

## DOSSIER

Open data, Open Law : un bouleversement pour la profession ?

# De l'ouverture des données juridiques publiques vers l'avènement des outils de justice prédictive



Par

**Nicolas Herzog**  
Avocat au Barreau  
de Paris – Cabinet  
H2O Avocats –  
[www.h2o-avocats.com](http://www.h2o-avocats.com)

Les données ouvertes (ou « *open data* ») correspondent aux données produites par l'État dont l'accès et l'usage sont laissés libres sans restriction technique, juridique ou financière<sup>1</sup>. Les articles 3, 20 et 21 de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique<sup>2</sup>, dite « loi Lemaire », ont posé le principe de la mise à disposition du public, à titre gratuit et dans un format ouvert, de l'exhaustivité des décisions de justice de l'ordre judiciaire et administratif et le droit de les réutiliser librement par des systèmes de traitement automatisé de données, sous réserve du respect de la vie privée des personnes concernées par les décisions. Cette généralisation de l'ouverture de la jurisprudence (ou « *open case law* ») a pour effet de déplacer encore davantage la valeur des services que les avocats fournissent à leurs clients, de la possession de l'information juridique vers la connaissance et l'analyse du droit et la stratégie qui en découle.

À l'heure de la société de l'information, et de la profusion de celle-ci, plus que jamais, la valeur de l'avocat réside moins dans la possession de l'information juridique que dans la connaissance du droit et sa capacité à l'exploiter pour résoudre des problématiques juridiques de plus en plus complexes, que ce soit dans le cadre de ses missions de conseil ou en matière de contentieux. Après avoir brièvement rappelé l'évolution réglementaire qui a conduit à l'ouverture totale des données juridiques publiques, nous présenterons les

principales applications de justice prédictive que celle-ci va permettre d'offrir aux avocats.

### L'OUVERTURE DES DONNÉES JURIDIQUES PUBLIQUES : UN PROCESSUS QUI A COMMENCÉ EN 1978 ET QUI EST DÉSORMAIS GÉNÉRALISÉ

Le principe de l'ouverture des données trouve son fondement dans l'article 15 de la Déclaration des droits de l'homme et du citoyen selon lequel « la société a le droit de demander compte à tout agent public de son administration ». Ce principe a conduit à considérer que les données publiques constituent un bien commun qui doit être mis à la disposition de tous afin de garantir à la fois les droits des citoyens, mais également la transparence nécessaire dans une société démocratique. C'est en vertu de ce principe que la loi dite CADA<sup>3</sup> a été adoptée en 1978. Celle-ci prévoyait à l'époque un simple droit d'accès indirect des administrés aux seules décisions administratives les concernant.

L'ouverture des données juridiques publiques a progressivement pris son essor dès lors qu'Internet s'est développé auprès du grand public. Ainsi, un arrêté du Premier ministre du 6 juillet 1999 a institué le site internet Legifrance dont l'objet était initialement essentiellement consacré à la diffusion du *Journal*

<sup>1</sup> « Moderniser l'État – L'ouverture des données publiques », <http://www.gouvernement.fr/action/l-ouverture-des-donnees-publiques>

<sup>2</sup> L. n° 2016-1321, 7 oct. 2016, pour une République numérique, dossiers législatifs, <https://www.legifrance.gouv.fr/affichLoiPreparation.do?idDocument=JORFDOLE000031589829&type=general&typeLoi=proj&legislature=14>

<sup>3</sup> L. n° 78-753, 17 juill. 1978, portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal, ayant créé la Commission d'accès aux documents administratifs (CADA), version consolidée au 9 janvier 2017, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000339241>

officiel de la République française<sup>4</sup>. Puis le décret n° 2002-1064 du 7 août 2002<sup>5</sup> pris en application de la loi n° 2000-321 du 12 avril 2000 relative aux droits des citoyens dans leurs relations avec l'administration<sup>6</sup> a créé un service public de la diffusion du droit par l'Internet qui a très largement étendu le périmètre des données juridiques librement accessibles sur le site Legifrance (à savoir ensemble des codes, lois et textes réglementaires en vigueur, traités internationaux, conventions collectives, une sélection de jurisprudence...). L'État a ensuite ouvert d'autres portails internet diffusant des données juridiques. Certains, destinés au grand public comme vie-publique.fr<sup>7</sup> et service-public.fr<sup>8</sup> et d'autres destinés aux professionnels comme le *Bulletin officiel des annonces des marchés publics*<sup>9</sup>, le *Bulletin officiel des annonces civiles et commerciales*<sup>10</sup> ou encore un portail diffusant des informations réglementaires sur les entreprises cotées en Bourse<sup>11</sup>.

Le 6 juin 2005, une ordonnance<sup>12</sup> transposant certaines dispositions de la directive 2003/98/CE du 17 novembre 2003<sup>13</sup> a consacré le principe du droit à la réutilisation des informations publiques. Néanmoins, ce droit de réutilisation n'était pas nécessairement octroyé à titre gratuit. En effet, l'administration pouvait demander le paiement d'une redevance par la conclusion d'une licence de réutilisation, si l'obtention des informations lui occasionnait un coût (recherche, traitement automatisé...), ou si elle souhaitait obtenir une rémunération de ses droits de propriété intellectuelle.

En 2013, c'est au tour du Sénat<sup>14</sup> et de l'Assemblée nationale<sup>15</sup> de publier des données relatives aux travaux législatifs, aux amendements, aux comptes rendus des débats et aux questions des sénateurs et des députés. Un nouveau tournant déterminant a été pris dans l'ouverture des données juridiques publiques le 24 juin 2014 avec l'adoption d'un arrêté consacrant le principe de la gratuité de la réutilisation des bases de données juridiques et associatives de la Direction de l'information légale et administrative (DILA)<sup>16</sup>. Puis, le 19 août 2014, un décret a posé le même principe concernant la réutilisation des informations juridiques issues des bases de données de l'Institut national de la propriété industrielle (INPI)<sup>17</sup>.

La loi Lemaire du 7 octobre 2016 marque le dernier pas de cette évolution en posant le principe de la mise à la disposition du public à titre gratuit et en format ouvert de l'exhaustivité des décisions de justice pro-

noncées par l'ensemble des juridictions de l'ordre judiciaire et administratif et librement réutilisables par des systèmes de traitement automatisé de données. Cette ouverture de la jurisprudence est soumise à une obligation spécifique d'anonymisation afin de veiller au respect du droit à la vie privée des personnes concernées par les décisions de justice. Un décret, dont la publication est envisagée pour le mois de janvier 2017, fixera les conditions d'application de cette mise à disposition de l'ensemble de la jurisprudence française<sup>18</sup>. Cette véritable « avalanche » de données juridiques désormais disponibles, et la possibilité de les réutiliser librement et gratuitement par des logiciels, ouvre la voie au développement des outils de justice prédictive.

## L'AVÈNEMENT DES OUTILS DE JUSTICE PRÉDICTIVE

Si l'État va gratuitement mettre à la disposition de tous les décisions de justice, il n'en demeure pas moins que ces données seront essentiellement brutes et non structurées. Il s'agira de données juridiques de masse (ou « big data ») qui ne pourront pas être traitées avec les outils logiciels classiques. Ainsi, la diversité des sources de données disponibles et leur volumétrie rendent inopérants les outils classiques de recherche d'informations juridiques. Rappelons en effet que, chaque année, ce sont plus de 4 000 000 décisions de justice qui sont rendues en France<sup>19</sup>.

Dans ce contexte d'explosion du volume de données juridiques disponibles, des entreprises intervenant dans le domaine dit de la « legaltech » (entreprises proposant des services juridiques par l'intermédiaire d'outils informatiques) commencent à commercialiser à destination des avocats, et plus généralement des juristes, des solutions logicielles d'intelligence artificielle basées notamment sur le « machine learning » (ou « apprentissage automatique »). Cette technologie permet aux ordinateurs d'apprendre à partir des données qui leur sont soumises, et non plus simplement d'exécuter des algorithmes<sup>20</sup>. L'ouverture des données juridiques fournit ainsi le « carburant », ou encore la matière première essentielle au fonctionnement de ces logiciels. Ces outils vont non seulement permettre aux avocats d'améliorer la pertinence de leurs recherches juridiques, mais aussi et surtout d'affiner la précision de leurs évaluations des risques d'un dossier.

<sup>4</sup> Arr. 6 juill. 1999, relatif à la création du site Internet Legifrance, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000396399&categorieLien=id>

<sup>5</sup> Décr. n° 2002-1064, 7 août 2002, relatif au service public de la diffusion du droit par l'Internet, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000413818&categorieLien=id>

<sup>6</sup> L. 2000-321, 12 avr. 2000, relative aux droits des citoyens dans leurs relations avec les administrations, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000215117>

<sup>7</sup> <http://www.vie-publique.fr/>

<sup>8</sup> <https://www.service-public.fr/>

<sup>9</sup> <http://www.boamp.fr/>

<sup>10</sup> <http://www.bodacc.fr/>

<sup>11</sup> <http://info-financiere.fr/>

<sup>12</sup> Ord. n° 2005-650, 6 juin 2005, relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000629684&categorieLien=id>

<sup>13</sup> Dir. n° 2003/98/CE du Parlement européen et du Conseil, 17 nov. 2003, concernant la réutilisation des informations du secteur public, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000521881>

<sup>14</sup> <http://data.senat.fr/>

<sup>15</sup> <http://data.assemblee-nationale.fr/>

<sup>16</sup> Arr. 24 juin 2014, relatif à la gratuité de la réutilisation des bases de données juridiques et associatives de la direction de l'information légale et administrative, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000029135212&fastPos=1&fastReqId=1203444049&categorieLien=cid&oldAction=rechTexte>

<sup>17</sup> Décr. n° 2014-917, 19 août 2014, relatif à la mise à disposition du public, pour un usage de réutilisation, d'informations publiques issues des bases de données de l'Institut national de la propriété industrielle, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000029378451&categorieLien=id>

<sup>18</sup> Échéancier de mise en application de la loi pour une République numérique, [https://www.legifrance.gouv.fr/affichLoiPubliee.do?sessionId=FEF89F05233456D836E6FE29AE53CBF7pdila14v\\_3?idDocument=JORFDOLE000031589829&type=recheancier&typeLoi=&legislature=14](https://www.legifrance.gouv.fr/affichLoiPubliee.do?sessionId=FEF89F05233456D836E6FE29AE53CBF7pdila14v_3?idDocument=JORFDOLE000031589829&type=recheancier&typeLoi=&legislature=14)

<sup>19</sup> « Les chiffres-clés de la Justice », éd. 2016, <http://www.justice.gouv.fr/budget-et-statistiques-10054/chiffres-clés-de-la-justice-10303/les-chiffres-clés-de-la-justice-edition-2016-29359.html>

<sup>20</sup> « Machine learning : décryptage d'une technologie qui monte », JDN.com, 10 avr. 2015, <http://www.journaldunet.com/solutions/analytics/machine-learning.shtml>

Les avocats pourront utiliser ces outils afin de :

- procéder à l'estimation du montant de l'indemnisation que leurs clients sont susceptibles de se voir allouer par une juridiction ;
- établir des probabilités sur les chances de gagner une procédure judiciaire ;
- répondre à une question juridique en croisant des informations jurisprudentielles et textuelles ;
- effectuer des prévisions sur les éléments de fait ou de droit qui seront susceptibles de convaincre une juridiction ;
- dresser une cartographie des juridictions les plus favorables selon la problématique rencontrée ;
- profiler les magistrats et même les confrères adverses en fonction des décisions rendues et des résultats obtenus.

Des éditeurs de documentation juridique comme Dalloz<sup>21</sup> ou LexisNexis<sup>22</sup>, mais également des *start-up* telles que Doctrine<sup>23</sup>, Predictice<sup>24</sup>, Case Law Analytics<sup>25</sup>, Tyr-legal<sup>26</sup> ou encore Supra Legem<sup>27</sup> commencent à proposer des solutions à destination des professionnels du droit. Ces outils dits « de justice prédictive », dont la pertinence ne va cesser de croître en même temps que l'augmentation du volume de données disponibles, vont inéluctablement conduire les avocats à avoir une vision plus mathématique de leur métier en incluant les statistiques et les probabilités dans leurs analyses et leurs stratégies de résolution des problématiques juridiques. Les avocats vont ainsi gagner en productivité grâce à l'utilisation de moteurs de recherche juridique d'une puissance inédite et permettre à leurs clients de mieux maîtriser leurs risques en les évaluant plus finement au moyen de statistiques et de probabilités.

Mais, attention, ne nous y trompons pas, ces outils ont vocation à assister les avocats, et non à se substituer à eux. Les conseils d'un avocat prennent en compte des facteurs bien plus subjectifs que les mathématiques appliquées, comme par exemple la psychologie. Par ailleurs, le contexte factuel d'une affaire est par essence unique, chaque détail étant susceptible d'influencer le juge à le qualifier juridiquement dans un sens ou dans un autre.

En outre, les règles régissant les relations humaines sont évolutives et derrière la plupart des revirements de jurisprudence se cache le travail créatif et collaboratif de l'avocat avec son client. C'est pourquoi les outils de justice prédictive seront d'une remarquable aide au travail d'analyse de l'avocat, mais ne se substitueront pas à son rôle de conseil, qui est de permettre à son client de prendre une décision éclairée, en toute connaissance des risques que celle-ci comporte. L'avènement du robot-avocat qui concurrencera, ou même se substituera, à l'humain avocat demeure à ce jour de l'anticipation ou même de la science-fiction.

Rappelons en effet qu'actuellement, et même si cela est une prouesse vertigineuse, la machine n'a réussi à battre l'humain que dans des activités exclusivement mathématiques comme le jeu d'échecs ou le jeu de go, et non dans des activités comportant du hasard, de l'inconnu et de la psychologie comme le poker par exemple<sup>28</sup>. Pour rester dans le présent, il est indéniable que l'avènement des outils de justice prédictive devrait permettre aux avocats de minimiser les coûts, la durée (en facilitant notamment les accords amiables) et les aléas d'une procédure judiciaire, et plus généralement d'augmenter la valeur ajoutée des conseils et des solutions qu'ils fournissent à leurs clients.

<sup>21</sup> Les éditions Dalloz proposent ainsi un service dénommé « Jurisprudence chiffrée » permettant de générer une liste de décisions rendues et d'obtenir ainsi une information chiffrée sur le montant des indemnités allouées par les tribunaux dans un cas d'espèce similaire à celui soumis à l'outil. Ce service est accessible à l'adresse suivante : <http://jurisprudencechiffree.dalloz.fr/>

<sup>22</sup> Le service « Données quantifiées JurisData » édité par LexisNexis est accessible à l'adresse suivante : [http://www.lexisnexis.fr/metiers/avocats/solutions/recherche\\_information/Donnees-Quantifiees-JurisData/index.html](http://www.lexisnexis.fr/metiers/avocats/solutions/recherche_information/Donnees-Quantifiees-JurisData/index.html)

<sup>23</sup> Le site Doctrine est accessible à l'adresse suivante : <https://www.doctrine.fr/>

<sup>24</sup> Le site Predictice est accessible à l'adresse suivante : <https://predictice.com/>

<sup>25</sup> Le site Case Law Analytics est accessible à l'adresse suivante : <http://caselawanalytics.com>

<sup>26</sup> Le site Tyr-legal est accessible à l'adresse suivante : <http://www.tyr-legal.fr/>

<sup>27</sup> Le site Supra Legem est accessible à l'adresse suivante : <http://www.supralegem.fr/>

<sup>28</sup> « Un ordinateur défie l'être humain au poker », LeFigaro.fr, 6 mai 2015, <http://www.lefigaro.fr/sciences/2015/05/06/01008-20150506ART-FIG00301-un-ordinateur-defie-l-etre-humain-au-poker.php>