titre, et selon l'article R. 6316-1, alinéa 1-b: ≪Chaque acte de télémédecine est réalisé dans des conditions garantissant: l'identification du patient≫.

Depuis le décret du 27 mars 2017[9], le numéro d'inscription au répertoire national d'identification des personnes physiques (NIRPP ou NIR, encore appelé numéro de sécurité sociale) est utilisé en tant qu'identifiant national de santé. Ce dernier permet de: ≪référencer les données de santé de toute personne bénéficiant ou ayant vocation à bénéficier d'acte de prévention, diagnostic, thérapeutique, de compensation du handicap [...] ou d'actions nécessaires à la coordination de plusieurs actes≫. Cet identifiant national de santé acte la volonté des pouvoirs publics d'assurer une meilleure coordination des soins. En effet, et avant ce décret, chaque établissement devait générer son propre numéro d'identification patient et n'était pas autorisé à utiliser le NIR. L'utilisation d'identifiants différents entraînait un problème d'interopérabilité

[7] Arrêté du 18 avril 2017 modifiant l'arrêté du 6 février 2009 modifié portant création d'un traitement de données à caractère personnel dénommé ≪Répertoire partagé des professionnels intervenant dans le système de santé≫ (RPPS), JORF №0093 du 20 avril 2017.

[8] ASIP santé. *Référentiel d'authentification des acteurs de santé* (2013)

[9] Décret №2017-412 du 27 mars 2017 relatif à l'utilisation du numéro d'inscription au répertoire national d'identification des personnes physiques comme identifiant national de santé, JORF №0075 du 29 mars 2017.

des systèmes d'information à communiquer l'identité d'un patient.

Disponibilité du service

Le notion informatique pour désigner l'accès permanent à un service avec un temps de réponse acceptable est la disponibilité du service. Elle est une condition à la réalisation des actes de télémédecine mentionnée à l'article R. 6316-1, alinéa 1-C: ≪Chaque acte de télémédecine est réalisé dans des conditions garantissant: l'accès des professionnels de santé aux données médicales du patient nécessaires à la réalisation de l'acte≫. Un service indisponible pour échanger les données médicales du patient rend l'acte de télémédecine impossible.

Sécurité des données de santé pendant l'acte

La sécurité des données de santé doit être également effective durant l'acte de télémédecine. A ce titre, il s'agit tant d'assurer la confidentialité que l'intégrité des données de santé.

Confidentialité des données de santé

La confidentialité n'est pas synonyme de secret. En effet, si le ≪secret≫ relève de ce qui est inconnaissable, la confidentialité désigne, quant à elle, ≪ce qui est limité à un cercle restreint≫[10].

La confidentialité des données, comme condition de leur sécurité, est formulée au sein de l'article 34 de la loi≪Informatique et libertés≫: ≪le responsable de traitement est tenu de prendre toutes précautions utiles, au regard de la nature des données et des risques présentés par le traitement, pour préserver la sécurité des données et, notamment empêcher [...] que des tiers non autorisés y aient accès≫.

Comme discuté plus haut, la carte CPS permet de chiffrer et de déchiffrer des messages. Ce chiffrement garantit que seules les personnes autorisées ont accès aux messages échangés. Ces étapes de chiffrement et de déchiffrement sont réalisées lors de la transmission des données. Les développeurs de service de télémédecine peuvent s'appuyer sur des protocoles de sécurisation des échanges sur Internet déjà utilisés dans d'autres domaines (paiement en ligne, consultation d'emails...). Par exemple, le protocole SSL (Secure Socket Layer) est le protocole le plus répandu actuellement. La première étape du protocole est l'échange, entre deux machines, des clés publiques. Chaque machine vérifie la clef publique du destinataire auprès de l'ASIP santé qui joue le rôle de tiers de confiance dans le domaine de la santé. Les machines s'échangent ensuite, de façon sécurisée avec leur clef publique, une clé de chiffrement dite symétrique qui sécurisera la transmission des données.

Intégrité des données de santé

Selon une commune acception, l'intégrité désigne: ≪L'Etat d'une chose, d'un tout, qui est entier, qui a toutes ses parties≫. Appliquée aux données de santé, cette entièreté demeure. En effet, l'intégrité des données de santé suppose qu'elles soient exactes et correctement utilisées[11].

L'intégrité des données de santé, comme condition de préservation de leur sécurité, est énoncée stricto sensu au sein de l'article 34 de la loi ≪Informatique et libertés≫: ≪Le responsable de traitement est tenu de prendre toutes précautions utiles, au regard de la nature des données et des risques présentés par le traitement, pour préserver la sécurité des données et, notamment, empêcher qu'elles soient déformées, endommagées[...]≫.

Lors de la transmission des données, le protocole TCP (transmission control protocol) est dit fiable car il s'assure que le message envoyé arrive à destination intègre. L'intégrité de la donnée est aussi vérifiée en

[10] Centre National de Ressources Textuelles et Lexicales, Confidentiel [en ligne]. Disponible sur: `http://www.cnrtl.fr/definition/confidentiel`

[11] BENEJAT Muriel, ≪Les droits sur les données personnelles ≫ In SAINT-PAU Jean-Christophe (dir), Droits de la personnalité, Paris: LexiNexis, coll. ≪Traités≫, 2013, p. 598

calculant une empreinte par une fonction mathématique appelée fonction de hachage. Une fonction de hachage prend en entrée un fichier (suite de bits) d'une taille quelconque et calcule une valeur, appelée empreinte, de longueur fixe (128 bits pour la fonction md5 par exemple). La moindre modification d'un bit d'un fichier entraîne une empreinte différente. En comparant l'empreinte du fichier de l'expéditeur et l'empreinte du fichier reçu, le destinataire est capable de savoir rapidement si le fichier a été modifié ou altéré.

Sécurité des données de santé après l'acte

La sécurité des données de santé ne prend pas fin à l'issue de l'acte de télémédecine. Elle perdure après l'acte c'est-à-dire au moment de la conservation et du stockage des données. La sécurité des données de santé après la réalisation de l'acte de télémédecine met en lumière une procédure, l'hébergement des données de santé, qui doit être conforme à des référentiels d'interopérabilité et de sécurité.

Hébergement des données de santé: signification et conséquences

L'hébergement des données de santé se définit comme: «L'externalisation des données de santé auprès d'un organisme spécialisé, distinct du professionnel ou de l'établissement de santé qui soigne le malade [...]. Le but recherché est d'amener les acteurs de l'hébergement au plus haut niveau de sécurité afin d'offrir un espace de confiance aux patients et aux professionnels de santé»[12].

Défini à l'article L. 1111-8 modifié du Code de la santé publique, l'hébergement des données de santé obéit à plusieurs conditions. D'une part, un organisme ne peut héberger des données de santé que s'il a préalablement obtenu un agrément. Cette obligation légale appelle une procédure spécifique explicitée aux articles R. 1111-9 à R. 1111-15 du Code de la santé publique. D'autre part, l'hébergement des données de santé ne nécessite plus le consentement préalable du patient. En effet, la loi du 26 janvier 2016 relative à la modernisation de notre système de santé[13] supprime la condition du consentement et la remplace par une obligation d'information assortie d'un droit d'opposition du patient, mais uniquement pour un motif légitime. In fine, toute prestation d'hébergement doit faire l'objet d'une convention.

La possibilité de faire héberger des données de santé issues d'acte de télémédecine est expressément énoncée à l'article R. 6316-10 du Code de la santé publique: «Les organismes et les professionnels de santé utilisateurs des technologies de l'information et de la communication pour la pratique d'acte de télémédecine s'assurent que l'usage de ces technologies est conforme aux dispositions prévues au quatrième alinéa de l'article L. 1111-8 du Code de la santé publique relatif aux modalités d'hébergement des données à caractère personnel».

L'hébergement des données de santé, qualifiées de données sensibles, appelle une politique de sécurité spécifique. En effet, et selon l'article R. 1111-9, alinéa 2 du Code de la santé publique: les candidats à l'obtention d'un agrément doivent: «Définir et mettre en œuvre une politique de confidentialité destinée à assurer des exigences de confidentialité et de secret, la protection contre les accès non autorisés et la pérennité des données».

En pratique, il est nécessaire de passer par une procédure d'agrément pour être autorisé à héberger des données de santé. Les dossiers de demande d'agrément sont traités par l'ASIP santé. Le dossier est analysé selon trois volets: juridique et éthique, économique et financier, sécurité et technique. Le dossier d'agrément est constitué de plusieurs formulaires décrivant ces différents volets. Par exemple, le formulaire P6 est un document de 31 pages décrivant en détail les dispositifs de sécurité mis en place.

[12] Commission nationale de l'informatique et des libertés, «Guide professionnels de santé» [en ligne]. Les guides de la CNIL, janvier 2011, p. 32. Disponible sur: `http://www.cnil.fr/fileadmin/documents/Guides_pratiques/CNILGuide_professionnels_de_sant.pdf/`

[13] Loi №2016-41 du 26 janvier 2016 de modernisation de notre système de santé, JORF №0022 du 27 janvier 2016

Un exemplaire est aussi transmis à la CNIL pour avis. Celle-ci est autorisée à réaliser des contrôles et à sanctionner les industriels lorsque des manquements sont identifiés.

Conclusion

Dans cet article, nous avons mis en correspondance les textes juridiques et les notions informatiques qui encadrent la sécurité des données lors d'un acte de télémédecine. Les principales mesures de sécurité que sont l'identification, l'authentification, la disponibilité, l'intégrité et la confidentialité sont abordées dans les textes juridiques même si ces termes ne sont pas toujours explicitement mentionnés. L'ASIP santé publie des référentiels pour aider les industriels à mettre en oeuvre ces mesures et la CNIL veille à la conformité des dispositifs mis en place pour garantir le secret médical.

Labellisation et certification des traitements de DACP : enjeux juridiques et techniques

Marcel Moritz†, Mathieu Cunche‡

†CERAPS, université de Lille
marcel.moritz@univ-lille2.fr

‡ Univ Lyon, INSA Lyon, Inria, CITI
mathieu.cunche@insa-lyon.fr

1 Introduction : quelques éléments de définition

La certification peut être définie comme une "Assurance écrite (sous la forme d'un certificat) donnée par une tierce partie qu'un produit, service ou système est conforme à des exigences spécifiques" [1]

La labellisation est une expression de cette certification, qui a le mérite d'être parlante pour le consommateur. "Ce "signe" est destiné à rassurer, à mettre en confiance et à aboutir à une transaction, un acte d'achat ou d'usage ou d'emploi"[7] [2].

Derrière ces pratiques, il y a nécessairement un référentiel technique, qui permet de s'assurer que les exigences requises pour l'obtention du certificat, du label, sont respectées à l'issue d'une procédure de certification.

Les données à caractère personnel (DACP) : Selon l'article 2 de la loi 78-17 du 6 janvier 1978 modifiée, constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou pouvant être identifiée, directement ou indirectement, que ce soit par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Il s'agit par exemple : d'un nom, d'un numéro de sécurité sociale, d'un numéro de téléphone, d'une photographie, d'éléments biométriques tels que l'empreinte digitale ou ADN, etc.

Mais un ensemble d'informations a priori anodines deviennent des données personnelles si, par recoupement, l'identification devient possible. Une date de naissance, couplée avec un lieu de naissance, permet ainsi l'identification de la personne concernée. Il s'agit donc bien d'une donnée à caractère personnel. En conséquence, plus les techniques progressent, plus le nombre d'informations personnelles stockées augmente et plus leur recoupement est facilité.

A ce jour certains traitements de données à caractère personnel doivent faire l'objet d'une déclaration auprès de la CNIL (Commission nationale de l'informatique et des libertés), voire d'une demande d'autorisation, par exemple lorsque sont concernées des données sensibles.

1. `iso.org`
2. E. Sutter, Certification et labellisation : un problème de confiance. Bref panorama de la situation actuelle, Documentaliste Sciences de l'information, 2005/4, p. 284.

2 L'importance de l'enjeu juridique

Le règlement européen 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (dit RGPD) entrera en application le 25 mai 2018. Il renouvelle en profondeur le droit applicable à la protection des DACP et accorde précisément une place importante à la certification.

Le considérant n° 100 (qui précède les articles du règlement et en explicite les motivations) dispose ainsi :

> "Afin de favoriser la transparence et le respect du présent règlement, la mise en place de mécanismes de certification ainsi que de labels et de marques en matière de protection des données devrait être encouragée pour permettre aux personnes concernées d'évaluer rapidement le niveau de protection des données offert par les produits et services en question".

Surtout, l'article 42 traite de la certification ; il peut être résumé ainsi :

- La protection des DACP est un enjeu juridique majeur, mais l'application du droit positif pur ne suffit pas à garantir son respect. Il suffit à cet égard de comparer les CGU de certains sites (ex. des réseaux sociaux) et de comparer cela au droit positif.
- Pour y remédier à moindre coût, l'UE mise sur une forme d'autodiscipline (la privacy by design en est une illustration)
- Cette autodiscipline est mieux acceptée par les entreprises car moins "autoritaire" et susceptible de faciliter la mise en conformité des PME. C'est la même logique quel celle qui préside au développement de codes de conduite (art. 40 et 41 du RGPD)
- La certification est une démarche volontaire
- Une certification est valable pour une durée maximale de trois ans
- Les certificats / labels peuvent viser tant les responsables de traitement (définis à l'article 3 de la loi de 1978) que les sous-traitants

3 Le choix des organismes de certification, labellisation

L'article 43 du RGPD porte sur la détermination des organismes de certification. Il laisse la porte grande ouverte aux entreprises privées (les démarches publiques n'ont pas eu beaucoup de succès à ce jour, voir Europrise [3]). Mais "l'autorité de contrôle compétente" (art. 42) a aussi un rôle à jouer, de sorte que va s'installer une concurrence entre personnes privées (ex : Bureau Veritas, AFNOR) et publiques. Et il n'y aura pas de place pour tout le monde car le public n'acceptera qu'un nombre limité de certifications, de labels.

Nous sommes donc face un à un problème : on a, d'une part, de fréquentes divergences de point de vue entre autorités de contrôle, et des moyens très relatifs ; d'autre part, un marché colossal de la labellisation / certification. Dans ce contexte, quelle place effective restera-t-il à la puissance publique ?

Certains risques sont déjà prévisibles :

3. `https://www.european-privacy-seal.eu/EPS-en/Home`

— L'agrément du certificateur est de 5 ans maximum ; il peut être renouvelé dans les mêmes conditions tant que l'organisme de certification satisfait aux exigences énoncées au RGPD. Ce délai peut être considéré comme trop long au regard de la rapidité des évolutions technologiques.
— Le coût de la certification / labellisation pourrait être prohibitif (ex. Europrise : un expert juriste, un expert technique, 1000 euros par jour chacun...).
— Le double niveau de certification (au niveau européen mais aussi national) risque de poser des problèmes :
  — difficultés de reconnaissance mutuelle des certifications nationales ;
  — prolifération contre-productive des labels et des marques ;
  — définition du périmètre de ces labels (simple conformité au RGPD ou plus étendu ?) ;
  — risque de "forum shopping" si certains États membres adoptent des critères de certification plus souples.

4 Les sanctions en cas de violation d'une certification

La certification fait présumer la légalité du traitement de DACP, mais elle n'enlève rien à la sévérité des sanctions qui pourront être prononcées en cas d'illégalité du traitement. L'article 42 est clair : "Une certification en vertu du présent article ne diminue pas la responsabilité du responsable du traitement ou du sous-traitant quant au respect du présent règlement et est sans préjudice des missions et des pouvoirs des autorités de contrôle qui sont compétentes en vertu de l'article 55 ou 56". Une certification non respectée pourrait même être interprétée de deux manières diamétralement opposées selon les faits :

— soit comme une preuve de bonne volonté de l'entreprise susceptible de minorer la sanction (voire même de l'exempter),
— soit comme une mauvaise foi caractérisée (l'entreprise s'est faite certifier, mais en pratique a agi en violation manifeste de ses engagements), auquel cas la sanction pourrait bien être accrue.

Fort logiquement, l'article 58 du RGPD permet aux autorités de contrôle (en France la CNIL) de "retirer une certification ou ordonner à l'organisme de certification de retirer une certification délivrée en application des articles 42 et 43, ou ordonner à l'organisme de certification de ne pas délivrer de certification si les exigences applicables à la certification ne sont pas ou plus satisfaites".

La responsabilité des organismes de certification est aussi susceptible d'être mise en jeu, l'article 83 du RGPD disposant :

> "Les violations des dispositions suivantes font l'objet, conformément au paragraphe 2, d'amendes administratives pouvant s'élever jusqu'à 10 000 000 EUR ou, dans le cas d'une entreprise, jusqu'à 2% du chiffre d'affaires annuel mondial total de l'exercice précédent, le montant le plus élevé étant retenu : [...] b) les obligations incombant à l'organisme de certification en vertu des articles 42 et 43".

La sanction est potentiellement très lourde pour l'organisme certificateur. Face à cela, le certificateur va probablement se placer dans une logique largement déclarative : l'entreprise s'engage à respecter, mais sans véritable contrôle

concret.

5 La base documentaire pour la constitution d'un référentiel technique

La mise en place d'une certification, ou d'une labellisation, va devoir reposer sur un référentiel d'exigence sur lequel s'appuieront les processus d'attribution et de renouvellement du certificat, ou du label. Ce référentiel devra s'appuyer sur d'autres documents décrivant l'état de l'art en matière de protection des données personnelles et de sécurité de l'information.

Parmi les documents pouvant servir de base à ce référentiel, on peut citer les documents de la CNIL traitant du Privacy Impact Assesment [4]. Ces documents ne traitent pas que du PIA, mais contiennent également une liste de bonnes pratiques dans le cadre de la gestion de DACP. Ces bonnes pratiques pourraient être transformées en exigences afin de constituer un référentiel.

Le G29 a également publié des avis sur le PIA (*Guidelines Data Protection Impact Assessment (DPIA)*). Ce document se focalise sur le PIA et précise les conditions et la mise en place de celui-ci. Il inclut une liste de point de vérification permettant de s'assurer qu'une technique de PIA permet de couvrir l'ensemble des points du GDPR. La mise en place d'un PIA étant un des éléments nécessaires à l'obtention du label, cette liste pourrait servir de support pour établir les points de vérifications du PIA lors d'un audit.

L'étude de ces documents fait apparaître que la sécurité du système d'information (SI) est un point central dans la protection de données personnelles. En effet, le système qui héberge ces données doit être en mesure de garantir la sécurité de celles-ci et en particulier leur confidentialité. Un certain nombre de mesures doivent être mise en place afin de garantir cette sécurité. Ainsi il est directement fait référence à des documents centraux dans la protection des systèmes d'information : le Référentiel Général de Sécurité (RGS) et la norme ISO-27001. Le Référentiel général de sécurité (RGS)[1] décrit un ensemble de bonnes pratiques pour la sécurisation d'un système d'information, tandis que la norme internationale ISO-27001[4] décrit un ensemble d'exigences visant à mettre en place un management de la sécurité de l'information. Le RGS et la norme ISO-27001 (et d'autres normes suite ISO/CEI 27000) sont abondamment cités par les documents PIA, en particulier lorsqu'il s'agit d'aborder les aspects techniques de la sécurisation du SI.

On peut également considérer l'Instruction interministérielle relative à la protection des systèmes d'informations sensibles (II901)[2]. Cette Instruction définit les objectifs et les règles relatifs à la protection des systèmes d'information sensibles, notamment ceux traitant des informations portant la mention Diffusion Restreinte[6, Annexe 3]. Un des objectifs visés, "prévenir la compromission d'informations sensibles", pourrait correspondre aux besoins de protection des DACP. Cette instruction vise la sécurisation et la gestion du système d'information au sens large (numérique et analogique). Les exigences de l'II901 pourraient être déclinées en exigences pour certification DACP. Elles sont par ailleurs déjà utilisées dans le cadre de la certification ANSSI des prestataires de

4. `https://www.cnil.fr/fr/PIA-privacy-impact-assessment`

détection des incidents (PDIS) et des prestataires de réponse aux incidents de sécurité (PRIS) [5].

6 Les exigences techniques requises pour la certification

La base documentaire permet de faire apparaître plusieurs axes autour desquels pourrait s'organiser un référentiel d'exigence relatif à la protection des DACP. Ces axes sont les suivants :

- Règles et principes fondamentaux relatifs aux DACP :
  - Minimisation et limitation de la durée de conservation des DACP
  - Information et consentement
  - Droits d'opposition, d'accès et de rectification
  - Anonymisation
- Sécurité et traçabilité :
  - Chiffrement des DACP
  - Traçabilité des opérations
  - Contrôle d'accès (logique et physique)
  - Marquage des documents
- Gestion des aspects humains :
  - Charte dédié aux enjeux des DACP
  - Formation du personnel (besoin de formations spécialisées)
  - Présence d'un CIL (Correspondant Informatique et Libertés) / *Data protection officer*
- Existence de processus :
  - Procédure de PIA
  - Procédure de gestion de fuites de DACP (généralisation des obligations de notification des fuites de DACP dans le cadre du RGPD)

7 Un exemple d'exigence : le cas de l'anonymisation

7.1 Principes de l'anonymisation

Le processus d'anonymisation des données est un point critique car il permet d'étendre l'utilisation des données, voire de les partager avec d'autres entités. De plus la mise en oeuvre de l'anonymisation est un processus complexe qui peut grandement varier en fonction du type de données.

L'anonymisation de DACP doit satisfaire à certains critères comme le rappelle le G29[3]. Elle doit prendre en considération les possibilités suivantes :

- l'individualisation, qui correspond à la possibilité d'isoler une partie ou la totalité des enregistrements identifiant un individu dans l'ensemble de données ;
- la corrélation, qui consiste dans la capacité de relier entre elles, au moins, deux enregistrements se rapportant à la même personne concernée ou à

5. `https://www.ssi.gouv.fr/entreprise/qualifications/qualifications-de-prestataires/`

un groupe de personnes concernées (soit dans la même base de données, soit dans deux bases de données différentes). Si une attaque permet d'établir (par exemple, au moyen d'une analyse de corrélation) que deux enregistrements correspondent à un même groupe d'individus, mais ne permet pas d'isoler des individus au sein de ce groupe, la technique résiste à l'"individualisation", mais non à la corrélation ;
— l'inférence, qui est la possibilité de déduire, avec un degré de probabilité élevé, la valeur d'un attribut à partir des valeurs d'un ensemble d'autres attributs.

7.2 Détails de l'exigence d'anonymisation

La conformité vis-à-vis de l'exigence sur l'anonymisation pourrait être contrôlée via les points de vérification suivants :

- Existence d'un processus d'anonymisation et modalité d'utilisation :
  - Les données concernées sont identifiées, et l'anonymisation apparaît à un moment précis dans le traitement des données.
  - La technique d'anonymisation employée doit être définie avec précision et une description de l'algorithme doit être fournie afin de réduire la part d'interprétation lors du passage à l'implémentation.
  - Les limites de la méthode d'anonymisation doivent être identifiées pour être prises en compte dans le PIA.
- Méthode d'anonymisation conforme aux standards (CNIL et G29).
  - La méthode d'anonymisation correspond aux standards existant, et/ou à l'état de l'art et est une technique éprouvée.
- Implémentation correcte du processus d'anonymisation :
  - Cette implémentation doit correspondre à la technique décrite dans les documents.
  - Cette implémentation doit faire l'objet de tests et si possible doit être soumise à des outils de vérification de code.
- Utilisation systématique du processus d'anonymisation dans les cas définis :
  - Le processus d'anonymisation est bien intégré au système de l'entreprise et à ses produits.
  - Les données qui doivent être anonymisées le sont effectivement.

7.3 Difficultés

Accès de l'auditeur à des DACP La vérification d'une anonymisation correcte va nécessiter l'accès aux DACP par l'auditeur. Par exemple, lorsqu'il est nécessaire de vérifier qu'il n'est pas possible de remonter aux données originales à partir des données supposées anonymisées. Si le certificateur a accès aux données, il est possible que cela ne respecte pas l'article 32 de la loi 1978 selon lequel : "I.-La personne auprès de laquelle sont recueillies des données à caractère personnel la concernant est informée, sauf si elle l'a été au préalable, par le responsable du traitement ou son représentant : (...) 5 Des destinataires ou catégories de destinataires des données".

BigData Le contexte du Big Data et la multiplication des jeux de données va engendrer des difficultés pour le contrôle de l'anonymisation. En effet, les risques de réidentification sont proportionnels au volume de données disponibles.

Plusieurs données non identifiantes peuvent ensemble identifier une personne[8, 5]. Le RGPD prend en compte cet enjeu : cons. 26 :

> "Les données à caractère personnel qui ont fait l'objet d'une pseudonymisation et qui pourraient être attribuées à une personne physique par le recours à des informations supplémentaires devraient être considérées comme des informations concernant une personne physique identifiable. Pour déterminer si une personne physique est identifiable, il convient de prendre en considération l'ensemble des moyens raisonnablement susceptibles d'être utilisés par le responsable du traitement ou par toute autre personne pour identifier la personne physique directement ou indirectement (...)".

L'accroissement rapide du nombre de sources de données auxiliaire va complexifier la tache de vérification et augmenter significativement son temps d'exécution. Cette tâche de vérification pourrait devenir impossible à effectuer avec des ressources raisonnables.

8 Des modalités concrètes de certification demeurant floues

A quelle fréquence effectuer les audits de rappel ? Le RGPD établit une simple durée maximale pour la validité de la certification (trois ans). Pour le reste, on peut se référer à l'article 32 RGPD sur la sécurité du traitement, qui est très flou :

> "Compte tenu de l'état des connaissances, des coûts de mise en œuvre et de la nature, de la portée, du contexte et des finalités du traitement ainsi que des risques, dont le degré de probabilité et de gravité varie, pour les droits et libertés des personnes physiques, le responsable du traitement et le sous-traitant mettent en œuvre les mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté au risque, y compris entre autres, selon les besoins : (...) d) une procédure visant à tester, à analyser et à évaluer régulièrement l'efficacité des mesures techniques et organisationnelles pour assurer la sécurité du traitement".

Qui pour faire cet audit ? Quelles compétences dans l'équipe d'audit ? A priori cela dépend beaucoup de la nature de l'audit. Certains sont fondés sur un mécanisme déclaratif d'auto-contrôle et proposés par des juristes spécialisés ou cabinets d'avocat. Il existe même des logiciels d'assistance à la mise en conformité. Mais cela est très différent si la certification repose sur un mécanisme d'audit technique. Des compétences techniques seront alors indispensables pour vérifier la sécurité du système d'information et le traitement/stockage des DACP.

Quel avenir pour la certification et la labellisation ? Les modalités concrètes de mise en oeuvre des mécanismes de certification et de labellisation restent donc encore très largement à définir. De leur pertinence dépendra grandement la réussite ou l'échec de cette formule, préconisée par le RGPD.

Références

[1] DGME / ANSSI. Référentiel général de sécurité (RGS), June 2014. v2.0.

[2] SGDSN / ANSSI. II901 : Instruction interministérielle relative à la protection des systèmes d'informations sensibles. PRMD1503279j, DGME / ANSSI, November 2015. v2.0.

[3] Groupe de travail ≪ARTICLE 29≫ sur la protection des données. Avis 05/2014 sur les Techniques d'anonymisation. Technical Report 0829/14/FR WP216, October 2014.

[4] Technologies de l'information - Techniques de sécurité - Systèmes de gestion de sécurité de l'information - Exigences. Standard, International Organization for Standardization, Geneva, CH, 2013.

[5] Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. Unique in the Crowd : The privacy bounds of human mobility. *Scientific Reports*, 3 :1376, March 2013.

[6] SGDSN/PSE/PSD. II1300 : Instruction interministérielle sur la protection du secret de la défense nationale. Technical report, SGDSN/PSE/PSD, November 2011.

[7] Éric Sutter. Certification et labellisation : un problème de confiance. Bref panorama de la situation actuelle. *Documentaliste-Sciences de l'Information*, 42(4) :284, 2005.

[8] Latanya Sweeney. Simple demographics often identify people uniquely. *Health (San Francisco)*, 671 :1–34, 2000.

Je sais ce que tu fais, je sais qui tu es…
Éthique et Données Personnelles

Linda Arcelin
Faculté de Droit, de Science Politique et de Gestion
Université de La Rochelle
45 rue François de Vaux de Foletier
17024 La Rochelle cedex 1
linda.arcelin@univ-lr.fr

Christophe Nicolle
Checksem – LE2I
Univ. Bourgogne Franche-Comté
Institut Marey, 64 rue de Sully
21000 Dijon, France
cnicolle@u-bourgogne.fr

Dans la mythologie grecque, Argos était un géant, capable de tout voir. Il possédait cent yeux, répartis sur toute la tête. En permanence, 50 d'entre eux surveillaient pendant que les autres dormaient. Nul ne pouvait tromper sa vigilance. Héra lui avait confié la tâche de prévenir les tentatives d'infidélités de son mari Zeus avec la prêtresse Io. De nos jours, nous souhaiterions tous pouvoir disposer d'un géant vigilant qui veille sur notre maison et nos proches. A contrario, nous savoir observés dans notre quotidien par un œil étranger représente une certaine angoisse.
A l'heure où les villes, les centres commerciaux, les maisons s'équipent d'œil électronique et de systèmes d'analyse de nos comportements (détection de présence, détection de chute, profilage de nos comportements routiers ou de nos comportements de consommation), la question du respect de notre vie privée se pose. La captation de nos actes, leur analyse et le potentiel jugement par un œil humain distant peut-elle se justifier ? Comment mesurer l'étendue ce que représentent nos « données personnelles », peut-on accepter que notre accès aux informations et produits soit limité par des algorithmes de recommandation qui nous profilent ?
Est-ce que nous offrir des choix liés à un profilage n'est pas réduire notre libre arbitre, améliorant à court terme l'efficacité, mais réduisant à long terme notre évolution personnelle ?

Il faut rappeler que le principe de base est que chaque individu puisse maîtriser la collecte et le traitement de ses données personnelles. La notion de données personnelles est encadrée. Pour le Règlement de 2016, il s'agit de *« toute information se rapportant à une personne physique identifiée ou identifiable ; est réputée être une «personne physique identifiable» une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale »*. Je souhaiterais faire observer que cette notion peut évoluer au fur et à mesure des évolutions technologiques. Ainsi, au rang des données personnelles qui sont de plus en plus utilisées par l'industrie publicitaire, figurent par exemple les émotions qui sont mesurées grâce à un protocole basé sur la technologie du FACS (facial action coding system). Les émotions sont classées en 7 catégories : joie, surprise, peur, dégoût, mépris, tristesse et colère.

Le problème dans la captation des émotions ou des données dites personnelles, quels que soient l'ensemble qu'elles représentent et la manière dont leur exploitation va impacter durablement notre libre arbitre. Dans une situation sans ordinateur, un vendeur peut exploiter la captation d'une émotion pour orienter le client vers un autre produit. Cette action va se restreindre à la durée de l'entretien. Face à un ordinateur, cette information sera conservée et partagée sans que le client puisse intervenir et peut être transformé son rapport émotionnel au

produit. L'algorithme va réduire de facto l'ensemble des options possibles à celles associées à des émotions positives.

Par principe, il est toujours possible de refuser de transmettre ses données, de refuser de s'inscrire sur un site, de refuser les cookies. Pour autant, ces facultés ne sont pas forcément effectives même si l'on tend à des améliorations. D'ailleurs, cette position peut aussi limiter l'accès aux services proposés par un appareil électronique, voir empêcher le fonctionnement de cet appareil. Cela peut créer une fracture numérique. On peut se trouver alors dans l'obligation d'accepter, bien obligé par le business model associé au système électronique. Un cas flagrant est SIRI, où l'ensemble des données personnelles peut être exploité par la société indépendamment de la volonté du client. Si dans votre carnet d'adresses électroniques le terme « Papa » est associé à un numéro, la société peut reconstruire votre filiation et même votre arbre généalogique avec le carnet d'adresses électroniques d'une sœur inconnue qui possède une entrée « papa » avec le même numéro de téléphone.

D'une part, si l'internaute doit donner son consentement à la collecte et au traitement de ses données personnelles, les informations peuvent être puisées aujourd'hui dans bien d'autres sources que celles dispatchées par l'internaute lui-même : les datas shadow, les métadonnées découlant des communications électroniques[1] sont légion. En particulier, ces dernières, les données sur « la donnée », peuvent fournir des informations personnelles sur les individus. Une étude réalisée par trois chercheurs de Stanford a ainsi montré qu'il est possible de déduire qu'une personne est cardiaque au regard des appels qu'elle a passés à un cardiologue et à une pharmacie et de leur durée[2].
Aujourd'hui il existe des algorithmes de vision artificielle qui, avec n'importe quelle webcam, peuvent capter vos paramètres vitaux et vous conseiller de prendre contact avec votre cardiologique.
La proposition de Règlement concernant le respect de la vie privée et la protection des données à caractère personnel dans les communications électroniques, abrogeant la directive 2002/58/CE (Règlement e-privacy) du 10 janvier 2017, considère ces métadonnées pour leur faire suivre le même régime que celui applicable aux données. En complément, les OTT (Over the Top service, de type Skype, Facebook messenger ou WhatsApp), qui jusqu'alors, échappaient à la directive de 2002, sont appréhendés. Dans le souci de promouvoir un développement équilibré du commerce électronique, la Commission préconise des *« conditions de concurrence équitables pour des services numériques comparables »*, c'est-à-dire entre les fournisseurs de services de communication traditionnels et les messageries par contournement[3]. La proposition de Règlement e-privacy aligne les régimes à l'ensemble des acteurs, c'est-à-dire les opérateurs de communication électronique tels que les fournisseurs d'accès à internet, mais également les services OTT.

D'autre part, si les cookies et autres traceurs – qui permettent de cibler la publicité – doivent être acceptés par l'internaute (art. 32 II de la loi de 1978), la réalité montre que celui-ci n'a guère le choix. Bien souvent, avant même que l'internaute n'ait donné son accord, il est déjà « tracé ».

---

[1] *« les numéros appelés, les sites Web visités, le lieu, la date, l'heure et la durée des appels passés par un individu, etc. »* et *« permettent de tirer des conclusions précises sur la vie privée des personnes intervenant dans la communication électronique, comme leurs rapports sociaux, leurs habitudes et activités au quotidien, leurs intérêts, leurs goûts, etc. » (*Considérant n° 2).

[2] J. Mayer, P. Mutchler and J.C. Mitchell, Evaluating the privacy properties of telephone metadata : PNAS, May 17, 2016.

[3] Communication – Les plateformes en ligne et le marché unique numérique – Perspectives et défis pour l'Europe, 25 mai 2016, SWD(2016) 172 final, p. 6.

Il existe même des systèmes capables de profiler le comportement d'un internaute au travers de ces lectures sur le net sans avoir besoin de cookies. Comme le disait François Mauriac, « Dis-moi ce que tu lis, je te dirais qui tu es. » Preuve que la volonté de profiler existe depuis longtemps, indépendamment de tout logiciel au système électronique.

La proposition de règlement présentée le 10 janvier 2017 envisage de permettre aux internautes d'accepter ou refuser les cookies et autres traceurs directement dans les paramètres de confidentialité de leur navigateur. L'acceptation par défaut serait permise, ce qui rassure l'industrie publicitaire, mais inquiète des associations telles que la Quadrature du net pour qui l'internaute lambda ne vérifie pas les paramètres de son ordinateur et sera donc aux prises de fuite de ses données, d'annonces ciblées... Ces questions de *"tracking*" sont au cœur des préoccupations du moment et certains opérateurs proposent des solutions apparemment « responsables ». Ainsi, Apple vient d'annoncer que la nouvelle version de son navigateur Safari intégrera un système "*anti-tracking*". Ce bloqueur de *"tracking*" serait fondé sur un système de *"deeplearning"* analysant les habitudes de navigation de l'internaute et déterminant les cookies qui doivent être bloqués de ceux qui ne doivent pas l'être. Certaines voix s'élèvent pour dénoncer une solution technique qui ne ferait que renforcer la position déjà dominante d'acteurs de la publicité en ligne comme Google et Facebook, sites sur lesquels les internautes se rendent au moins une fois par jour[4].

Le problème de la traçabilité est qu'aujourd'hui l'internaute n'a plus conscience du risque qu'il prend en diffusant lui-même les pièces du puzzle de sa vie, librement et massivement. Il n'y a pas d'éducation de la préservation de la donnée personnelle. Photo, commentaires, discussions sont autant d'informations capitales pour un profilage que des informations de cookies. Beaucoup d'adolescents considèrent l'exposition de leur vie privée comme un must have. Ils n'ont pas conscience de mettre à crédit leur libre arbitre futur.

Ces captations des données sans l'accord des intéressés et pour des raisons commerciales peuvent choquer. Mais en revanche, peut-on imaginer que l'on aille contre la volonté de l'individu, pour son bien ou pour le bien de la société, en dévoilant et traitant ses données personnelles contre son gré ? Nous aurions donc un algorithme éthique qui ferait lui-même la part des choses… Comme il a déjà été observé, *« il convient surtout de ne pas confondre la robot-éthique avec l'éthique des robots (…). L'éthique des machines implique que le robot lui-même doive respecter des règles éthiques, alors que la robot-éthique s'applique à l'homme, qui va concevoir ou utiliser des robots »*[5].

Cette hypothèse n'est pas farfelue dans le domaine de la santé. Le Règlement de 2016 prévoit en particulier que le traitement des données de santé d'une personne est interdit sauf s'il *« est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée ou d'une autre personne physique, dans le cas où la personne concernée se trouve dans l'incapacité physique ou juridique de donner son consentement »* (art. 9). Deux conditions sont requises pour se passer du consentement de l'individu : ses intérêts vitaux sont menacés, il ne peut donner son consentement pour des raisons soit physiques soit juridiques. Par exemple, une personne âgée est tombée chez elle et ne peut se relever, cette information est transmise à son médecin, à un centre médical… En revanche, les données issues des échanges entre robots communicants sur les risques de santé encourus par la personne ne peuvent être utilisées. Par exemple, dans

[4] https://www.theverge.com/2017/6/6/15747300/apple-safari-ad-tracking-cookie-blocker-google-facebook-privacy

[5] N. Nevejans, in Rapport au nom de L'Office parlementaire d'évaluation des choix scientifiques et technologiques sur Les Robots et la loi, 2016, p. 65.

un futur pas si loin, on peut imaginer la communication entre l'assistant personnel, la balance, le frigo, le tensiomètre, le podomètre qui décèle un comportement alimentaire de la personne et une activité physique la conduisant directement vers un cas d'obésité inquiétant.
Ce futur existe déjà (voir les produits de la gamme Withings). Certains travaillent sur le bouchon connecté pour veiller à ce que chacun boive suffisamment d'eau.
L'assistant centralisant les informations pourrait-il les transmettre au médecin ? à un centre diététique ? à une entreprise de produits amincissant ? à Air France qui, au vu des courbes alarmantes d'augmentation du nombre des personnes en surpoids, pourrait être amenée, peut-être pour des raisons de sécurité, sans doute pour des raisons de confort et finalement d'image de marque, à répartir les voyageurs dans ses avions ? L'assistant pourrait-il, toujours dans l'intérêt de l'assisté, prendre le pas en refusant l'accès au frigo et comme ce dernier sera connecté à un supermarché, en commandant des yaourts 0 % et non plus les crèmes desserts ou du Nutella. Ces robots et ces algorithmes peuvent-ils aller contre la volonté de l'individu pour son bien ? La question a déjà été posée : *« Un robot d'assistance aux soins pourrait-il interdire à une personne de boire un verre d'alcool ou la forcer à marcher, si elle reste allongée toute la journée, au prétexte que le médecin l'a déconseillé ? »*[6]. Le règlement ne le permet pas pour le moment. Mais avant que le droit ne réponde, c'est avant tout à la société d'y réfléchir. C'est une question plus éthique que juridique.

C'est aussi une question économique. L'assistant fonctionne dans le cadre d'une société qui fixe les règles. Il existe déjà aux États-Unis un suivi par objets connectés du comportement à risque de clients. Si ce comportement est avéré par les données remontées, le montant de la police d'assurance est directement réévalué.
Aux États-Unis, l'apprentissage de l'écriture cunéiforme n'est plus une priorité à l'école qui favorise l'usage des claviers. La lecture par des algorithmes de texte écrit au clavier et bien plus facile que celle d'une écriture à la main. Tout est construit pour digitaliser le libre arbitre.

Dans un domaine lié, la protection de la vie privée, les interrogations sont les mêmes. La surveillance de personne par des robots, par exemple des robots de soins, ne va pas sans poser de sérieux doutes quant au respect de leur vie privée. Il a été *proposé « d'instaurer la nécessité d'un consentement à la surveillance de son intimité par un robot de soins personnels, afin qu'un tiers ne puisse porter atteinte à celle-ci que dans des circonstances très précises, c'est-à-dire dans les périodes d'alerte d'urgence »*[7].
Demeure toujours à définir les périmètres de ces « périodes d'alerte d'urgence », ou pour l'application du règlement de 2016, la notion de « sauvegarde des intérêts vitaux ». Par exemple, ne faudrait-il pas contrôler le temps de jeu qui fait basculer l'individu du joueur compulsif au joueur addict et qui peut, de nombreuses études le démontrent, causer des troubles psychologiques importants et même des troubles de la santé ? Selon l'étude menée par MM. Salles et Durain, en 2014, près de 35 millions de Français jouent, dont la moitié à des jeux payants. 4,5 millions seraient spectateurs de compétitions de jeux vidéo et 850 000 seraient des joueurs de jeux vidéo compétitifs[8] dont la plupart sont des MMORPG (Massevely Multiplayer Online Role Playing Game – Jeux de rôle massivement multijoueurs). S'il n'existe pas d'études complètes sur la question de l'addiction à ces jeux, l'INPES appelle toutefois à la prudence[9]. Sans aller jusqu'à adopter un comportement pathologique, le joueur peut simplement céder à une période de « jeu excessif » pouvant engendrer des risques en

[6] N. Nevejans, in Rapport au nom de L'Office parlementaire d'évaluation des choix scientifiques et technologiques sur Les Robots et la loi, 2016, p. 65.
[7] N. Nevejans, Traité de droit et d'éthique de la robotique civile : LEH Edition, 2017, n° 1064, p. 889.
[8] R. Salles et J. Durain, E-Sport. La pratique compétitive du jeu vidéo, mars 2016, p. 2.
[9] http://inpes.santepubliquefrance.fr/10000/themes/addiction-jeux/jeux-video/index.asp

termes financiers (endettement) ou sociaux (rupture du lien familial et/ou scolaire, associabilisation « physique »…). Par le biais de l'algorithme vigilant, pourrait-on envoyer un message d'alerte au médecin ? Aux sites ? À une association de prévention ? Deux conditions sont exigées avons nous vu pour que le consentement du joueur ne soit pas requis : s'il peut y avoir une menace pour les intérêts vitaux de l'individu (1ére condition), la seconde condition est moins évidente : la volonté n'est pas impossible physiquement, mais psychologiquement. Le règlement ne peut donc s'appliquer. Il est peut-être alors regrettable que le règlement n'ait pas intégré cette hypothèse d'empêchement psychologique, car c'est peut-être une addiction à l'alcool, au tabac… qui aurait pu être traitée. Mais resterait la difficulté de déterminer les addictions acceptables ou pas (je bois trop de café ? Je mange trop de chocolat ? Trop de fromages ?...) Mais c'est aussi admettre que l'individu ne peut plus avoir de comportement libre : brûler la vie par les deux bouts ne serait plus possible sans de continuels rappels à l'ordre culpabilisants… Nous arriverions à un monde policé, lisse, sans génie ? Le génie ne s'exprime que dans la liberté et parfois dans l'excès. Or, peut-on inculquer cet état à un algorithme ? Le choix ne serait plus que d'être connecté ou pas ? d'utiliser ces objets connectés ou pas ?

Mais un assureur ne pourrait-il pas l'imposer ou, comme les opérateurs de téléphonie mobile en GB qui proposent des réductions de facture si les abonnés acceptent de recevoir des publicités, un assureur ne serait-il pas tenté établir des conditions tarifaires préférentielles sur les individus qu'il peut surveiller et dont il peut « limiter » les risques ? Est-ce là bien éthique ?

Les Convergences du Droit et du Numérique – 11 au 13 septembre 2017 – Bordeaux.

Interception de données à caractère personnel sur internet à des fins de renseignement.

Olivier DELMAS,
Maître de conférences, LaBRI, Université de Bordeaux.

Maxime KHELOUFI,
Doctorant en droit public, CRDEI – CMRP, Université de Bordeaux.

« Jamais nos concitoyens ne pourront faire l'objet d'un espionnage massif puisqu'ils ne sauraient être suspectés de constituer une menace potentielle pour l'État et, par voie de conséquence, pour eux-mêmes »[1]

C'est en ces termes, a priori éloignés de la philosophie américaine, que la délégation parlementaire au renseignement[2] (ci-après : DPR) vient définir l'esprit général de la politique française du renseignement. Il s'agit en effet de défendre et promouvoir les intérêts fondamentaux de la Nation[3], en recourant à des techniques contingentées et limitées, contre une menace préalablement identifiée et correspondant au cadre légal[4]. C'est précisément sur la condition de l'identification préalable d'une menace que semblent se distinguer les politiques française et américaine du renseignement. Il semblerait en effet que règne, outre-Atlantique, une « idéologie de la capture »[5]. En d'autres termes, il ne s'agit pas de mobiliser des techniques de renseignement à l'encontre de menaces préalablement identifiées mais, au contraire, de déployer un dispositif visant à détecter des menaces. Or cette dimension proactive suppose « un espionnage massif, sans réelle restriction autre que celle induite par les limites technologiques »[6], restriction à laquelle il faudrait sans doute ajouter celles de nature budgétaires.

Lors des débats relatifs à la « loi renseignement » de 2015, le Patriot Act[7], symbole de la philosophie américaine en ce sens qu'il visait à lutter mais aussi – et de façon proactive – déceler le terrorisme, fût régulièrement érigé en « ligne rouge » à ne pas franchir pour les uns[8], tandis que d'autres appelaient de leurs vœux à ce que la France se dote d'un texte comparable[9]. Il faut dire que les questions de renseignement suscitent inquiétudes et passions. Sans doute, le caractère confidentiel de la matière n'est-il pas étranger à la survenance de quelques exagérations ou idées fantasmagoriques. Pour autant, les enjeux sont bien réels. Le renseignement est en effet un puissant outil au service de la prise de

1 Rapport 2014 de la délégation parlementaire au renseignement, Assemblée nationale, document n° 2482 et Sénat, document n° 201 du 18 décembre 2014, p. 71.
2 Cette délégation, commune au Sénat et à l'Assemblée Nationale, se compose de quatre sénateurs et quatre députés. Voir article 6 nonies de l'ordonnance n° 58-1100 du 17 novembre 1958 relative au fonctionnement des assemblées parlementaires, modifiée par la loi n° 2007-1443 du 9 octobre 2007 portant création d'une délégation parlementaire au renseignement.
3 1er Rapport d'activité 2015 – 2016 de la Commission Nationale de Contrôle des Techniques de Renseignement, p. 30.
4 Rapport 2014 de la DPR, *Op. Cit.*, p. 69.
5 CHARDEL P.-A., « Données personnelles et devenir des subjectivités : questions d'éthique », *Sécurité et stratégie*, n° 17, octobre-décembre 2014, p. 7.
6 Rapport 2014 de la DPR, *Op. Cit.*, p. 69.
7 USA PATRIOT Act : Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001, *Loi pour unir et renforcer l'Amérique en fournissant les outils appropriés pour déceler et contrer le terrorisme.*
8 CHARMET-ALIX A., « Le Patriot Act à la française, on n'en est pas loin », *Le Monde Pixels*, [en ligne], 13 janvier 2015, [consulté le 03 septembre 2017], https://www.lemonde.fr
9 LECOQ T., « Les dangers d'un Patriot Act à la française », *Libération*, [en ligne], 19 janvier 2015, [consulté le 03 septembre 2017], http://www.liberation.fr

décision politique. Ici la prise de décision doit s'entendre comme relevant du pouvoir exécutif. Il s'agit en quelque sorte de donner des yeux et des oreilles à un individu qui n'aurait autrement que des bras et des jambes. Aussi puissants soient-ils, ils ne peuvent agir sans indications fiables. Connaître son ennemi pour anticiper ses actions, préserver et promouvoir ses propres intérêts, voici rapidement la raison d'être du renseignement. En fait, le renseignement peut être envisagé comme un cycle. Un besoin particulier est identifié et donne lieu à une prise de décision de la part du pouvoir exécutif visant à la collecte de données susceptibles d'apporter un éclairage à une future prise de décision, ou pouvant venir influencer cette future prise de décision. Ce processus s'autoalimente. Mais la collecte de données ne fournit pas immédiatement un renseignement exploitable. En effet, il convient d'évaluer et traiter ces données afin d'en faire des informations fiables, lesquelles seront elles-mêmes analysées et synthétisées afin de constituer un renseignement utile à la prise de décision politique. C'est ce que décrivent à travers le schéma ci-dessous Eric Boutin et Franck Bulinge[10].

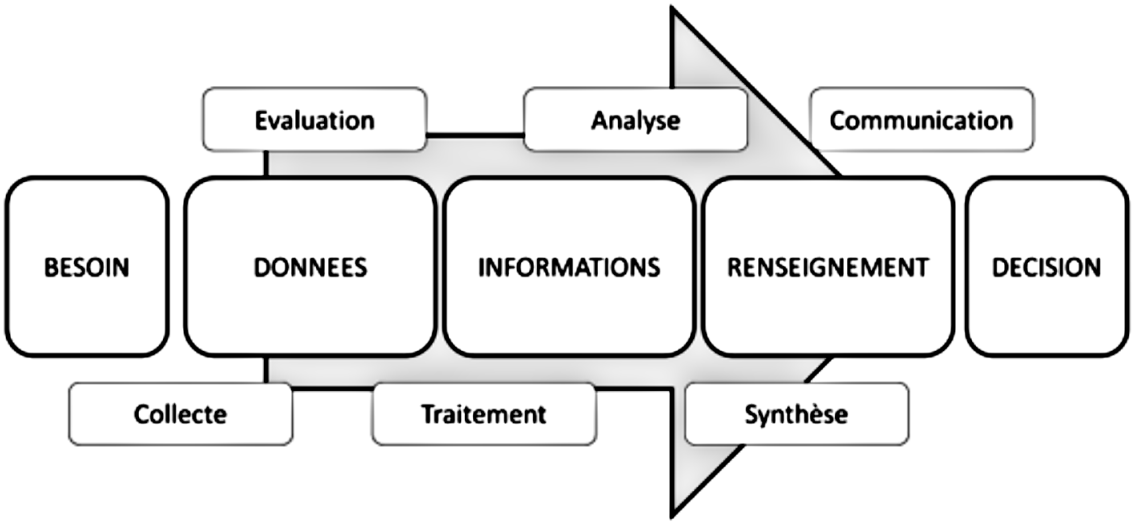


Schéma : Le processus de renseignement d'après Franck Bulinge.

Le renseignement est donc un outil indispensable de nos sociétés démocratiques. Mais il s'agit également d'un puissant outil pouvant faire peser un risque réel sur la protection de nos droits et libertés fondamentaux s'il était utilisé à mauvais escient. Le développement des technologies numériques offrent d'intéressantes perspectives pour la matière du renseignement. Toutefois, au regard du caractère potentiellement très intrusif de ces nouveaux moyens, il devient nécessaire de se doter d'un cadre juridique permettant, dans l'idéal, d'offrir à nos services de renseignement, les outils adaptés à l'exercice de leurs missions, sans pour autant mettre en péril nos valeurs en matière de protection des droits fondamentaux.

C'est dans ce contexte, aggravé par le climat terroriste, que la France s'est dotée en 2015 d'une loi venant définir un cadre légal pour les activités de renseignement[11] et mettant fin à un retard certain puisque nous étions alors, comme le rappelle justement Floran Vadillo, l'une des dernières démocraties occidentales à ne pas avoir défini de cadre juridique pour nos activités de renseignement[12], à l'exception, très ciblée toutefois, de la loi de 1991 relative aux interceptions de sécurité[13], qui avait alors été adoptée sous le coup d'une condamnation de la France par la Cour européenne des droits de l'Homme (ci-après

10 BOUTIN E. et BULINGE F., « Le renseignement comme objet de recherche en SHS : le rôle central des SIC », *Communication et Organisation*, Presses universitaires de Bordeaux, 2015, n° 47, p. 184.

11 Loi n° 2015-912 du 24 juillet 2015 relative au renseignement.

12 VADILLO F., « Une loi relative aux services de renseignement : l'utopie d'une démocratie adulte ? », *Fondation Jean-Jaurès* [en ligne], 2012, note n° 130, p. 1, [consulté le 03 septembre 2017], www.jean-jaures.org

13 Loi n° 91-646 du 10 juillet 1991 relative au secret des correspondances émises par la voie des communications électroniques.

CEDH)[14], et de toute évidence inadaptée au développement des nouvelles technologies. Les exigences de protection posées par la Convention européenne des droits de l'Homme (ci-après ConvEDH) telles qu'interprétées par la CEDH ne sont certainement pas innocentes dans le fait qu'aujourd'hui, le premier article du code de la sécurité intérieure (ci-après : CSI) consacré au renseignement dispose *« Le respect de la vie privée, dans toutes ses composantes, notamment le secret des correspondances, la protection des données personnelles et l'inviolabilité du domicile, est garanti par la loi »*[15]. La question de la protection des données à caractère personnel, c'est-à-dire « *toute information se rapportant à une personne physique identifiée ou identifiable, directement ou indirectement »*[16], est envisagée par le CSI comme une des composantes du droit au respect de la vie privée.

Il s'agit donc de dresser l'inventaire des techniques numériques entrées dans le champ du droit par les lois relatives au renseignement de 2015 en France permettant l'interception de données à caractère personnel sur internet à des fins de renseignement et d'analyser cette convergence entre droit et numérique afin de voir si le *Droit* parvient à réguler les larges possibilités offertes par le *Numérique*.

1. Panorama des techniques de renseignement soumises à autorisation prévues par la loi renseignement de 2015.

La loi de 2015 prévoit quatre catégories de « techniques de recueil de renseignement soumises à autorisation » :

- l'accès administratif aux données de connexion **(1.1.)**
- les interceptions de sécurité **(1.2.)**
- la sonorisation de certains lieux et véhicules et captation d'images et de données informatiques **(1.3.)**
- les mesures de surveillance des communications électroniques internationales **(1.4.)**

Pour chacune de ces catégories, il s'agira de s'intéresser en particulier aux techniques impliquant une interception de données à caractère personnel via le réseau internet. On laissera donc de côté l'interception du contenu des communications qui fait référence à la protection d'une autre liberté fondamentale : celle du secret des correspondances.

1.1. L'accès administratif aux données de connexion.

Les données de connexion sont également connues sous le nom de « métadonnées ». On peut les définir en tant que données sur des données[17]. Si le code de la sécurité intérieure ne pose pas de définition précise, ce que l'on peut regretter, on trouve toutefois à l'article L. 851-1 du CSI une liste de *données de connexion* susceptibles de faire l'objet d'une interception. Néanmoins, les termes utilisés dans cette liste reste eux-même relativement vagues, à l'image de ce qu'il faudrait entendre par *« données techniques relatives à l'identification des numéros d'abonnement ou de connexion à des services de communications électroniques ».* Cette définition englobante a l'avantage de laisser une certaine marge de manœuvre aux services de renseignement lorsqu'ils souhaitent accéder à de telles données.

Si l'on entre dans le détail des techniques pouvant être mobilisées dans le cadre du chapitre relatif à l'accès administratif aux données de connexion, on note que les possibilités sont assez diverses.

14 CEDH, *Kruslin c/ France*, 24 avril 1990, n° 11801/85 et CEDH, *Huvig c/ France*, 24 avril 1990, n° 11105/84.

15 Art. L. 801-1 du code de la sécurité intérieure.

16 Art. 4, 1) du règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (règlement général sur la protection des données).

17 On peut prendre l'exemple d'une conversation vidéo via un réseau social. Les métadonnées ne font pas référence au contenu des échanges au cours de la conversation mais à la durée de la conversation, l'heure à laquelle il s'est déroulé, leur fréquence, …

L'article L. 851-1 du CSI envisage la possibilité d'un accès aux données de connexion en temps différé. Cette technique n'implique pas *d'interception* de données à caractère personnel entendu dans le sens où ces dernières seraient relatives à un contenu circulant d'un point A vers un point B. En revanche, cette technique implique que les données de connexion aient été conservées au préalable par les opérateurs de communications électroniques.

L'article L. 851-2 du CSI envisage lui aussi d'exploiter les données de connexion mais en temps réel cette fois-ci. Ceci implique donc bien l'interception de données de connexion relatives à un contenu circulant d'un point A vers un point B sur un réseau. Il convient de souligner que cette technique, plus intrusive en terme de violation des libertés fondamentales, ne peut être mobilisée que pour la seule finalité de prévention du terrorisme[18]. Cette technique fait, elle aussi, peser une charge sur les opérateurs de communications électroniques puisque ces derniers doivent prévoir les moyens d'intercepter les données de connexion en temps réel.

L'article L.851-3 du CSI prévoit la possibilité d'un traitement de données de connexion par algorithme. Cette disposition de la loi relative au renseignement fût l'objet de nombreux débats en ce sens qu'elle se rapproche de l'idée portée par la philosophie américaine, à savoir mobiliser une technique de renseignement afin de déceler, détecter une menace. Il ne s'agit donc plus de mobiliser une technique à l'encontre d'une personne préalablement identifiée susceptible d'être en lien avec une menace, comme le voudrait la philosophie française.
On constate alors que cette technique s'avère particulièrement intrusive dans le champ des libertés fondamentales. C'est pourquoi le législateur a prévu un certain nombre de garde-fous. Tout d'abord, seule la finalité de prévention du terrorisme peut justifier la mobilisation de cette technique. Ensuite, la première autorisation de mise en œuvre de cette technique ne doit pas pouvoir *« permettre l'identification des personnes auxquelles les informations […] se rapportent »*[19]. Enfin, ce n'est qu'au terme d'une seconde autorisation du premier Ministre, et dans le cas où le premier traitement par algorithme aurait permis de détecter *« des données susceptibles de caractériser l'existence d'une menace terroriste »*[20], que peut être révélée l'identité de la personne en question.
Si cette technique peut sembler séduisante pour anticiper les menaces, il ne faut pas sous estimer les contraintes matérielles de stockage des informations qu'elle présente, ainsi que les contraintes budgétaires qu'elle implique.

L'article L.851-4 du CSI permet l'interception, en temps réel, de données techniques relatives à la localisation d'équipements terminaux en temps réel. Il s'agit ici d'autres données de connexion que celles déjà mentionnées à l'article L. 851-1 du CSI et développées ci-dessus.

L'article L.851-5 du CSI prévoit que *« peut être autorisée l'utilisation d'un dispositif technique permettant la localisation en temps réel d'une personne, d'un véhicule ou d'un objet »*. Nous ne détaillerons pas cette technique en ce sens qu'elle vise essentiellement la pose de balises[21]. Internet n'est donc pas le vecteur central pour la mise en œuvre de cette technique.

L'article L.851-6 du CSI envisage la possibilité de mettre en œuvre un *IMSI catcher*. Il s'agit d'un dispositif technique permettant de capter les données de connexion d'appareils téléphoniques mobiles et, notamment, d'identifier son utilisateur via le numéro de sa carte SIM[22]. Il est bien question ici d'intercepter des données à caractère personnel sur internet puisque, quand bien même il s'agirait de données d'appareils téléphoniques mobiles, circule malgré tout du flux internet via les smartphones. Ce flux peut alors être intercepté par un *IMSI catcher*.

---

18 Concerne la *« personne préalablement identifiée susceptible d'être en lien avec une menace »* mais également *« les personnes appartenant à l'entourage de la personne concernée par l'autorisation », « lorsqu'il existe des raisons sérieuses de penser [que ces personnes] sont susceptibles de fournir des informations au titre de la finalité qui motive l'autorisation ».* Voir Art. L.851-2, I. du CSI.

19 Art. L.851-3, I. al. 2 du CSI.

20 Art. L.851-3, IV. du CSI.

21 CNCTR, *1er Rapport d'activité 2015-2016*, p. 40.

22 CNCTR, *1er Rapport d'activité 2015-2016*, p. 41.

1.2. Les interceptions de sécurité.

Le chapitre relatif aux interceptions de sécurité renvoie aux techniques pouvant être mobilisées afin d'accéder non seulement au contenu de communications mais également aux données de connexion qui y sont associées[23].
L'interception du contenu de communications est particulièrement intrusif dans le droit fondamental au respect de la vie privée puisqu'il viole sa composante relative au secret des correspondances. On peut prendre l'exemple de l'interception d'un courriel. Mais il peut également violer le droit fondamental à la protection des données à caractère personnel puisque certains éléments de ces communications peuvent également renvoyer à de telles données. Pour reprendre l'exemple précédent, le courriel intercepté peut tout à fait présenter des données à caractère personnel permettant d'identifier directement ou indirectement un individu[24]. Les interceptions de sécurité peuvent également porter atteinte à la protection des données à caractère personnel lorsqu'il s'agit d'intercepter les données de connexion relatives aux communications interceptées[25].
Il est à noter que les interceptions de sécurité peuvent également viser l'entourage de la personne visée par la technique de renseignement[26].

Le code de la sécurité intérieure prévoit également que peuvent être mobilisés des dispositifs techniques permettant d'intercepter les correspondances émises. Il s'agit une nouvelle fois des *IMSI catchers*. Néanmoins, compte tenu du caractère particulièrement intrusif de cette technique dans le champ de la protection des droits fondamentaux, ici encore, les finalités permettant de mobiliser cet outil ont été restreintes par le législateur. On retrouve trois finalités[27] :

- L'indépendance nationale, l'intégrité du territoire et la défense nationale ;
- La prévention du terrorisme ;
- La prévention des atteintes à la forme républicaine des institutions ;

1.3. La sonorisation de certains lieux et véhicules et captation d'images et de données informatiques.

Sans possibilité de recourir à une autre technique de renseignement, sauf à porter atteinte à l'efficacité du travail des services de renseignement, le code de la sécurité intérieure envisage que puisse être mobilisée la captation de paroles, d'images[28] ou de données informatiques[29].

La captation de paroles et d'images nécessite souvent la pose de dispositifs techniques et implique donc bien souvent de s'introduire dans un lieu privé[30]. Ceci porte atteinte de façon particulièrement grave au droit à la protection de la vie privée. De son côté, le droit à la protection des données à caractère personnel est ici bien moins concerné. C'est pourquoi nous ne détaillerons pas davantage cette technique de renseignement.
En revanche, ce qui peut se faire pour la captation de paroles et d'images peut également se faire pour la captation de données informatiques. C'est là que le réseau internet peut constituer une plus-value. En effet, il n'est plus systématiquement nécessaire de s'introduire dans un lieu privé pour mettre en place la technique[31]. Nombre de données à caractère personnel peuvent être interceptées directement via le

23 Art. L.852-1 du CSI.
24 On peut penser aux signatures électroniques en fin de courriel ou figurent bien souvent le nom et les coordonnées de l'expéditeur.
25 Voir *supra* point 1.1.
26 Si cet entourage est susceptible de *« fournir des informations au titre de la finalité qui motive l'autorisation »* Voir : Art. L.852-1, I. du CSI.
27 Art.L. 852-1, II. du CSI.
28 Art. L.853-1 du CSI.
29 Art. L.853-2 du CSI.
30 L'article L.853-3 du CSI fait état de la procédure particulière à suivre lorsque le déploiement d'une technique de renseignement implique de s'introduire dans un lieu privé ou un véhicule.

réseau internet. Cette captation de données informatiques peut se faire en temps réel[32] ou sur des données stockées[33]. On bascule donc davantage sur une ingérence dans le droit au respect des données à caractère personnel et, pour ce qui concerne le contenu de certaines données informatiques, sur une composante particulière du droit au respect de la vie privée : le secret des correspondances.

1.4. Les mesures de surveillance des communications électroniques internationales.

Cette technique de renseignement, qui vise à la surveillance des communications émises ou reçues à l'étranger[34], fait l'objet d'un régime juridique spécial[35]. La loi du 24 juillet 2015 relative au renseignement, dont il est question jusqu'à présent, consacrait déjà cette technique de renseignement. Toutefois, le Conseil Constitutionnel, dans sa décision du 23 juillet 2015[36] est venu censurer ce volet au motif que *« en ne définissant dans la loi ni les conditions d'exploitation, de conservation et de destruction des renseignements collectés en application de l'article L.854-1, ni celles du contrôle par la Commission nationale de contrôle des techniques de renseignement de la légalité des autorisations délivrées en application de ce même article et de leurs conditions de mise en œuvre, le législateur n'a pas déterminé les règles concernant les garanties fondamentales accordées aux citoyens pour l'exercice des libertés publiques »*[37].
C'est donc par la loi du 30 novembre 2015[38], déclarée conforme à la Constitution par le Conseil Constitutionnel[39], que le législateur est venu poser un nouveau régime relatif à la surveillance des communications électroniques internationales.

Les finalités permettant de mobiliser cette technique sont nombreuses puisqu'il s'agit de l'ensemble des finalités prévues à l'article L.811-3 du CSI[40].
Par ailleurs, cette technique peut être employée à l'encontre d'individus utilisant des numéros d'abonnement ou des identifiants techniques rattachables au territoire national s'ils *« communiquent depuis l'étranger et […] sont identifiés comme présentant une menace au regard des intérêts fondamentaux de la Nation »*[41]. Il en va de même si un tel individu fait l'objet d'une mesure d'interception de sécurité au moment où il quitte le territoire français.
L'interception des communications électroniques internationales peut viser à la fois le contenu des communications mais aussi les données de connexions qui y sont associées[42]. Ainsi cette technique, peu limitée par les finalités de sa mise en œuvre, peut conduire à ce que nombre de données à caractère personnel soient interceptées sur le réseau internet, y-compris les données de nationaux français.
On note par ailleurs que ce qu'il faut entendre au sens du code de la sécurité intérieure par *« identifiant technique rattachable au territoire national »* est particulièrement vague. Une définition plus précise aurait été la bienvenue. Ici encore, cette formulation laconique a l'avantage d'offrir une marge de manœuvre confortable aux services de renseignement.

---

31 En revanche, s'il est nécessaire de s'introduire dans un lieu privé ou un véhicule, la procédure prévue à l'article L.853-3 du CSI s'applique.

32 Art. L.853-2, I. 2° du CSI : « Peut être autorisée […] l'utilisation de dispositifs techniques permettant d'accéder à des données informatiques, de les enregistrer, de les conserver et de les transmettre, telles qu'elles s'affichent sur un écran pour l'utilisateur d'un système de traitement automatisé de données, telles qu'il les y introduit par saisie de caractères ou telles qu'elles sont reçues et émises par des périphériques audiovisuels ».

33 Art. L.853-2, I. 1° du CSI : « Peut être autorisée […] l'utilisation de dispositifs techniques permettant d'accéder à des données informatiques stockées dans un système informatique, de les enregistrer, de les conserver et de les transmettre ».

34 Art. L.854-1, al. 1er. du CSI.

35 Art. L.854-1, al. 2 du CSI : *« Cette surveillance* [des communications électroniques internationales], *[...] est exclusivement régie par le présent chapitre ».*

36 Cons. Const., décision n° 2015-713 DC, *Loi relative au renseignement*, 23 juillet 2015

37 Cons. Const., décision n° 2015-713 DC, *Loi relative au renseignement*, 23 juillet 2015, cons. 78.

38 Loi n° 2015-1556 du 30 novembre 2015 relative aux mesures de surveillance des communications électroniques internationales.

39 Cons. Const., décision n° 2015-722 DC, *Loi relative aux mesures de surveillance des communications électroniques internationales*, 26 novembre 2015.

40 Voir *infra* point 2.3.

41 Art. L.854-1, al. 3 du CSI.

42 Art. L.854-1, al. 2 du CSI.

Enfin, ce chapitre soulève également la problématique de la territorialité du droit face au réseau internet. Si l'élément territorial est un critère classique d'applicabilité du droit, l'identification des frontières et d'un territoire est parfois très délicate sur internet. A titre d'exemple, un des principes du réseau TOR réside en la possibilité d'émettre une communication[43], sans qu'il ne soit techniquement possible[44] de localiser l'origine de cette communication. Dans ce cas, comment envisager l'applicabilité d'une technique de renseignement, elle même fondée sur une loi dont le champ d'application s'appuie en partie d'un critère territorial ? Cette situation met en évidence une vrai difficulté à concilier efficacement droit et numérique.

2. Les conditions à la mise en œuvre de techniques de renseignement visant l'interception de données à caractère personnel sur internet.

La mise en œuvre de ces techniques est soumises à plusieurs conditions. Tout d'abord, il convient d'en obtenir l'autorisation auprès du premier Ministre **(2.1.)**. Ensuite seuls certains services de renseignement sont habilités à solliciter une telle mise en œuvre **(2.2.)**. Enfin, la mise en œuvre de ces techniques doit répondre à une des nombreuses finalités prévues par le code de la sécurité intérieure **(2.3.)**.

2. 1. L'autorisation de mise en œuvre de techniques de renseignement.

Les quatre techniques de renseignement prévues par le code de la sécurité intérieure sont mises en œuvre sur autorisation du premier Ministre. Ce dernier agit en tant qu'autorité de police administrative. Il convient de distinguer la police administrative ayant une vocation préventive, de la police judiciaire à vocation répressive[45]. La mise en œuvre de ces techniques n'est donc pas soumise à autorisation du juge judiciaire mais relève du pouvoir exécutif lui-même. Ceci ne signifie pas pour autant que la délivrance d'autorisations est exempte de tout contrôle, y compris de contrôle juridictionnel. En effet, la décision du premier Ministre d'autoriser la mise en œuvre d'une technique de renseignement est soumise à l'avis d'une autorité administrative indépendante (ci-après : AAI) : la Commission nationale de contrôle des techniques de renseignement (ci-après : CNCTR)[46], laquelle peut effectuer un contrôle *a priori*[47], c'est à dire avant la mise en œuvre de la technique de renseignement, et *a posteriori*[48], c'est-à-dire après la mise en œuvre de la technique de renseignement avant de rendre son avis.
Le contrôle de la CNCTR porte à la fois sur des éléments de légalité externe[49], mais aussi et surtout, sur des éléments de légalité interne[50].
Lorsque l'avis émis par la CNCTR n'est pas suivi par le premier Ministre, il s'agirait essentiellement de l'hypothèse où le premier Ministre autoriserait la mise en œuvre d'une technique de renseignement soumise à autorisation alors que la CNCTR aurait rendu un avis négatif, le Conseil d'État peut alors être saisi[51]. La juridiction suprême de l'ordre administratif français assure donc elle aussi, à travers une

43 Une description synthétique des propriétés recherchés par le réseau TOR est disponible sur https://www.torproject.org/about/overview.html.en et http://www.bortzmeyer.org/blog-tor-onion.html

44 Attention, cependant aucune technique n'est parfaite et il est quasiment impossible de garantir un anonymat total et en permanence.

45 TRUCHET D., *Droit administratif*, 4° éd., Paris, PUF, 2011, p. 303 et 304.

46 La CNCTR se compose de 4 parlementaires (2 députés, 2 sénateurs), 2 membres du Conseil d'État, 2 magistrats de la Cour de Cassation, 1 personnalité qualifiée. Voir : Art. L.831-1 du CSI.

47 *« La mission de contrôle a priori confiée par la loi à la CNCTR consiste à examiner la légalité des demandes tendant à la mise en œuvre de techniques de renseignement »*. Voir : CNCTR, *1er Rapport d'activité 2015-2016,* p. 63

48 Il appartient à la CNCTR de contrôler l'exécution des techniques autorisées. Voir : CNCTR, *1er Rapport d'activité 2015-2016*, p. 76.

49 *« Compétence de l'auteur de la demande, régularité de la procédure, […] caractère suffisant de la motivation, […] »*. Voir : CNCTR, *1er Rapport d'activité 2015-2016*, p. 63

50 *« Justification de la demande au regard des finalités invoquées, […] proportionnalité des atteintes portées aux droits fondamentaux aux motifs invoqués et aux buts poursuivis, […] »*. Voir : CNCTR, *1er Rapport d'activité 2015-2016*, p. 63

51 Art. L.833-8 du CSI : *« Le Conseil d'État peut être saisi d'un recours [...] soit par le président de la commission lorsque le Premier ministre ne donne pas suite aux avis ou aux recommandations de la commission ou que les suites qui y sont données sont estimées*

formation spécialisée, un contrôle, lequel prend cette fois la forme d'un contrôle juridictionnel. Afin de permettre à la CNCTR et à la formation spécialisée du Conseil d'État d'effectuer un contrôle dans le respect de la protection des éléments classifiés, leurs travaux sont couverts par le secret de la défense nationale[52] et leurs membres ont accès aux éléments classifiés[53].
Au-delà des attributions de la CNCTR visant à vérifier que les techniques de renseignement soumises à autorisation sont bien mises en œuvre dans le respect du cadre juridictionnel dessiné par la loi relative au renseignement de 2015, les individus peuvent, eux- aussi, bénéficier d'une voie de recours afin de vérifier qu'aucune technique n'a été illégalement mise en œuvre à leur encontre. Ils peuvent ainsi, dans un premier temps, saisir la CNCTR, puis, dans un second temps, saisir le Conseil d'État.

Les procédures décrites ci-dessus connaissent toutefois quelques aménagements pour ce qui concerne la surveillance des communications électroniques internationales.
En effet, l'accès aux données de connexion ou au contenu des communication se fait en deux temps. Le premier Ministre doit délivrer, dans un premier temps, une autorisation d'interception des communications[54]. Ce n'est que dans un second temps et après avoir obtenu du premier Ministre une seconde autorisation qu'il est possible d'accéder aux données de connexion[55] ou au contenu des communication[56]. Il est à noter ici que la CNCTR n'émet pas d'avis *a priori*. Ceci pourrait constituer une faiblesse en terme de garanties. Toutefois, la CNCTR reçoit notification de chaque autorisation délivrée en matière d'interception et d'exploitation des communications électroniques internationales[57]. En outre, depuis deux délibérations classifiées adoptées en formation plénière les 28 avril et 19 mai 2016, la CNCTR accepte d'émettre un avis *a priori* en matière de surveillance des communications électroniques internationales[58]. Cette pratique, non prévue par le législateur, est le fruit de la volonté du premier Ministre[59]. Elle s'inscrit par ailleurs dans la pratique mise en œuvre par la Commission nationale de contrôle des interceptions de sécurité (ci-après CNCIS), « ancêtre » de la CNCTR, qui – bien que la loi de 1991[60] ne l'ait pas prévu – rendait elle aussi un avis *a priori* de la mise en œuvre d'écoutes téléphoniques[61].
De même, les voies de recours ouvertes aux individus connaissent ici certains particularismes. En effet, si un individu peut saisir la CNCTR, afin que cette dernière vérifie qu'aucune interception de ses communications électroniques internationales n'a été illégalement conduite, il ne peut pas, sous peine d'irrecevabilité[62], saisir le Conseil d'État dans un second temps, alors que comme cela est pourtant permis pour les autres techniques de renseignement. Seule CNCTR pourra saisir le Conseil d'État au nom de l'individu souhaitant vérifier qu'aucune mesure de surveillance de ses communications électroniques internationales n'a été illégalement mise en œuvre[63].

Ceci étant exposé, reste encore à préciser la liste des services pouvant prétendre à obtenir une autorisation de la part du premier Ministre concernant la mise en œuvre de techniques de renseignement.

---

insuffisantes, soit par au moins trois membres de la commission ».

52 Art. L.832-5 al. 3 et 4 du CSI pour la CNCTR, Art. L.773-2 al. 3 du code de justice administrative (ci-après CJA) pour le Conseil d'État.

53 Art. L.832-5 al. 1er et 2 du CSI pour la CNCTR, Art. L.773-2 al. 3 du CJA pour le Conseil d'État.

54 Art. L.854-2, I. du CSI.

55 L'exploitation des données de connexion peut être menée de façon non individualisée selon les dispositions de l'Art. L.854-2, II. du CSI ou viser une zone géographique, une organisation, un groupe de personne ou un individu particulier selon les dispositions de l'Art. L.854-2, III. du CSI.

56 L'exploitation des communications peut viser une zone géographique, une organisation, un groupe de personne ou un individu particulier selon les dispositions de l'Art. L.854-2, III. du CSI.

57 Art. L.854-9, al. 1er du CSI.

58 CNCTR, *1er Rapport d'activité 2015-2016,* p. 45 et 46.

59 CNCTR, *1er Rapport d'activité 2015-2016,* p. 45.

60 Loi n° 91-646 du 10 juillet 1991 relative au secret des correspondances émises par la voie des communications électroniques.

61 VADILLO F., « Essai de modélisation d'une théorie du contrôle des activités de renseignement », in LAURENT S.-Y. (dir.) et WARUSFEL B. (dir.), *Transformations et réformes de la sécurité et du renseignement en Europe*, Pessac, Presses Universitaires de Bordeaux, 2016, p. 238.

62 CE, *Mme B. A.* 19 octobre 2016, n°397623, cons. n°3.

63 Art. L.854-9, al. 5 du CSI.

2.2. Les services de renseignement habilités à mettre en œuvre les techniques de renseignement soumises à autorisation.

On retrouve l'ensemble des services spécialisés dits du « premier cercle » au sens de l'article L.811-2 du CSI :

- Direction générale de la sécurité extérieure (DGSE)
- Direction de la protection et de la sécurité de la défense (DPSD)
- Direction du renseignement militaire (DRM)
- Direction générale de la sécurité intérieure (DGSI)
- Direction nationale du renseignement et des enquêtes douanières (DNRED)
- Traitement du renseignement et action contre les circuits financiers clandestins (Tracfin)

A l'exception de la DRM et de Tracfin, ces services peuvent potentiellement recourir à l'ensemble des techniques de renseignement soumises à autorisation[64]. Ce n'est pas le cas des services dits du « second cercle »[65] dont la partie réglementaire du CSI fixe les techniques auxquels ils peuvent prétendre[66].

2. 3. Les finalités justifiant la mise en œuvre de techniques de renseignement.

Les finalités justifiant la mise en œuvre de techniques de renseignement sont limitativement énumérées par la loi. Il convient de se référer à l'article L.811-3 du CSI, lequel prévoit sept grandes catégories de finalités :

1) L'indépendance nationale, l'intégrité du territoire et la défense nationale ;
2) Les intérêts majeurs de la politique étrangère, l'exécution des engagements européens et internationaux de la France et la prévention de toute forme d'ingérence étrangère ;
3) Les intérêts économiques, industriels et scientifiques majeurs de la France ;
4) La prévention du terrorisme ;
5) La prévention :
    a) Des atteintes à la forme républicaine des institutions ;
    b) Des actions tendant au maintien ou à la reconstitution de groupements dissous en application de l'article L.212-1 ;
    c) Des violences collectives de nature à porter gravement atteinte à la paix publique ;
6) La prévention de la criminalité et de la délinquance organisées ;
7) La prévention de la prolifération des armes de destruction massive.

Une première remarque conjoncturelle peut immédiatement être dressée. Contrairement à ce que pourrait laisser entendre l'actualité politique et médiatique ayant accompagné l'élaboration de la loi de 2015 relative au renseignement, la question du terrorisme est bien loin d'être la seule raison d'être des activités de renseignement[67]. Les finalités pour la mise en œuvre des techniques de renseignement sont nombreuses, ce qui indique, au passage, la diversité des missions que les services sont susceptibles de devoir assurer. Toutefois, on a pu noter, lors de l'inventaire des techniques de renseignement soumises à autorisation[68], qu'un certain nombre d'entre elles ne peuvent être mobilisées que pour une finalité de prévention du terrorisme[69].

---

64 CNCTR, *1er Rapport d'activité 2015-2016,* p. 32.

65 Pour la liste complète, voir : Décret n° 2015-1639 du 11 décembre 2015 relatif à la désignation des services autres que les services spécialisés de renseignement, autorisés à recourir aux techniques mentionnées au titre V du livre VIII du code de la sécurité intérieure, pris en application de l'article L. 811-4 du code de la sécurité intérieure ou ; CNCTR, *1er Rapport d'activité 2015-2016,* p. 34 et 35.

66 Art. R.851-1 et suivants du CSI.

67 CNCTR, *1er Rapport d'activité 2015-2016,* p. 69, 70 et 73.

68 Voir *supra* point 1.

69 Interception des données de connexion en temps réel, traitement de données de connexion par algorithme.

Une deuxième remarque ayant trait à une question légistique doit ensuite être soulignée. Si la liste de finalité est bien limitative, la dimension à donner à chaque catégorie reste, en l'absence de toute définition précise, relativement hasardeuse. Ceci n'est pas anodin lorsque l'on sait que la mise en œuvre de techniques de renseignement peut être particulièrement intrusive dans le droit fondamental à la protection de ses données à caractère personnel et les autres composantes du droit au respect de la vie privée. L'interprétation de la CNCTR et, en dernier recours, du Conseil d'État sera capitale et bienvenue sur ce point.

Une troisième remarque peut enfin être apportée concernant cette large liste de finalités car, comme nous avons pu le voir[70], bien souvent, les techniques de renseignement les plus intrusives ne peuvent pas être mobilisées pour l'ensemble de ces sept catégories. Par ailleurs, les techniques de *sonorisation de certains lieux et véhicules et captation d'images et de données informatiques* sont conditionnées à une certaine forme de subsidiarité. C'est-à-dire que lorsqu'un service de renseignement souhaite mettre en place une technique soumise à autorisation, le choix doit se porter sur la technique la moins intrusive en terme de protection des droits fondamentaux, sans pour autant que cela ne vienne altérer l'efficacité de leur travail. Ainsi, si à résultat égal, une technique de renseignement soumise à autorisation moins intrusive est envisageable, c'est cette dernière qui doit être privilégiée.

70 Voir *supra* points 1.1. et 1.2.

Comité de pilotage

- Anne CADIOT-FEIDT, Avocate au barreau de Bordeaux, ancienne Bâtonnière de Bordeaux
- Clothilde CAZAMAJOUR, Avocate à la Cour, barreau de Bordeaux
- Jean-François DESRAMÉ, Président du Tribunal administratif de Bordeaux
- Olivier DUBOS, Professeur de droit public, chaire Jean Monnet, Coordonnateur du Forum Montesquieu, université de Bordeaux
- Anne GUÉRIN, Présidente de la Cour administrative d'appel de Bordeaux
- Vivianne LE HAY, Enseignante-chercheur en sociologie au Centre Emile Durkheim, IEP Bordeaux
- Fabrice HOURQUEBIE, Professeur de droit public, directeur de l'Ecole Doctorale de droit, université de Bordeaux
- Valérie MALABAT, Professeur de droit privé et de sciences criminelles, Directrice de l'Institut de sciences criminelles et de la justice, université de Bordeaux
- Isabelle MONTEILS, Sous-directrice, chef du département recherche et documentation, Ecole nationale de la magistrature
- François PELLEGRINI, Professeur d'informatique, Vice-président en charge du numérique, université de Bordeaux
- Benjamin PELLETIER, Directeur exécutif du Forum Montesquieu
- Olivier PUJOLAR, Maitre de conférences en droit privé, Vice-président en charge des partenariats, université de Bordeaux
- Jean-Christophe SAINT PAU, Professeur de droit privé et de sciences criminelles, Doyen de la faculté de droit et de science politique, université de Bordeaux
- Laura SAUTONIE-LAGUIONIE, Professeur de droit privé, Vice-doyen de la Faculté de droit et science politique en charge de la professionnalisation, université de Bordeaux
- Manuel TUNON de LARA, Président de l'université de Bordeaux

Comité scientifique

- Linda ARCELIN, Professeur de droit privé, université de La Rochelle
- Thierry DAUPS, Maître de conférences de droit public, Université Rennes 2
- Gilles GUGLIELMI, Professeur de droit public, université Paris 2 Panthéon–Assas
- Benjamin JEAN, Fondateur du cabinet Inno³
- Bernard LAMON, Avocat au barreau de Rennes
- Daniel LE METAYER, Directeur de recherche, Inria
- Valérie MALABAT, Professeur de droit privé et de sciences criminelles, directrice de l'Institut de sciences criminelles et de la justice, université de Bordeaux
- Nathalie MITTON, Chercheur, Inria
- Nathalie NEVEJANS, Maître de conférences en droit privé, HDR, Université d'Artois
- François PELLEGRINI, Professeur d'informatique, Vice-président en charge du numérique, université de Bordeaux
- Bertrand RIOU, Vice-président au Tribunal administratif de Bordeaux
- Thierry WICKERS, Avocat au barreau de Bordeaux, ancien bâtonnier de Bordeaux