



---

**CENTRE PERELMAN  
DE PHILOSOPHIE DU DROIT**

**L'INTELLIGENCE ARTIFICIELLE EN  
DROIT : MYTHES, RÉALITÉS ET  
CONSÉQUENCES.**

**Travail de fin d'étude réalisé en vue de l'obtention du grade de master  
en droit public et international**

**Par William-James KETTLEWELL**



[www.philodroit.be](http://www.philodroit.be)

# Table des matières

<b>1. Introduction .....</b>	<b>3</b>
<b>2. Etat de l'art : le droit saisi par l'algorithmique .....</b>	<b>7</b>
2.1. <i>Inventer l'intelligence artificielle en droit : la recherche académique.....</i>	7
2.1.1. Qu'est-ce que le domaine de recherche AI & Law ? .....	8
2.1.2. Les systèmes-experts : le paradigme initial.....	10
2.1.3. Le nouveau paradigme : le machine learning.....	13
2.1.4. Conclusion concernant la technique : machine learning et big data .....	15
2.2. <i>Utiliser l'intelligence artificielle en droit : le marché des legaltechs.....</i>	17
2.2.1. Une grande diversité de l'offre.....	17
2.2.2. Quelques tendances générales .....	18
2.2.3. Un marché en évolution .....	20
2.2.4. La réalité des acteurs de la nouvelle vague .....	21
2.2.5. Conclusion concernant le marché des legaltechs .....	25
<b>3. La portée de la justice analytique : implémentation, pratiques et futur du droit... 26</b>	
3.1. <i>Analyse de l'implémentation technique.....</i>	27
3.1.1. Esquisse des principes du machine learning supervisé .....	28
3.1.2. Illustrations d'applications utilisant du machine learning.....	30
3.1.3. Commentaires sur les données .....	35
3.1.4. Portée concrète de l'implémentation sur le droit et le monde juridique .....	36
3.2. <i>Analyse de l'utilisation par les acteurs du droit .....</i>	39
3.2.1. Le point de vue de la sphère politique.....	40
3.2.1.1. Une réification de « La Jurisprudence ».....	41
3.2.1.2. De nouvelles interactions entre pouvoirs .....	42
3.2.2. Le point de vue des milieux du droit.....	44
3.2.2.1. Juges, justice et le poids de la comparaison. ....	44
3.2.2.2. De nouvelles pratiques chez les avocats.....	46
3.2.2.3. Nouveaux objets et nouveaux outils pour le monde académique .....	49
3.2.3. Le point de vue des utilisateurs du droit, particuliers et entreprises .....	49
3.2.3.1. La performativité et l'impact de l'implémentation .....	50

3.2.3.2. Une vision actuarielle du droit .....	51
3.3. <i>La justice analytique et le futur du droit</i> .....	52
<b>4. Conclusion</b> .....	<b>54</b>
<b>5. Bibliographie</b> .....	<b>56</b>
5.1. <i>Introduction</i> .....	56
5.2. <i>Inventer l'intelligence artificielle</i> .....	56
5.3. <i>Utiliser l'intelligence artificielle</i> .....	58
5.4. <i>Analyse de l'implémentation</i> .....	59
5.5. <i>Analyse de l'utilisation</i> .....	60
5.6. <i>Futur du droit</i> .....	61

## 1. Introduction

Depuis une dizaine d'années, la rapidité avec laquelle les avancées scientifiques ont permis à des machines de résoudre des problèmes jusqu'alors exclusivement accessibles à l'intelligence humaine est difficile à croire pour quiconque ne suit pas attentivement cette évolution. Ainsi, l'« intelligence artificielle » semble désormais être devenue capable d'accomplir des tâches motrices fines de nature générale<sup>1</sup>, de conduire une voiture en conditions réelles de manière autonome<sup>2</sup>, de maîtriser le langage humain naturel pour battre les meilleurs joueurs du monde à *Jeopardy*<sup>3</sup> ou même de composer de la musique indiscernable d'une création humaine<sup>4</sup>. Etant donné que les recherches dans la formalisation du droit et dans les liens entre droit et informatique<sup>5</sup> ont précédé ou accompagné ces changements, il n'est pas étonnant que ces domaines aient également été sujet à d'importantes évolutions.

A la fois en parallèle et au vu de ces progrès, une « bulle » médiatique s'est progressivement formée autour de la prédiction de bouleversements gigantesques et imminents. Ceux-ci concerneraient presque tous les domaines de l'économie et de la société, en ce compris le monde juridique. Des exemples tels que « *Des robots juges pourraient bientôt aider avec les affaires des tribunaux* »<sup>6</sup>, « *Les robots seront-ils vraiment les avocats de demain ?* »<sup>7</sup>, « *Une IA capable de prédire le sort d'une requête devant la CEDH* »<sup>8</sup>, « *L'intelligence artificielle, l'avocat et le*

---

<sup>1</sup> Voyez, par exemple, W., KNIGHT, « This Robot Could Transform Manufacturing », *MIT Technology Review*, 2012. Dernière consultation le 11 avril 2017, en ligne, à l'adresse <https://www.technologyreview.com/s/429248/this-robot-could-transform-manufacturing/#comments>.

<sup>2</sup> Voyez, par exemple, M., WALDROP, « Autonomous vehicles: No drivers required », *Nature*, Vol. 518, N°7537, 2015, pp. 20-23.

<sup>3</sup> Voyez, par exemple, S., BAKER, *Final Jeopardy: The Story of Watson, the Computer that Will Transform Our World*, Boston, Mariner Books, 2012

<sup>4</sup> Voyez, par exemple, P., MISHRA *et al.*, « Of Art and Algorithms: Rethinking Technology & Creativity in the 21st Century », *TechTrends*, 2013, Vol. 57 N°3, pp. 10-14.

<sup>5</sup> Voyez la revue *Artificial Intelligence and Law* de Springer ou l'un des ouvrages suivants: A., NARAYANAN, *Law, Computer Science, and Artificial Intelligence*, Chicago, Intellect, 1998 ; A., SILVERMAN, *Mind, machine, & metaphor: an essay on artificial intelligence and legal reasoning*, Colorado, Westview press, 1993 ; G., SARTOR, et BRANTING, L., « Judicial Applications of Artificial Intelligence », Springer Netherlands, 1998 ;

<sup>6</sup> Traduction personnelle du titre de l'article de presse « *Robots judges could soon be helping with court case* », paru en ligne le 24 octobre 2016 à l'adresse : <https://www.independent.co.uk/life-style/gadgets-and-tech/news/ai-judge-robot-european-court-of-human-rights-law-verdicts-artificial-intelligence-a7377351.html>, consulté pour la dernière fois le 28 mars 2018.

<sup>7</sup> Titre de l'article de presse paru, en ligne, le 10 janvier 2017 à l'adresse : <https://www.lesechos.fr/10/01/2017/LesEchos/22358-045-ECH-les-robots-seront-ils-vraiment-les-avocats-de-demain--.htm>, dernière consultation le 28 mars 2018

<sup>8</sup> Titre de l'article de presse paru, en ligne, le 25 octobre 2016 à l'adresse : <https://www.nextinpact.com/news/101870-une-ia-capable-predire-recevabilite-devant-cedh.htm>, dernière consultation le 4 avril 2018. CEDH est ici, bien sûr, l'acronyme de « Cour européenne des droits de l'Homme ».

*juge* »<sup>9</sup>, « *Quand les algorithmes entrent dans les prétoires* »<sup>10</sup>, « *L'aube de l'intelligence artificielle marque-t-elle le crépuscule du monde du droit ?* »<sup>11</sup>, illustrant l'exorbitance de certaines affirmations, ne constituent qu'un échantillon des titres accrocheurs d'articles de presse parus récemment. . Souvent, ceux-ci se reposent sur une compréhension et une analyse bien mince des phénomènes en jeu, particulièrement dans le cadre du traitement de ce qui est fréquemment dénommé la « justice prédictive »<sup>12</sup>. Malgré cela, comme le fait remarquer Antoine Garapon<sup>13</sup> (magistrat français spécialiste de ces questions), la prolepse, une figure de style consistant à parler au présent de ce qui va nécessairement advenir, est très souvent utilisée dans ce genre d'articles, les auteurs parlant alors tels des prophètes prédisant le futur.

Toutefois, il est vrai que de nombreux domaines de la société subissent des transformations sous l'influence de l'évolution technologique. Le droit, dans sa dimension économique, sociale et politique, n'y fait effectivement pas exception. Or, une grande partie du discours médiatique sur le sujet<sup>14</sup> reflète souvent une image déformée des évolutions technologiques d'intérêts ainsi que de l'impact concret de celles-ci sur la sphère juridique. En effet, tout d'abord, ce discours se focalise presque exclusivement sur « l'intelligence artificielle », raccourci souvent approximatif qui recouvre parfois des techniques variées même si souvent liées à l'apparente faculté de « prendre des décisions » de manière autonome. Ensuite, les prouesses des nouvelles technologies sont souvent présentées sans contexte, sur base d'essais ou d'affaires particulières, rendant approximative ou impossible l'évaluation des performances réelles et l'imminence de leurs arrivées sur le marché. Enfin, l'absence presque totale d'un traitement technique des méthodes utilisées pour produire lesdites prouesses conduit à un autre type d'erreur : la méconnaissance de l'impact du fonctionnement de ces technologies, souvent présumé « neutre », sur le domaine du praticien.

---

<sup>9</sup> Titre de l'article de presse paru, en ligne, le 10 octobre 2017 à l'adresse : <https://theconversation.com/lintelligence-artificielle-lavocat-et-le-juge-85426>, dernière consultation le 2 avril 2018.

<sup>10</sup> Titre de l'article de presse paru, en ligne, le 22 mars 2017 à l'adresse : <https://www.lesechos.fr/22/03/2017/LesEchos/22409-043-ECH-quand-les-algorithmes-entrent-dans-les-pretoires.htm>, dernière consultation le 28 mars 2018.

<sup>11</sup>Titre de l'article de presse paru, en ligne, le 22 mars 2018 à l'adresse : <https://www.lemondedudroit.fr/decryptages/56842-aube-intelligence-artificielle-marque-elle-crepuscule-monde-droit.html>, dernière consultation le 2 avril 2018.

<sup>12</sup> Nous verrons que la dénomination elle-même du phénomène est en jeu et fait débat.

<sup>13</sup> Propos tenus lors d'une conférence organisée par le Centre de droit public de l'ULB le 22 février 2018.

<sup>14</sup> Nous ne prétendons pas ici qu'il n'existe aucune publication tentant de nuancer les propos tenus dans d'autres journaux ou qu'aucun auteur n'a jamais critiqué certains excès journalistiques dans ce domaine. Au contraire, il est possible de trouver ce genre de critique, par exemple dans cet article de presse publié, en ligne, le 27 mars 2018 à l'adresse : <https://www.lecho.be/opinions/carte-blanche/droit-et-intelligence-artificielle/9995832.html>. Le discours majoritaire se dégageant de la lecture de médias « grands publics » ou à destination des juristes reste néanmoins peu éclairant quant aux tenants et aboutissants concrets des évolutions en cours.

Sur base de ces constats, un juriste curieux pourrait légitimement se poser la question de la réalité que cachent de telles hyperboles médiatiques ? Quel impact concret auront l'implémentation, la dissémination et l'utilisation de « l'intelligence artificielle en droit » sur le monde juridique et le droit lui-même ?

C'est à cette question que la présente étude entend apporter une réponse. Toutefois, de prime abord, cette question semble trop ambiguë pour faire l'objet d'un traitement analytique sérieux. En effet, nous venons à l'instant de critiquer l'absence d'une définition rigoureuse de « l'intelligence artificielle » par les médias. Cette ambiguïté est en réalité volontaire. En effet, lever celle-ci nous permettra de préciser l'objet réel, si souvent flou voire caché, du discours médiatique sur « l'intelligence artificielle en droit » participant ainsi à l'un des l'objectifs de notre travail: la démythification d'un pan de l'ingénierie informatique à destination du juriste.

Pour ce faire, nous utiliserons une approche pragmatique du droit global et emprunterons donc sa méthodologie<sup>15</sup>. En effet, la montée des nouvelles technologies et leur impact sur le monde juridique *est* un phénomène global, tout comme l'est la recherche en science informatique. De plus, et encore plus à propos, le domaine du droit global tente de comprendre et conceptualiser l'impact de la mondialisation (dont l'évolution technique fait partie) sur le droit et de penser différents types de nouveaux objets juridiques (les « objets juridiques non-identifiés »<sup>16</sup>). Il semble donc être un cadre théorique adéquat pour tenter d'aborder l'impact des nouvelles technologies sur le monde juridique *et* sur le droit lui-même. Car, en effet, la première hypothèse de ce travail est que l'implémentation technique, c'est-à-dire le fonctionnement, de « l'intelligence artificielle en droit » ne pourra pas rester neutre vis-à-vis de la substance et la pratique du droit positif. En cela, sans être en elle-même normative, elle devient donc un objet que la doctrine juridique se devra d'investiguer. Notre deuxième hypothèse postule que l'impact des technologies ne consistera pas en une révolution des pratiques, mais se présentera plutôt comme une reconfiguration (c'est-à-dire un renforcement, une modification ou un approfondissement) de pratiques déjà existantes, à cause des nouvelles possibilités ouvertes par les technologies, et autour de celles-ci.

La méthodologie du droit global recommande de procéder par l'étude de cas et de se garder

---

<sup>15</sup> B., FRYDMAN, *Comment penser le droit global?*, Working Papers du Centre Perelman de Philosophie du Droit, Vol.1, 2012. Publié en ligne à l'adresse : [http://www.philodroit.be/IMG/pdf/comment\\_penser\\_le\\_droit\\_global\\_2011.pdf](http://www.philodroit.be/IMG/pdf/comment_penser_le_droit_global_2011.pdf). Dernière consultation le 20 mars 2018. Pour la partie spécifiquement sur la méthodologie, voyez les pages 5-10.

<sup>16</sup> *Ibid*, *spec.* p. 5.

d'en faire une étude trop abstraite. Elle nous invite à procéder à notre investigation au plus près « du terrain » et à multiplier les angles d'approche pour ensuite seulement procéder à un traitement analytique du sujet<sup>17</sup>.

En suivant ces directives, la réponse à notre question de recherche se construira en deux temps. Le premier temps consistera dans une investigation de deux domaines liés à l'intelligence artificielle et aura comme objectif la construction de notre objet d'étude, c'est-à-dire, la levée de l'ambiguïté de la question concernant la nature de « l'intelligence artificielle en droit ». Pour se garder d'aborder le sujet de manière trop abstraite, nous ne partirons pas d'une définition théorique du concept « d'intelligence artificielle ». Au contraire, nous investiguerons, d'une part, la recherche académique à l'intersection de l'intelligence artificielle et du droit et, d'autre part, l'évolution du marché des nouveaux outils juridiques et des entreprises qui les commercialisent. Lors de cette investigation nous mobiliserons donc une approche « informatique » du droit (envisageant le droit comme une « méthode automatisable ») ainsi qu'une approche économique du droit (envisageant le droit comme un « service consommable »). Cette dernière investigation sera accompagnée d'une « enquête de terrain » constituée de diverses interviews de fondateur ou directeurs desdites entreprises ce qui nous permettra de confirmer les études de marchés et d'en approfondir certains aspects.

Bien que l'objectif de ces investigations soit de déterminer les techniques et les pratiques à l'origine du déclenchement de la bulle médiatique, et d'obtenir une compréhension concrète, technique et approfondie du phénomène supposément récent que constitue « l'arrivée de l'intelligence artificielle dans le monde juridique », il ne sera pas procédé à une analyse des articles de presse en eux-mêmes. Au contraire, une meilleure compréhension du contenu de ces évolutions se doit de procéder au plus près des domaines évoluant eux-mêmes, plutôt que de se reposer sur une analyse des rapports journalistiques les concernant. Nous y ferons toutefois parfois référence afin de nous aiguiller dans notre recherche.

Dans un deuxième temps, nous présenterons, détaillerons et analyserons les caractéristiques techniques de ces technologies et de ces pratiques concrètes pour en déduire les conséquences sur le monde juridique et la substance du droit. Nous analyserons ensuite les conséquences de leurs utilisations par les acteurs classiques du droit. Enfin, nous mentionnerons brièvement les liens potentiels entre ces techniques et un autre grand courant d'évolution du droit, pour finir par indiquer quelques pistes d'approfondissement de nos recherches.

---

<sup>17</sup> *Ibid.*

## 2. Etat de l'art : le droit saisi par l'algorithmique

C'est par l'étude de ces deux facettes de l'évolution des technologies, académique et technique d'un côté, économique et pratique de l'autre, que nous commencerons notre investigation.

Tout d'abord, en abordant la recherche académique pour mieux comprendre l'intelligence artificielle en droit, nous présenterons le champ de recherche au cœur de ces questions dénommé « Intelligence Artificielle et Droit » (plus connu sous le nom anglo-saxon de « *AI & Law* », utilisé par la suite). Nous mentionnerons brièvement son histoire, les problèmes qu'il aborde et les modèles informatiques du droit qu'il a développé. Nous présenterons ensuite les méthodes appliquées à la résolution de problèmes juridiques en fonction des deux paradigmes dominants du domaine : les « systèmes-experts » et le « *machine learning* ». Enfin, nous conclurons de l'investigation de ce domaine de recherche le type de techniques algorithmiques<sup>18</sup> qu'il nous faudra analyser.

Ensuite, nous présenterons une synthèse de l'état et l'évolution du marché des « *legaltechs* » européennes et américaines. Les *legaltechs* sont des entreprises, souvent jeunes, censées offrir de nouveaux outils technologiques (ou services faisant usage de ceux-ci) sur le marché juridique. Nous nous appuyerons également, ce faisant, sur plusieurs interviews de directeurs ou fondateurs de *legaltechs*, principalement (mais pas exclusivement) celles faisant usage du *machine learning*. Nous dégagerons, ce faisant, les pratiques au cœur du phénomène que nous cherchons à étudier.

### 2.1. Inventer l'intelligence artificielle en droit : la recherche académique.

Contrairement à ce que peuvent laisser penser certains médias, les tentatives d'application des techniques de la science informatique au domaine du droit sont loin d'être neuves. Au contraire, dès la fin des années cinquante<sup>19</sup>, l'objectif d'automatiser certains pans du raisonnement humain, et plus particulièrement du raisonnement juridique, était présent dans la communauté scientifique. Cet objectif a amené de nombreuses évolutions aussi bien dans le domaine de

---

<sup>18</sup> La définition de « algorithmique » du dictionnaire Larousse est « [q]ui concerne les algorithmes, peut être exprimé par un algorithme ». Un algorithme se définit comme un « [e]nsemble de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations ». Voyez le site internet du dictionnaire Larousse aux adresses <http://www.larousse.fr/dictionnaires/francais/algorithmique/2239> et <http://www.larousse.fr/dictionnaires/francais/algorithme/2238>, dernières consultations le 5 avril 2018.

<sup>19</sup> De nombreux auteurs considèrent que la naissance de cette ambition coïncide avec la publication de l'article L., MEHL, « Automation in the Legal World », présenté à la conférence *Mechanisation of Thought Processes*, Vol. 2, 1959, pp. 755-787. En effet, celui-ci aborde pour la première fois l'idée d'utiliser la logique formelle à des fins d'extraction d'informations juridiques et d'inférences.



recherche générale sur l'intelligence artificielle que dans le domaine plus restreint de l'application de l'intelligence artificielle au droit : *AI & Law*.

### 2.1.1. *Qu'est-ce que le domaine de recherche AI & Law ?*

Il fallut toutefois attendre les années quatre-vingt<sup>20</sup> pour voir le pan juridique de cette ambition donner naissance à un domaine de recherche entier, dénommé habituellement *AI & Law*, accompagné de sa propre communauté, de ses propres conférences internationales et de ses propres journaux.

Ce domaine de recherche est vaste. En effet, il englobe<sup>21</sup> tout ou partie de sujets aussi variés que la théorie formelle des systèmes normatifs, la logique informatico-juridique, les représentations des systèmes d'arguments juridiques, l'ontologie juridique ou la modélisation formelle d'institutions. Bien qu'il soit généralement considéré comme un domaine de recherche extérieur à la recherche proprement juridique, certaines contributions publiées dans les livres ou journaux qui y sont liés abordent des questions profondes sur la nature du droit et sont fréquemment signées de la main de juristes et de philosophes du droit<sup>22</sup>.

Aucune synthèse de quelques lignes ne saurait fidèlement présenter la substance ou l'évolution d'un tel champ de recherche<sup>23</sup>. Toutefois, en simplifiant, il est possible d'affirmer que l'un des objectifs principaux des chercheurs dans ce domaine est le développement d'un modèle informatique du raisonnement juridique<sup>24</sup> qui permettrait à la fois d'aider dans la résolution justifiée et expliquée de problèmes juridiques et de prédire les résultats de ceux-ci<sup>25</sup>. Pour ce faire, le problème est décomposé en différents sous-problèmes, intéressants et utiles en eux-mêmes, qui se sont développés comme des sous-champs de recherches bien établis : comment

---

<sup>20</sup> La tenue la première « *International Conference on AI & Law* » en 1987 est souvent présentée comme le point de départ réel de la communauté de recherche *AI & Law*. En effet, même si depuis 1959 différents travaux publiés peuvent être rattachés à ce domaine, ceux qui la précèdent constituaient souvent des efforts isolés et non la participation à une communauté active. Sur ce point voyez T., BENCH-CAPON *et al.*, « A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law », *Artificial Intelligence and Law Journal*, September, Vol. 20, N°3, 2012, pp. 1-101, spec. pp. 2-4.

<sup>21</sup> Voyez G., SARTOR *et al.*, ROTOLO, « AI and Law », *Agreement technologies*, P. Casanovas, G., Sartor, dir., Londres, Springer, 2013, pp. 199-207, spec. pp. 199-200.

<sup>22</sup> Par exemple, voyez l'article de M., FERNÁNDEZ-BARRERA *et G.*, SARTOR, « The Legal Theory Perspective: Doctrinal Conceptual Systems vs. Computational Ontologies », *Approaches to legal ontologies: theories, domains, methodologies*, P. Casanovas, G., Sartor, dir., Londres, Springer, pp. 15-23.

<sup>23</sup> Nous renvoyons donc le lecteur intéressé par un approfondissement vers les références suivantes : E., RISSLAND *et al.*, « AI and Law: A fruitful synergy », *Artificial intelligence*, Vol. 150, N°1, 2003, pp. 1-15 ainsi que T., BENCH-CAPON *et al.*, *op. cit.*

<sup>24</sup> La langue véhiculaire de ce domaine étant l'anglais, ce nom est traduit de « *Computational Model of Legal Reasoning* » (CMLR).

<sup>25</sup> Voyez G., SARTOR *et al.*, ROTOLO, *op. cit.*, spec. p 200.

modéliser l'information comprise dans une norme<sup>26</sup> (telle qu'une loi ou un règlement) et parvenir à son application à des faits ? Comment modéliser l'information contenue dans un jugement<sup>27</sup> ? Comment extraire l'information de textes juridiques<sup>28</sup> ? Est-il possible de modéliser un argument juridique et les interactions entre différents arguments<sup>29</sup> ? Comment utiliser ces modèles pour prédire l'issue d'un litige ?

Pour résoudre ces sous-problèmes, les chercheurs développèrent, au fil des années de recherche, diverses techniques de représentation du droit : formalisation du raisonnement juridique à base de logique des prédicats et de logique floue, analogies entre deux cas pratiques réalisées sur base de leur « dimensionnalité » et de leurs « facteurs légaux », élaboration de schémas d'argumentations « calculables » ou encore d'ontologies légales et déontiques permettant de représenter informatiquement les éléments constitutifs du droit (tel que les droits et obligations). Ces modèles, dont nous n'avons malheureusement pas le temps d'aborder ici les détails techniques<sup>30</sup>, permettent à un ordinateur de procéder à telle ou telle étape de l'analyse juridique<sup>31</sup>. Bien sûr, une fois les modèles développés, ceux-ci peuvent se combiner pour permettre de résoudre des problèmes plus généraux.

Cette brève présentation soulève deux interrogations. La première cherche à comprendre pourquoi, au vu de l'état d'avancement théorique des modèles informatiques du droit, n'est-il pas encore possible de résoudre informatiquement tous les types de problèmes juridiques ? La deuxième cherche à comprendre, compte tenu de l'ancienneté du domaine *AI & Law*, où trouver la cause de l'engouement médiatique actuel relatif à l'application de l'intelligence artificielle en droit ? La réponse à ces deux questions se trouve, non pas dans la théorie de la représentation informatique du droit, mais dans sa mise en pratique via au travers de deux paradigmes distincts.

---

<sup>26</sup> Ce qui implique de répondre à des questions épineuses telles que : comment gérer la texture ouverte du droit sous forme informatique ? Comment gérer les références croisées des textes légaux ?

<sup>27</sup> Ce qui implique de modéliser mathématiquement différents types de raisonnements possibles tels que l'interprétation téléologique ou contextuelle.

<sup>28</sup> Ce qui doit être fait pour différents types de documents : contrats, actes unilatéraux, documents factuels, textes normatifs. C'est également ce sous-champ de recherche qui, en tentant de résoudre ce type de problème, a donné naissance aux moteurs de recherche juridique tels que ceux proposés par Stradalex ou Jura, largement utilisés par la communauté des juristes belges.

<sup>29</sup> L'interaction de différents arguments doit, entre autre, pouvoir prendre en compte des conflits de valeurs et doit pouvoir gérer des arguments en faveur et opposés à une cause.

<sup>30</sup> Nous renvoyons le lecteur intéressé vers l'excellent livre de K., ASHLEY, *Artificial intelligence and legal analytics : new tools for law practice in the digital age*, Cambridge: Cambridge University Press, 2017. La présentation et l'explication des détails techniques de ces modèles informatiques du droit occupe la première moitié du livre, c'est-à-dire les chapitres 1 à 6 : pages 51 à 201.

<sup>31</sup> Sans que les résultats en soient pour autant exempts de défauts !

### 2.1.2. *Les systèmes-experts : le paradigme initial*

Jusque récemment, la mise en pratique des solutions à tous les problèmes présentés ci-dessus avait une caractéristique commune : toutes impliquaient l'utilisation d'un certain type de programme, dénommé « système-expert ». Celui-ci peut se définir comme « *une application informatique qui contient des représentations de connaissance et d'expertise (...) pouvant être appliquée - à la manière des humains - à la résolution de problèmes, au conseil ou à une variété de tâches. En droit, l'idée est d'utiliser les technologies de l'information pour rendre l'expertise et la connaissance [juridique] plus largement et facilement accessibles* »<sup>32</sup>.

La caractéristique principale de ce type de système est donc la *reproduction* d'une expertise préalablement existante par des moyens informatiques. Ces applications se concentrent le plus souvent sur des domaines juridiques précis et utilisent des procédés qui ne sont pas facilement généralisables à d'autres domaines du droit<sup>33</sup>. Pour comprendre, il est utile de présenter une vue d'ensemble<sup>34</sup> d'un exemple de conception d'un système-expert permettant de résoudre un problème pratique mettant en jeu l'un des sous-problèmes mentionnés ci-dessus.

L'une des applications les plus utilisées de modélisation de l'information comprise dans une norme légale est celle qui consiste à déterminer si une pratique donnée d'une certaine entreprise est conforme à la loi<sup>35</sup>. Pour automatiser cette tâche à l'aide d'un système expert, deux modélisations sont nécessaires. En effet, il faut :

#### **1. Modéliser informatiquement les normes applicables.**

Dans ce contexte, la modélisation d'une norme va consister dans un exercice de traductions successives. Le texte légal sera d'abord traduit vers des « *Business rules* », qui seront elles-mêmes traduites vers un langage formel informatique. Les *Business rules* sont des lignes directrices spécifiques à une entreprise et une situation constituant une application concrète de la règle légale. Le langage formel informatique est un ensemble de symboles qui représente des concepts et des règles liant ces symboles entre

---

<sup>32</sup> Traduction personnelle de la définition de « *expert system* » de Richard Susskind dans R. SUSSKIND, *The End of Lawyers? : Rethinking the Nature of Legal Services*, Oxford: Oxford University Press, 2010, spec. p. 120.

<sup>33</sup> K., ASHLEY, *op.cit.*, spec. p. 25.

<sup>34</sup> Pour des raisons liées au caractère mathématique et technique de l'implémentation, l'exemple ne pourra être traité en détail et la description du système-expert restera donc assez générale.

<sup>35</sup> Mieux connu sous son nom anglo-saxon : *Business Process Compliance Modeling* (BPCM). Cette utilisation de la « simple » modélisation de normes pour ce genre d'utilisation échappe à certaines critiques des systèmes-experts qui seront présentés ci-après, notamment celle liée à la vision strictement positiviste du droit. Pour cette raison, ce genre d'application est encore d'actualité. Certaines firmes, telle que Neota Logic, font d'ailleurs de la BPCM une bonne partie de leurs activités. Voyez le site internet <https://www.neotalogic.com/solution/expert-advisors/> (consulté pour la dernière fois le 3 avril 2018).

eux. Celui-ci doit satisfaire à certaines propriétés pour pouvoir interpréter correctement l'application des *Business rules* à une situation<sup>36</sup>. Ces traductions ne sont pas automatisables et requièrent la participation d'experts humains. Le tableau, ci-dessous, offre un exemple de ce processus de traduction appliqué à l'évaluation de la conformité, à certaines dispositions du droit allemand, du processus de traitement des déclarations des sinistres par les assurances allemandes.

Paraphrased regulations	Business rules	As applied to business process	Predicate logic form
GDV Code of Conduct §§5–8: customer who provides personal data must be asked for agreement if this data is to be used for marketing purposes. This agreement has to be solicited within a short time span.	R1: After activity Receive claim an activity asking the claimant for agreement has to follow.	R1: Follow this by sending a data privacy notification	followedBy("Receive claim," "Send claim and data privacy notification") AND unknown
	R2: Activity asking for agreement has to be performed at most 14 days after activity Receive claim.	R2: Send at most 14 days after claim is received	followedBy("Receive claim," "Send claim and data privacy notification") AND maxTime BetweenActivities ("Receive claim," "Send claim and data privacy notification," "14 days")
German company outsourcing its data processing must ensure service providers comply with German FDPA §4b II sentence 1 BDSG re processing, storage, and exposure of personal data.	R3: Customer DB shall not be hosted outside of Germany.	R3: Do not store outside of Germany	hostingRegion ("CustomerDB," "Germany")

Figure 1. Illustration<sup>37</sup> du processus de traduction du texte légal au langage formel<sup>38</sup>, appliquée à la traduction du code allemand régulant le traitement des déclarations de sinistres par les assurances<sup>39</sup>.

## 2. Modéliser informatiquement la pratique de l'entreprise (le processus).

Pour ce faire, on utilise la compétence des membres de l'entreprise ayant une expertise

<sup>36</sup> En plus de certaines propriétés liées à la pratique (comme le fait de faire référence, pour chaque règle en langage formel, à la source légale *ad hoc* afin de pouvoir plus simplement l'adapter en cas de changement législatif), le langage formel utilisé pour le BPCM doit également respecter les contraintes découvertes par la recherche pour mieux modéliser le langage du droit. Nous renvoyons le lecteur intéressé vers K., ASHLEY, *op. cit.*, spec. pp. 72-73.

<sup>37</sup> L'illustration provient de K., ASHLEY, *op. cit.*, spec. p. 70.

<sup>38</sup> Ici, le langage formel est la « simple » logique des prédicats. Celle-ci fut largement utilisée comme langage formel au commencement du domaine *AI & Law*. Toutefois, cette utilisation ne parvenait pas à correctement représenter bon nombre de propriétés du langage juridique. Les termes tels que « followedBy (...) » sont ici les noms des fonctions informatiques qui prennent des « valeurs de vérité » booléennes.

<sup>39</sup> GDV est l'acronyme du nom de l'association des assurances allemandes.

du processus en question. On convertit celle-ci dans un certain type de formalisme informatique qui permet de le représenter de manière *légalement pertinente*. Les détails de l'implémentation exacte importent peu pour l'exemple<sup>40</sup>, mais les concepts utilisés lors de la modélisation doivent être identiques à ceux utilisés pour la traduction des « *Business rules* » en langage formel.

L'équipe de recherche ayant automatisé l'évaluation de la procédure de gestion des plaintes d'une compagnie australienne de télécommunication soumise à l'*Australian Telecommunication Consumers Protection Code*<sup>41</sup>, a abouti à une décomposition du texte légal en 176 « *Business rules* » et 223 règles en langage formel. La procédure de gestion des plaintes, quant à elle, fut modélisée par six processus distincts en langage informatique. Une fois ce travail fait, la légalité des six processus modélisés a pu être établie automatiquement en quelques secondes.

Bien qu'assez théorique et général, cet exemple fait ressortir deux caractéristiques importantes de la création et de l'utilisation de systèmes-experts en droit : la nécessité d'une forte participation humaine et un domaine d'application relativement restreint. En effet, d'une part, la modélisation nécessite l'intervention de plusieurs experts humains et, d'autre part, les choix au niveau du langage formel, liés à la modélisation des processus, restreignent la généralité de telles applications. Les systèmes-experts présentent, en plus de leur absence de généralité, un certain nombre de défauts rédhibitoires qui rendent leur extension à tous les domaines juridiques pratiquement impossible. Les principaux<sup>42</sup> sont les suivants :

- Les techniques permettant aux systèmes experts de gérer les situations (très fréquentes en pratique) où l'information est incomplète sont le plus souvent *ad hoc* et peu fiables.
- L'acquisition de l'expertise nécessaire, manuelle la plupart des cas, est un processus long, lourd et coûteux. Ceci constitue un goulot d'acquisition et de dissémination de la connaissance et de l'expertise juridique.
- De nombreux systèmes-experts se fondent sur une conception philosophique du droit qui est souvent problématique comme, par exemple, le positivisme juridique de Harts<sup>43</sup>.

---

<sup>40</sup> L'utilisation d'une représentation en réseaux de Petri est fréquente. L'explication des raisons de ce choix ou de ce qu'est précisément un réseau de Petri sort néanmoins du cadre de ce travail. Nous renvoyons donc le lecteur intéressé vers K., ASHLEY, *op. cit.*, spec. pp. 74-75.

<sup>41</sup> Seule la section 8 du code fut utilisée. Voyez l'article publié sur le processus d'automatisation en question : M., HASHMI, G., GOVERNATORI, « Normative Requirements for Business Process Compliance », *Australasian Symposium on Service Research and Innovation*, 2013, spec. p. 13.

<sup>42</sup> Les défauts principaux mis en évidence ici s'inspirent fortement de K., ASHLEY, *op. cit.*, spec. pp. 27-28.

<sup>43</sup> En effet, l'utilisation de traductions successives présentées ci-dessus suppose, tout d'abord, que le droit est un ensemble cohérent de règles et, ensuite, que les mots dont sont composées ces règles ont un sens « clair » ou, du

Celui-ci sous-tend, par exemple, le processus de traduction du texte légal en règle formelle dans l'exemple présenté ci-dessus.

Toutefois, il est important de comprendre que l'exemple ci-dessus, et plus généralement toutes les applications de systèmes-experts à la résolution de problèmes juridiques sont considérés comme des exemples d'utilisation de l'intelligence artificielle en droit. La modélisation par des systèmes-experts est d'ailleurs bel et bien considérée comme une branche de l'intelligence artificielle. C'est toutefois les défauts de ceux-ci qui ont incité les chercheurs à explorer d'autres alternatives.

### 2.1.3. *Le nouveau paradigme : le machine learning*

C'est sur un autre paradigme d'utilisation de l'intelligence artificielle que se fonde les applications qui déclenchent l'engouement médiatique<sup>44</sup> relatif à cette question.

Pour trouver des traces de ce nouveau paradigme, il suffit de regarder les publications les plus récentes du journal de la communauté *AI & Law* : une part importante des articles qui y sont publiés est liée au *machine learning*<sup>45</sup>. L'article introductif du numéro de 2017 mentionne même explicitement : « *L'intelligence artificielle est actuellement au centre de l'attention des professionnels du droit. De nombreuses startups explorent les potentielles applications de l'IA au domaine juridique (...). Les facteurs à l'origine de cette attention accrue comprennent : des avancées importantes en machine learning et traitement du langage naturel (...)* »<sup>46</sup>.

Le domaine *AI & Law* n'a, en effet, pas échappé à la tendance récente dans le domaine de la recherche en intelligence artificielle (et en dehors) consistant en une forte augmentation de

---

moins, que l'indétermination qui les caractérise se situe autour d'un sens intrinsèque (la fameuse « pénombre »). Dans le cadre de lois liées à la régulation économique, ces hypothèses (qui trouvent leurs inspirations dans une conception strictement positiviste du droit) peuvent très bien fonctionner en pratique. Toutefois, celles-ci ne donnent pas d'aussi bons résultats dans d'autres domaines. Or, ces hypothèses sont difficiles à contourner lors de la création de systèmes-experts. Voyez, pour plus d'informations, G., BOELLA *et al.*, « A Critical Analysis of Legal Requirements Engineering from the Perspective of Legal Practice », *7th IEEE International Workshop on Requirements Engineering and Law (RELAW)*, 2014, pp. 14-21, spec. pp. 16-18 et A. PALIWALA, « Rediscovering artificial intelligence and law: an inadequate jurisprudence? », *International Review of Law, Computers & Technology*, Vol. 30, N°3, 2016, pp.107-114, spec. p. 108.

<sup>44</sup> Une partie de l'explication doit certainement aussi être trouvée dans la conjoncture médiatique et dans une intense campagne de communication de la part d'IBM, voyez sur ce point O. EZRATTY, *Les avancées de l'intelligence artificielle - 3*, 2016. Articles publiés en ligne à l'adresse <http://www.oezratty.net/wordpress/2016/avancees-intelligence-artificielle-3/>, dernière consultation le 3 avril 2018.

<sup>45</sup> Par exemple, parmi les publications disponibles sur le site de Springer, mais non encore parues dans un journal, 30% concernent directement l'utilisation du *machine learning*. Aucune autre catégorie n'obtient un tel score puisque les articles abordent des sujets très divers tels des thèmes théoriques ou des comptes rendus sur l'activité de la recherche elle-même.

<sup>46</sup> Traduction personnelle de l'article : F. BEXHENRY *et al.*, « Introduction to the special issue on Artificial Intelligence for Justice (AI4J) », *Artificial Intelligence and Law*, 2017, Vol. 25, N°1, pp. 1–3, spec. p. 1.

l'intérêt pour les techniques traditionnellement associées à ce domaine<sup>47</sup>. De manière très générale, le domaine du « *machine learning* »<sup>48</sup> constitue l'étude et la mise en pratique de l'ensemble des techniques algorithmiques dont l'objectif est de faire « apprendre » à une machine, sur la base de données, certaines connaissances ou certains comportements sans lui donner explicitement de savoir ou d'instructions. Cet apprentissage peut avoir comme finalité la prédiction de propriété d'une donnée sur base d'autres données connues ou la découverte de propriétés dans un ensemble de données<sup>49</sup>.

Ce qui est frappant c'est que l'histoire de ce domaine est aussi ancienne que la recherche en intelligence artificielle<sup>50</sup> et suit une évolution dont certains aspects sont parallèles au domaine *AI & Law*. Il développe, en effet, ses propres journaux et sa propre communauté au milieu des années quatre-vingt, notamment à cause de la domination de la mise en pratique de la recherche en intelligence artificielle par l'approche par systèmes-experts<sup>51</sup>. Or, comme cela sera expliqué plus en détails section 3.1.1, le *machine learning* présente des liens très étroits avec le domaine de la statistique<sup>52</sup>. Il n'est donc, en réalité, pas surprenant que la croissance de l'intérêt dans ce domaine intervienne en parallèle de la plus grande explosion du nombre de données disparates que l'humanité ait jamais connue<sup>53</sup>. Cette explosion de la quantité de données produites est plus connue sous le nom de « *big data* »<sup>54</sup> et il est clair que les deux tendances sont liées. En effet, le *machine learning* est adapté aussi bien au traitement de ces données (pour les rendre utilisables et compréhensibles) qu'à leur exploitation. Par exemple, pour les mêmes raisons que

---

<sup>47</sup> Pour ce qui concerne la recherche, voyez M., JORDAN et T., MITCHELL, « Machine learning: Trends, perspectives, and prospects », *Science*, Vol. 349, N°6245, 2015, pp. 255-260, spec. p. 256. Pour ce qui concerne d'autres domaines, tel que le monde des affaires, voyez par exemple : J. HURWITZ ET D. KIRSCH, *Machine learning for dummies : IBM limited edition*, New Jersey, John Wiley & Sons, Inc., 2018, spec. pp. 1, 3-4.

<sup>48</sup> Il n'existe pas de définition « officielle » du domaine du *machine learning*. Cette définition, centrée sur son objectif, s'inspire néanmoins de la présentation donnée dans M., JORDAN et T., MITCHELL, *op. cit.*, spec. p. 255 et dans R., MICHALSKI, et al. *Machine Learning: An Artificial Intelligence Approach*, Springer Berlin Heidelberg, 2013, spec. pp. 3-6.

<sup>49</sup> Une distinction devrait en réalité être faite entre le *machine learning* qui se référerait uniquement à la première finalité et le *data mining* qui serait lié à la deuxième. Nous avons volontairement décidé d'omettre cette distinction dans un objectif de simplification.

<sup>50</sup> Par exemple, le terme « *machine learning* » lui-même était déjà utilisé dans les laboratoires d'IBM en 1959. Voyez, par exemple, A. SAMUEL, « Some Studies in Machine Learning Using the Game of Checkers », *IBM Journal of Research and Development*, Vol.3, N°3, 1959, pp. 210-229.

<sup>51</sup> Voyez sur ce point, et pour plus d'informations, P., LANGLEY, « The changing science of machine learning », *Machine Learning*, Vol. 82, N°3, 2011, pp. 275-279, spec. p. 275.

<sup>52</sup> On le présente même parfois comme à l'intersection de la statistique et de la science informatique, voyez M., JORDAN et T., MITCHELL, *op. cit.*, spec. p. 255.

<sup>53</sup> Voyez, sur ce point, J. DEAN, *Big Data, Data Mining and Machine Learning : Value Creation For Business Leaders and Practitioners*, New Jersey, John Wiley & Sons, Inc., 2014, spec. pp. 1-9.

<sup>54</sup> *Big data* est ici défini comme un phénomène. A l'origine, ce terme se référerait en réalité à un ensemble de données d'une grande taille et le terme est encore souvent utilisé en ce sens. Cette polysémie étant parfois source de confusion, nous n'utiliserons ce terme que pour désigner le phénomène et non des données en elles-mêmes.

dans le cas de l'inférence statistique<sup>55</sup>, les prédictions par *machine learning* sur le comportement général de données étudiées sont généralement meilleures avec un échantillon de données plus grand. La gestion et l'utilisation d'une telle quantité de données est également fortement demandeuse en puissance de calcul, raison supplémentaire pour laquelle le *machine learning* fait parler de lui aujourd'hui plutôt que dans les années nonante.

Le *machine learning*, en « apprenant » à partir de données (par un mécanisme qui sera expliquée section 3.1.1) contourne le problème de l'obtention et de la formalisation de l'expertise humaine qui constituait un des problèmes majeurs des systèmes-experts. De plus, celui-ci se combine bien avec d'autres algorithmes, dans une grande diversité de champs de recherche. Il a ainsi permis, par exemple, de grands progrès<sup>56</sup> dans la recherche sur le traitement du langage naturel ou l'extraction d'informations textuelles.

Toutefois, l'utilisation du *machine learning* implique aussi un changement d'objectif, du moins à moyen terme. Tout d'abord, au niveau des problèmes traités, l'utilisation du *machine learning* incite à se focaliser sur la résolution de sous-problèmes qu'il permet de résoudre de manière très performante, plutôt que sur la réalisation d'un système informatique pouvant performer l'ensemble des composantes de l'analyse juridique. Ensuite, au niveau de l'évaluation des performances, le *machine learning* tente d'émuler les *résultats* produits par l'intelligence, plutôt que d'imiter les processus utilisés par celle-ci<sup>57</sup>. Nous reviendrons sur ce point important section 3.1.4.

Entre statistiques, prédictions et compréhension du langage naturel, nous retrouvons ici les grands thèmes médiatiques qui ont constitué le point de départ de notre investigation<sup>58</sup>. C'est donc bien ce phénomène, qu'il nous faut analyser.

#### 2.1.4. *Conclusion concernant la technique : machine learning et big data*

La conclusion de notre investigation des techniques informatiques et de la recherche

---

<sup>55</sup> Le domaine de l'inférence statistique est un ensemble de procédés permettant de tirer des conclusions sur une « population » (un grand ensemble de données qui forme le sujet d'étude) à partir d'un échantillon (un petit nombre de données issues de cette population). C'est donc par nature une technique « prédictive ». L'inférence est généralement « meilleure » lorsque l'échantillon est grand.

<sup>56</sup> Voyez M., JORDAN et T., MITCHELL, *op. cit.*, spec. pp. 255-256, 258.

<sup>57</sup> Voyez, sur ce point, D., KATZ, « Quantitative legal prediction—or—how I learned to stop worrying and start preparing for the data-driven future of the legal services industry », *Emory law journal*, Vol. 62, pp. 909-966, spec. p. 918.

<sup>58</sup> La plupart des articles de presse cités en introduction concernent, en effet, ces trois thèmes. Soit ils mentionnent directement la prédiction des décisions de justice (plus précisément, celles de la Cour européenne des droit de l'Homme) soit ils parlent de l'application *Watson*, création d'IBM pour faciliter et améliorer le traitement automatique de textes désorganisés via une meilleure compréhension du langage naturel écrit soit, enfin, ils annoncent l'arrivée de statistiques très fines de la jurisprudence.



académique est que « l'intelligence artificielle en droit » tourne autour de l'utilisation du *machine learning* pour la résolution de problèmes juridiques. Cette utilisation est liée à deux phénomènes, d'une part l'augmentation de la puissance de calcul et, d'autre part et crucialement, au *big data*. Les systèmes-experts, bien qu'encore utilisés en pratique, ne semblent pas (encore ?) participer à la révolution supposée du droit par les technologies.

Toutefois, la combinaison du *big data* avec l'utilisation croissante du *machine learning* laisse place à une grande variété potentielle de mise en pratique de ces techniques. C'est pour cette raison que nous aborderons l'analyse du marché des nouveaux outils et services juridiques dans la section suivante.

Si c'est bien l'influence de l'utilisation du *machine learning* que nous étudierons dans la suite de ce travail, il serait une erreur, sur base de ce qui vient d'être présenté ici, de déconsidérer les systèmes-experts et les modèles informatiques du raisonnement juridique développées par la communauté *AI & Law*. En réalité, il eut été possible de présenter directement le phénomène *big data* et la montée du *machine learning*, en négligeant les courants liés aux modèles informatiques du droit et aux systèmes experts. Or, l'étude de ces outils et techniques par la communauté des juristes demeure importante pour au moins deux raisons<sup>59</sup>. Premièrement, comme cela a déjà été mentionné, l'utilisation de ce genre de système reste répandue<sup>60</sup>. Deuxièmement, et de manière cruciale, bien que les exemples concrets existants soient encore peu nombreux, le potentiel d'une combinaison des approches est énorme<sup>61</sup>. Certains auteurs considèrent même que cette combinaison constitue la réelle révolution à venir. Ceux-ci mettent en avant, par exemple, que la possibilité d'extraire des informations textuelles ayant un important contenu sémantique, comme des arguments juridiques, à l'aide du *machine learning* puis d'intégrer cet outil extractif à ceux de gestion des arguments déjà développés par la communauté scientifique pourrait donner naissance à un nouveau type d'outils juridiques révolutionnaire. Ceux-ci, pour simplifier, élimineraient l'idée même de recherche juridique telle que conçue actuellement<sup>62</sup>.

---

<sup>59</sup> Toutefois, les limitations liées à l'ampleur de notre travail nous empêchent d'en faire un traitement complet.

<sup>60</sup> Voyez ci-dessus l'exemple de Neota Logic.

<sup>61</sup> Pour un exemple d'une première tentative de combinaison des approches, déjà en 2005, voyez M. MOZINA *et al.*, « Argument Based Machine Learning Applied to Law », *Artificial Intelligence and Law*, Vol. 13, N°1, pp. 53-73. Cette combinaison des approches est aussi le sujet des chapitres 10 à 12 du livre de K., ASHLEY, *op. cit.*

<sup>62</sup> Au sens de « recherche de textes ou arguments légaux, jurisprudentiels ou doctrinaux », pas au sens de « recherche académique » en droit. Pour plus de détails sur cette vision, voyez K., ASHLEY, *op. cit.*, spec. pp. 28-30.

## 2.2. Utiliser l'intelligence artificielle en droit : le marché des *legaltechs*

Contrairement à ce que d'aucuns pourraient supposer « naïvement », l'état du marché des outils juridiques à destination des juristes et non-juristes ne suit pas fidèlement l'évolution de la recherche académique sur l'intelligence artificielle en droit. Toutefois, ce constat n'est pas si surprenant s'il est pris en compte que ce marché suit une autre logique que celle de la recherche : la logique économique. Réalisant cela, il semble même normal de trouver sur le marché une grande diversité d'outils proposés : pourquoi les outils faisant usage de l'intelligence artificielle seraient-ils seuls à être utiles au consommateur-juriste<sup>63</sup> ?

### 2.2.1. Une grande diversité de l'offre

Différents types de technologies sont utilisés par les jeunes entreprises techno-juridiques<sup>64</sup>. Certaines utilisent des techniques qui peuvent être considérées comme de l'intelligence artificielle, mais beaucoup en utilisent d'autres qui sont parfois tout autant « à la mode » que ces dernières. Ainsi retrouve-t-on l'utilisation de plateformes collaboratives (dans les produits offerts, par exemple, par Legal Space ou Legal Salah<sup>65</sup>), de moteurs de recherche de « nouvelle génération » (comme celui proposé par lex.be<sup>66</sup>), ou celles de technologies de type « *blockchain* » (proposées, par exemple, par SmartContracts<sup>67</sup> ou BitProof<sup>68</sup>).

Parmi celle usant de l'intelligence artificielle, on retrouve certaines *legaltechs* qui exploitent de « nouveaux » systèmes-experts ou des anciennes techniques développées par la recherche en *AI & Law*. Celles-ci proposent, par exemple, le « remplissage automatique » de contrats standardisés sur base de réponses à des questions préétablies (DoxMaster<sup>69</sup>) ou des dialogues avec un automate répondant à certaines questions précises sur un domaine spécifique ou pour un problème précis (comme le désormais célèbre automate de DoNotPay<sup>70</sup>). Si elles n'ont donc

---

<sup>63</sup> La présentation qui suit se fonde en majorité sur quelques enquêtes de marché concernant les *legaltechs* : celle de l'entreprise Day One, celle de l'entreprise Tracxn et celle de Charpentier Consulting. Toutes sont consultables en lignes, aux adresses référencées aux notes de bas de page *ad hoc*, et listent un bien plus grand nombre d'entreprises que nous ne le faisons ici.

<sup>64</sup> Voyez l'étude de DAY ONE, *Droit & digital : réalité et perspectives*, 2017, publiée en ligne à l'adresse : [http://www.actuel-direction-juridique.fr/sites/default/files/droit\\_digital\\_realites\\_prospectives\\_vf.pdf](http://www.actuel-direction-juridique.fr/sites/default/files/droit_digital_realites_prospectives_vf.pdf), spec. p. 8.

<sup>65</sup> Voyez l'enquête de marché TRACXN, *Legal Tech Report*, 2016, publiée, en ligne, à l'adresse : <https://s3.amazonaws.com/tracxnblog/wp-content/uploads/2016/12/Tracxn-Research-Legal-Tech-Landscape-December-2016-1.pdf>, spec. p. 130.

<sup>66</sup> Lex.be se définit elle-même comme « nouvelle plateforme de recherche juridique innovatrice », voyez son site internet à l'adresse : <https://lex.be/>.

<sup>67</sup> TRACXN, *op. cit.*, spec. p. 56.

<sup>68</sup> DAY ONE, *op. cit.*, spec. p. 25. Le site web de cette entreprise est accessible à l'adresse : <http://dev.bitproof.io/>

<sup>69</sup> *Ibid.*

<sup>70</sup> Une plateforme qui aide les particuliers à ne pas payer leurs amendes de stationnement. Voyez leur site internet à l'adresse : <https://donotpay.com/>.

rien à voir avec le nouveau paradigme présenté ci-dessus, ces *legaltechs* peuvent aussi avoir une influence sur la pratique du droit, son efficacité et l'accès à la justice. Une étude approfondie de celles-ci devrait donc dans l'idéal être menée. Toutefois, encore une fois par manque de temps, ces *legaltechs* ne seront pas étudiées dans le cadre de ce travail.

Enfin, il existe aussi des *legaltechs* faisant usage du *machine learning* et donc exploitant le nouveau paradigme technique. Ces entreprises ne représentent néanmoins pas une grande proportion de l'ensemble des *legaltechs*, à savoir moins de 15%<sup>71</sup>, bien qu'elles réalisent généralement des tâches d'une complexité supérieure à celles réalisées par les autres *legaltechs*. On trouve tout d'abord de l'analyse sémantique et automatique de documents légaux (tels que des contrats, des statuts,...) comme celle proposée par eBrevia<sup>72</sup> ou Kira Systems<sup>73</sup>. Ce type de service est particulièrement utile, par exemple, dans le cadre d'une procédure de « *due diligence* »<sup>74</sup> avant une fusion ou une acquisition. L'analyse des « émotions » de ces textes est même possible, par exemple, dans le but de détecter les courriels agressifs dans une abondante correspondance électronique. Une autre tâche nécessitant le *machine learning* est l'analyse fine et automatique de milliers de décisions de justice. Cette analyse peut mener à la production de statistiques concernant celles-ci ou la production de prédictions algorithmiques sur l'issue d'un litige à l'aide de ces dernières. Comme nous l'avons déjà mentionné, cette pratique (les deux aspects statistique et prédictif) est plus souvent appelée « justice prédictive » et est le domaine de *legaltechs* telles que Lex Machina, LexPredict, Predictice<sup>75</sup> ou encore Case Law Analytics<sup>76</sup>.

### 2.2.2. *Quelques tendances générales*

Malgré la diversité des types d'outils proposés, certaines tendances générales peuvent être dégagées. Pour les présenter, il est utile de suivre la décomposition analytique fournie par la « matrice Day One », présentée Figure 2.

Celle-ci présente une décomposition des *legaltechs* selon trois critères.

- **Objectif de l'entreprise – remplacer ou améliorer.**

Représentés sur l'axe vertical de la matrice, les termes « remplacer » ou « amélioration » se

---

<sup>71</sup> Plus précisément, 20 entreprises sur les 140 recensées. Voyez DAY ONE, *op. cit.*, spec. p. 38.

<sup>72</sup> TRACXN, *op. cit.*, spec. p. 54.

<sup>73</sup> *Ibid*, spec. p. 98.

<sup>74</sup> *Due diligence* est le nom anglais (et internationalement utilisé dans d'autres langues) pour le processus de vérification qu'une entité va réaliser avant de procéder à une transaction, afin de prévenir de potentiels problèmes ou surprises.

<sup>75</sup> Voyez DAY ONE, *op. cit.*, spec. p. 32.

<sup>76</sup> CHARPENTIER CONSULTING, *French Legaltech List*, 2017, publié, en ligne, à l'adresse: [https://media.wix.com/ugd/c21db1\\_14b04c49ba7f46bf9a5d88581cbda172.pdf](https://media.wix.com/ugd/c21db1_14b04c49ba7f46bf9a5d88581cbda172.pdf), spec. p. 2.

réfèrent non pas au juriste mais à une de ses tâches<sup>77</sup> (écrire un contrat, communiquer avec des collègues, ...). On constate que l'écrasante majorité (84%) des *legaltechs* cherchent à améliorer (ou à accélérer) ce que le juriste fait et non pas à rendre cette tâche totalement automatique.

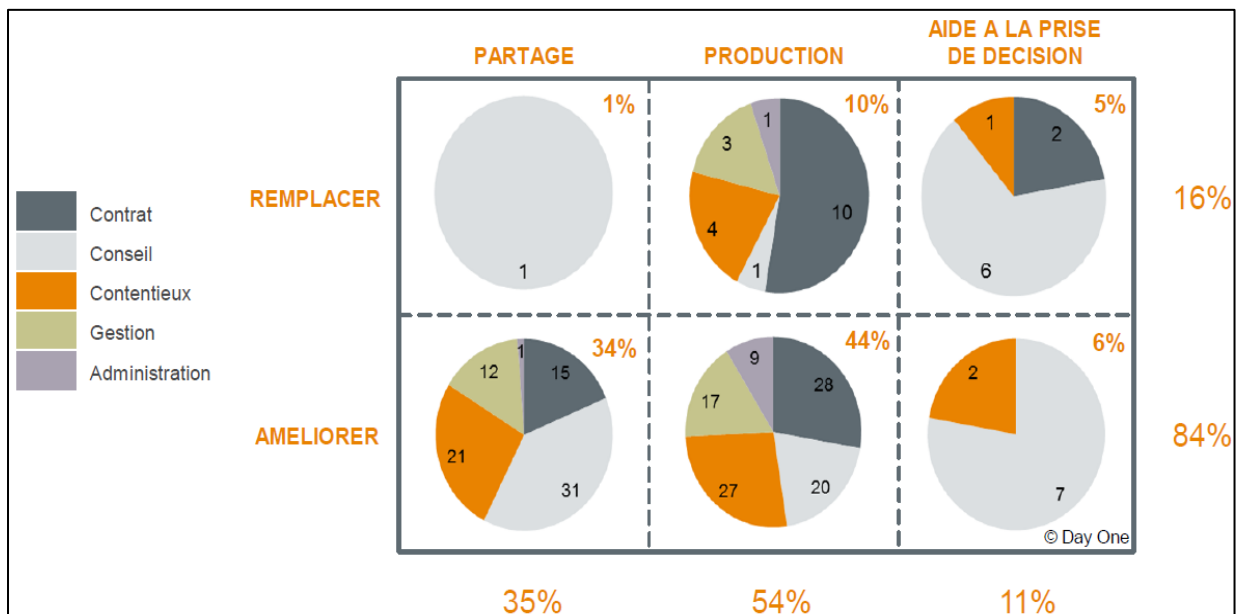


Figure 2. Décomposition des *legaltechs* par l'étude de l'entreprise Day One selon trois critères : objectifs de l'entreprise, type d'outils ou de services fournis et domaine juridique de ces derniers<sup>78</sup>.

- **Type(s) de service(s) fourni(s) : le partage, la production ou l'aide à la décision.**

L'axe horizontal de la matrice présente un découpage selon trois catégories de services fournis (directement ou par l'outil vendu). La première catégorie, le partage, se réfère à la diffusion d'information (mise en contact, partage des résultats d'une recherche,...) tandis que les deux autres catégories se réfèrent à des tâches juridiques. La production est une catégorie générique qui se réfère à tous les services juridiques (tel que l'écriture de contrat(s) ou l'exécution de formalités administratives, ...) et l'aide à la décision se focalise sur cette tâche précise qu'est la prise de décision juridique. Comme cela ressort clairement de la Figure 2, les catégories production et partage rassemblent l'écrasante majorité (89%) des *legaltechs*. La plus petite catégorie de l'aide à la décision rassemble, quant à elle, la majorité (65%) des *legaltech* utilisant de l'intelligence artificielle liée au nouveau paradigme<sup>79</sup>.

- **Domaine juridique des services fournis : conseils, contrats, contentieux, ...**

<sup>77</sup> *Ibid*, p. 14.

<sup>78</sup> DAY ONE, *op. cit.*, spec. p. 16.

<sup>79</sup> DAY ONE, *op. cit.*, spec. p. 39.

Enfin, un dernier critère faisant appel à une légende colorée classe les *legaltechs* selon le « domaine juridique » auquel elles sont associées. On peut constater sur la Figure 2 que les domaines du contrat et du conseil dominant. Si le rapport de One Day précise bien que certaines des *legaltechs* auraient pu (ou devraient) être catégorisées dans plusieurs « cases » et qu'un choix concernant la classification « principale » a dû être fait, on constate que la plupart des *legaltechs* sont concentrées sur un petit nombre de domaines spécifiques<sup>80</sup>.

Enfin, deux autres tendances n'étant pas illustrées par la Figure 2 méritent d'être mentionnées. La première est que la plus grande partie des *legaltechs* tend à offrir des services à d'autres entreprises, ou aux professionnels du droit, plutôt que directement aux justiciables<sup>81</sup>. La deuxième se retrouve dans la non-homogénéité de la répartition géographique des *legaltechs*. En effet, celle-ci sont nettement plus nombreuses aux Etats-Unis qu'en Europe ou dans le reste du monde<sup>82</sup>.

### 2.2.3. *Un marché en évolution*

Les principes présentés ci-dessus ne sont que peu, voire pas, remis en question par l'évolution du marché. En effet, l'asymétrie géographique du nombre de *legaltechs*, leur tendance à fournir des services aux professionnels<sup>83</sup> et la domination des domaines du contrat et du conseil demeurent inchangées sur les dernières années. Production et partage restent également les types de service proposés majoritairement<sup>84</sup>.

Tout n'est pas statique pour autant. Le rythme de création de *legaltechs*, aussi bien en Europe qu'aux Etats-Unis, s'est énormément accéléré entre 2008 et 2015, avec une multiplication au moins par sept du nombre de *legaltechs* fondées chaque année, avant de ralentir durant l'année 2016<sup>85</sup>. Dans la catégorie production, les services d'automatisation du traitement de documents et ceux de gestion de contrats ont connu la plus forte croissance<sup>86</sup>. Dans la catégorie partage, ce sont les entreprises offrant de mettre en contact les consommateurs, entreprises et

---

<sup>80</sup> Le nombre de *legaltechs* analysées dans l'étude de One Day est de 140. Or on compte 216 recensements dans la matrice. En effet, les *legaltechs* couvrant plusieurs domaines sont recensées plusieurs fois dans la même case. Le ratio entre le nombre de recensement et le nombre de *legaltechs* analysées, 1.54, nous donne une idée du nombre de domaines couverts en moyenne par chaque *legaltech*.

<sup>81</sup> DAY ONE, *op. cit.*, spec. p. 5.

<sup>82</sup> TRACXN, *op. cit.*, spec. pp. 35, 38, 41, 44, 47.

<sup>83</sup> DAY ONE, *op. cit.*, spec. p. 11.

<sup>84</sup> *Ibid*, spec. p. 13.

<sup>85</sup> Voyez TRACXN, *op. cit.*, spec. pp. 38, 41, 47 et DAY ONE, *op. cit.*, spec. p. 39.

<sup>86</sup> Voyez TRACXN, *op. cit.*, spec. pp. 38, 41.

professionnels du droit qui ont connu la plus forte croissance<sup>87</sup>.

De manière générale, on ne constate pas une « course vers l'intelligence artificielle ». Le nombre de *legaltechs* faisant usage de technologies du nouveau paradigme, en plus de ne pas être très grand, ne semble pas être en forte croissance<sup>88</sup>, en tout cas en Europe<sup>89</sup>. Aux Etats-Unis, cette croissance est toutefois plus prononcée<sup>90</sup>. Ces deux phénomènes sont probablement dus à la quantité d'investissement requis pour démarrer une telle entreprise, l'utilisation poussée de nouvelles techniques étant souvent très coûteuse en expertise et en récolte de données<sup>91</sup>. D'ailleurs, la tendance au développement accru de l'intelligence artificielle par les *legaltechs* se voit nettement mieux lorsque, au lieu de se concentrer sur le nombre d'entreprises créées, on analyse les capitaux investis. En effet, les *legaltechs* utilisant le nouveau paradigme de l'intelligence artificiel concentre la plus grande part des capitaux investis, du moins aux Etats-Unis<sup>92</sup>. Une explication possible de la différence d'investissement entre l'Europe et les Etats-Unis provient de la taille du marché juridique (notamment au niveau de la langue utilisée). Ceci expliquerait la quasi-absence de *legaltechs* belges exploitant le nouveau paradigme de l'intelligence artificielle<sup>93</sup>.

En analysant l'utilisation du *machine learning* par les *legaltechs* existantes, on constate que celle-ci consiste donc, principalement, dans l'amélioration et le remplacement de trois types de tâches<sup>94</sup> : la recherche juridique, le traitement de documents juridiques et la production de statistiques ou de prédictions liées aux décisions de justices.

#### 2.2.4. *La réalité des acteurs de la nouvelle vague*

Pour affiner encore notre diagnostic des pratiques sur lesquelles repose « l'intelligence artificielle en droit », ainsi que du marché des *legaltechs* en générale, nous avons réalisé une

---

<sup>87</sup> Ibid, p. 47.

<sup>88</sup> DAY ONE, *op. cit.*, spec. p. 39.

<sup>89</sup> Les *legaltechs* analysées par Day One sont plus européennes qu'américaines. Le constat de Day One est d'ailleurs confirmé par CHARPENTIER CONSULTING, *op. cit.* spec. pp. 2-5, qui liste uniquement les *legaltechs* françaises. Il est, toutefois, fortement nuancé par l'analyse de Tracxn, qui se concentre plus sur les Etats-Unis.

<sup>90</sup> TRACXN, *op. cit.*, spec. pp. 53-54.

<sup>91</sup> DAY ONE, *op. cit.*, spec. p. 38.

<sup>92</sup> Ibid et TRACXN, *op. cit.*, spec. p. 7.

<sup>93</sup> Quasi-absence car, d'une part, certains nouveaux moteurs de recherche juridiques font potentiellement une utilisation sommaire du *machine learning* et, d'autre part, des produits de « justice prédictive » sont en développement à l'intérieur de groupes d'édition tels que Larcier ou Wolter Kluwer. Ce dernier vient d'ailleurs de lancer son nouveau produit *Legal Insight*, dont la description est disponible, en ligne, à l'adresse <https://legalworld.wolterskluwer.be/fr/solutions/outils-de-calcul/legal-insights/>. Pour plus de détails sur les différents types de (nouvelles) *legaltechs* belges voyez la « carte du marché belge des *legaltechs* » en ligne, à l'adresse <https://legaltech.be/the-rise-of-legaltech-the-belgian-market-overview-map/>.

<sup>94</sup> Voyez DAY ONE, *op. cit.*, spec. pp. 42-43 et TRACXN, *op. cit.*, spec. pp. 38-42.

série d'interviews<sup>95</sup> d'acteurs de terrain. Leur objectif était de comprendre la nature précise du produit développé ou commercialisé et certains détails techniques, particulièrement l'utilisation faite du *machine learning*. Nous présentons, ci-dessous, une synthèse de ces interviews avec quatre de ces acteurs.

- *Jonathan Unikowski*, fondateur de LexLoci<sup>96</sup>.

Le produit proposé par LexLoci est un moteur de recherche de nouvelle génération. Celui-ci tente une reconceptualisation de la manière dont l'utilisateur cherche l'information. Abandonnant l'idée de mots clefs, le nouveau moteur de recherche permet, sur la base d'un (extrait de) document déjà considéré comme pertinent, l'obtention d'autres documents similairement pertinents. Il affiche ensuite, pour l'utilisateur, en parallèle, le texte donné et les (extraits pertinents) des documents trouvés.

Contrairement à d'autres *legaltechs* se concentrant sur l'amélioration de la recherche juridique<sup>97</sup>, l'algorithme utilisé ne se repose pas sur du *machine learning* mais uniquement sur des algorithmes de traitement de texte. Pour ce faire, le document (ou l'extrait) pertinent est comparé à ceux présents dans la base de données sur base de trois critères : l'utilisation des mêmes mots, le contexte lié à chaque mot et le format du texte<sup>98</sup>. La comparaison utilise principalement ainsi l'indicateur « *Tf/idf* »<sup>99</sup> qui mesure le rapport entre la fréquence d'un mot dans un document et sa fréquence dans le reste de la base de données. Il permet ainsi de caractériser à quel point un mot est spécifique au document considéré<sup>100</sup>. Sur base de ces trois comparaisons, l'algorithme classe alors les documents de la base de données par ordre de pertinence. Aux dires de M. Unikowski, les résultats concrets, sur base de tests comparatifs avec les moteurs de recherche existants, sont excellents. Toutefois, M. Unikowski note la grande réticence de la communauté juridique américaine à la nouveauté et, ce, malgré les tests et comparaisons favorables aux nouveaux produits.

---

<sup>95</sup> Les interviews furent tenues par courriel, téléphone ou *de visu*. Elles prenaient pour point de départ une série de questions: « Quel produit offrez-vous ? », « Quelle utilisation du *machine learning* faites-vous ? », « Faites-vous de l'analyse et du traitement de textes et si oui, comment ? » puis évoluait en fonction des réponses.

<sup>96</sup> Jonathan Unikowski est juriste diplômé de l'Université Libre de Bruxelles. C'est un entrepreneur belge qui s'est lancé sur la scène *legaltech* américaine. Voyez le site internet de LexLoci à l'adresse : <https://www.lexloci.com/view/hospitality>.

<sup>97</sup> Comme par exemple, Doctrine.fr ou ROSS, voyez DAY ONE, *op. cit.*, spec. p. 42.

<sup>98</sup> L'algorithme repose notamment sur certaines caractéristiques particulières que partagent le format de beaucoup de documents juridiques.

<sup>99</sup> Acronyme de « *Term frequency* » (tf) / « *inverse document frequency* » (idf).

<sup>100</sup> Cet indicateur permet, par exemple, d'éviter de considérer les termes fréquents de la langue anglaise (tels que « *the* », « *a* », « *is* », etc).

- *Louis Larret-Chahine*, Directeur Général de Predictice<sup>101</sup>.

Bien que son site internet suggère que Predictice propose de la « justice prédictive », décrire son produit comme de la « justice statistique » serait probablement plus approprié. En effet, Predictice présente à l'utilisateur des statistiques détaillées sur les décisions de justice appartenant à certains contentieux, mais n'utilise pas le *machine learning* pour produire une quelconque prédiction<sup>102</sup>. Le développement de celle-ci n'est d'ailleurs pas envisagé, en raison, d'après M. Larret-Chahine, de la réticence de la communauté juridique française à ce genre d'offre. L'appréciation de la pertinence des statistiques produites est d'ailleurs laissée à l'utilisateur. En effet, celles-ci ne s'adaptent pas automatiquement aux changements législatifs ou revirements de jurisprudence pouvant rendre certaines décisions de justice dépourvues d'intérêt.

Pour produire ces statistiques, Predictice combine le *machine learning* à des techniques d'extraction de l'information<sup>103</sup> afin d'automatiser le traitement des textes de jurisprudence. En effet, ceux-ci, obtenus au format XML ou HTML<sup>104</sup> via leur collaboration avec Wolters Kluwer, sont analysés automatiquement pour en extraire les « caractéristiques importantes » (le contentieux, les faits pertinents, ...). Ces caractéristiques sont d'abord choisies par les créateurs de l'outil<sup>105</sup> puis adaptées en fonction des commentaires des utilisateurs et de la faisabilité technique de l'extraction.

Chaque décision est alors associée informatiquement aux catégories extraites de celle-ci, ce qui permet d'en faire usage lors de la présentation de statistiques sur le contentieux *ad hoc*. Les statistiques produites selon cette méthode sont donc « classiques » mais très détaillées. Il devient, par exemple, possible de connaître la proportion de décisions favorables à l'employeur ayant été rendues ces dix dernières années dans le cadre de litiges pour licenciement abusif d'un ouvrier de plus de 40 ans gagnant mensuellement moins de 2000€ brut<sup>106</sup>.

---

<sup>101</sup> Voyez le site internet de l'entreprise Prédicite à l'adresse : <https://predictice.com/>

<sup>102</sup> Aucun autre outil de prédiction n'est d'ailleurs proposé.

<sup>103</sup> Telle que le « *semantic role labelling* ». Voyez, pour plus d'informations sur cette technique : M., PALMER, *et al.*, « Semantic Role Labelling », *Synthesis Lectures on Human Language Technologies*, Vol. 3, N°1, 2010, pp. 1-103.

<sup>104</sup> Les formats XML (acronyme de *Extensible Markup Language*) et HTML (acronyme de *HyperText Markup Language*), sont deux types de formats informatiques permettant de partager des textes avec plus d'informations que les seules lettres. L'information supplémentaire fournie peut être de nature variée mais, selon toute vraisemblance, dans notre cas, celle-ci aide à l'extraction des caractéristiques pertinentes. Le format HTML est par exemple utilisé dans la transmission de « pages » internet.

<sup>105</sup> Probablement sur base d'une analyse de la doctrine ou de conseil d'experts.

<sup>106</sup> En effet, il suffit de compter les décisions répertoriées sous les catégories extraites « contentieux : licenciement abusif », « types de liens : ouvrier », « age : 40 ans et plus », « salaire : mois de 2000€ » et



- *Mike Kramer*, Directeur des ventes chez Lex Machina<sup>107</sup> :

Le produit proposé et les méthodes employées par Lex Machina sont très proches de celles de Predictice. En effet, Lex Machina utilise également des méthodes de *machine learning* combinées à de l'extraction d'information de documents juridiques. Ses principales spécificités sont une focalisation sur la procédure et une catégorisation d'un litige fortement axée sur ses acteurs (avocat(s) adverse(s), juge(s), ...).

Pour obtenir ces informations, Lex Machina analyse les résumés de procédure des litiges judiciaires qui sont publiquement disponibles aux Etats-Unis. Elle parvient ainsi à produire, en fonction des juges et des avocats participants, des statistiques concernant les types de « *motions* »<sup>108</sup> couronnées de succès, la durée des procès, etc. Les techniques algorithmiques employées, du moins celles qui ne sont pas des secrets d'affaires, sont par contre identiques dans leur principe à celles de Predictice.

- *Jacques Levy-Vehel*, fondateur de Case Law Analytics<sup>109</sup> :

Le produit proposé par Case Law Analytics est ce qui se rapproche le plus de la « justice prédictive » à proprement parler<sup>110</sup>. Selon M. Levy-Vehel, son entreprise serait d'ailleurs la seule *legaltech* française du genre à ne pas produire « que » des statistiques détaillées concernant certains contentieux. En effet, Case Law Analytics utilise bien du *machine learning* pour produire une prédiction ou, plus précisément, cent prédictions, c'est-à-dire la solution fournie à un litige par un échantillon représentatif de cent juges.

Pour ce faire, le *machine learning* est utilisé à deux reprises. Une première phase l'utilise pour extraire, automatiquement, des informations pertinentes de textes de jurisprudence pour les

---

« dispositif : favorable à l'employeur ». On divise alors le nombre de décisions obtenues par l'ensemble des décisions des décisions (favorables et défavorables à l'employeur) aux mêmes caractéristiques pour obtenir la fraction désirée.

<sup>107</sup> Voyez le site internet de Lex Machina, et plus particulièrement les pages liées aux « *legal analytics* » à l'adresse : <https://lexmachina.com/legal-analytics/>. Lex Machina a récemment été acquise par LexisNexis, une compagnie américaine qui se spécialise dans l'utilisation de l'informatique en droit dans tous ses aspects.

<sup>108</sup> Nom anglais donné à une variété d'instruments visant, aux Etats-Unis, à obtenir du juge l'activation d'une procédure donnée.

<sup>109</sup> Voyez le site internet de Case Law Analytics à l'adresse : <http://www.caselawanalytics.com/>

<sup>110</sup> De manière surprenante, dès le début de notre interview, M. Levy-Vehel insistait fortement sur l'inexistence la « justice prédictive ». Sa propre dénomination était celle de « justice actuarielle ». Ses objections étaient à la fois théorique (il n'est, fondamentalement, pas possible de prédire le futur à partir du passé, sans même considérer l'importante part d'impondérable que présentent les décisions de justice) et éthique (il ne serait pas moral de prétendre prédire la justice). Il est vrai qu'il est toujours utile de rappeler que l'utilisation des termes « justice prédictive » ne se rapporte pas à une prétention de prédire la décision de justice comme un physicien prédit la position des astres mais plutôt à celle d'un actuair prédisant statistiquement (et donc de manière uniquement valable pour un grand nombre de clients) les pertes et le profit de sa compagnie d'assurance selon qu'elle adopte telle ou telle stratégie.

catégoriser selon celles-ci. Cette phase est similaire à ce que font Predictice et Lex Machina. Toutefois, chez Case Law Analytics, ce traitement automatique n'est qu'une partie du travail de « catégorisation »<sup>111</sup> des décisions. La majeure partie est réalisée « à la main » par des experts, sur base de consultations des juges de ce contentieux ou d'avocats spécialisés. Ceci permet de catégoriser les décisions de justice de manière plus fine (avec entre 20 et 90 caractéristiques prises en compte) et en fonction de critères qui ne peuvent pas techniquement être obtenus à l'heure actuelle par leurs algorithmes<sup>112</sup>. Le désavantage de cette méthode est que la prédiction n'est possible que dans un nombre plus réduit de contentieux où ce travail d'expert a pu être réalisé.

Lors d'une deuxième phase, un algorithme de *machine learning*<sup>113</sup> « apprend »<sup>114</sup>, sur base des catégorisations effectuées, les schémas de décision. Il permet ensuite la prédiction de ce que diraient 100 juges représentatifs en face d'un litige représenté, bien sûr, selon les caractéristiques extraites initialement. Toutefois, à l'instar de Predictice, il revient à l'utilisateur de faire preuve de vigilance quant à la pertinence du résultat car celui-ci ne s'adapte que partiellement aux revirements de jurisprudence et pas du tout aux changements législatifs.

#### 2.2.5. *Conclusion concernant le marché des legaltechs*

En bilan de cette investigation du marché des *legaltechs*, il semble que l'utilisation de l'intelligence artificielle en droit ne constitue, à l'heure actuelle, qu'un pan relativement restreint, en ampleur, de l'utilisation des technologies récentes dans le domaine juridique. En effet, les analyses de marché ont montré, et l'interview de Jonathan Unikowski a confirmé, que l'absence d'utilisation du *machine learning* ou de systèmes-experts (les deux catégories « d'intelligence artificielle » que nous avons ciblées dans la section 2.1) n'empêche pas l'introduction intéressante d'autres technologies informatiques dans la sphère juridique.

Les utilisations du *machine learning* en droit semblent ne pas encore avoir atteint leur pleine maturité. En effet, celles-ci se concentrent d'une part sur l'amélioration de la recherche juridique à l'aide du *machine learning* (permettant une meilleure évaluation de la pertinence du résultat) et, d'autre part, sur deux types de nouvelles pratiques. La première, et la plus importante, consiste dans l'analyse et le traitement textuel automatique de documents

---

<sup>111</sup> Cette catégorisation fait partie du travail de représentation informatique des données. Ceci sera développé plus en détails section 3.1.1.

<sup>112</sup> Comme, par exemple, une indemnité de licenciement en fonction du nombre de mois de salaire plutôt qu'en somme d'argent.

<sup>113</sup> Dont les détails sont un secret d'affaire, mais dont les principes ne diffèrent pas des autres techniques de *machine learning* supervisées expliquées plus loin.

<sup>114</sup> Nous verrons en détails ce que cela signifie section 3.1.1.

juridiques, afin de permettre une amélioration sensible de la connaissance de ceux-ci. Dans le cadre d'analyse de décisions de justice, cette connaissance passe particulièrement (mais pas exclusivement) par la production de statistiques sur les informations extraites de ces documents. Cette pratique, bien qu'encore récente, est déjà bien présente sur le marché et semble, sur base des interviews recueillies, fonctionner très correctement. L'autre pratique nouvelle est celle de l'utilisation d'algorithmes de *machine learning* pour prédire les décisions de justice. Bien qu'également présente sur le marché, cette pratique-là est moins répandue et n'en est qu'à ses débuts.

Enfin, malgré le fait que ces évolutions semblent freinées par la réticence des juristes à utiliser des technologies récentes<sup>115</sup> qu'ils comprennent mal, l'investissement (si pas le rythme de création d'entreprise) est en forte augmentation, particulièrement aux Etats-Unis. Si l'on peut donc prévoir une augmentation continue de la commercialisation et de l'utilisation concrète de ces technologies à court et moyen terme<sup>116</sup>, il ne semble pas que l'intelligence artificielle soit sur le point de conquérir le monde (juridique).

### **3. La portée de la justice analytique : implémentation, pratiques et futur du droit.**

Au cours des deux sections précédentes, nous avons déterminé le type d'objet à analyser pour répondre à notre question sur l'impact de l'intelligence artificielle en droit. Dans un premier temps, nous avons conclu que l'utilisation du *machine learning* et le phénomène du *big data* étaient au cœur des techniques liées à la notion d'intelligence artificielle telle que comprise implicitement dans la question originelle. Nous avons ensuite étudié le marché des *legaltechs* et déduit que, bien que la majorité de ces dernières n'utilisent pas ce nouveau paradigme technique, celles qui en font usage le font selon trois types de pratiques : l'amélioration de la recherche, l'analyse automatique de documents et la prédiction de, ou la production de statistiques sur, l'issue d'un litige. Bien que ces trois types de pratiques laissent encore la place à une variété importante de mises en œuvre concrètes, nous avons considérablement circonscrit ce dont il nous faudra analyser les conséquences. De plus, ce faisant, nous nous sommes assurés que notre conception de « l'intelligence artificielle en droit »

---

<sup>115</sup> Ce qui ressort de nos interviews mais qui était déjà annoncé dans certaines études sur les nouveaux outils juridiques, voyez par exemple l'enquête publiée sur le site internet du journal le Monde à l'adresse : <http://internetactu.blog.lemonde.fr/2017/09/13/la-justice-predictive-23-predictions-et-regulations/>.

<sup>116</sup> Au fur et à mesure que les investissements actuels aboutissent à des projets concrets et que les réticences des communautés juridiques s'estompent.

est cohérente avec les évolutions théoriques récentes du domaine *AI & Law* et en phase avec la pratique technologique et commerciale des *legaltechs* déjà présentes sur le marché. Nous avons, finalement, pris une mesure de l'ampleur actuelle et future du phénomène.

Un point de terminologie reste toutefois à régler. Aussi bien dans les médias grands publics que dans les publications à destination des juristes, la désignation la plus fréquente pour la nouvelle pratique juridique basée sur l'intelligence artificielle est « justice prédictive », que nous avons déjà mentionnée à plusieurs reprises. Nous pensons toutefois avoir démontré dans les sections précédentes que, si la prédiction fait bien partie des nouvelles pratiques, elle n'en est ni l'unique composante ni même la composante principale. En réalité, la caractéristique commune aux trois nouvelles pratiques décrites ci-dessus est non pas la prédiction, mais *l'analyse* de données juridiques (particulièrement textuelles). La dénomination la plus appropriée, qui sera utilisée dans le reste du présent travail, est donc, selon nous, « justice analytique », qui possède trois faces pratiques : la recherche, le traitement et la prédiction.

Dans cette troisième partie, nous nous concentrerons donc sur l'analyse des trois faces de la justice analytique. Notre objectif sera de répondre à notre question de recherche (à savoir, quel sera l'impact de l'intelligence artificielle en droit ?) en testant nos deux hypothèses. Pour mémoire, celles-ci postulent, d'une part, que l'implémentation ne sera pas neutre du point de vue de la pratique juridique et de la substance du droit produit en lui-même et, d'autre part, que l'impact principal consistera en une reconfiguration de pratiques existantes autour de ces nouveaux outils. Pour ce faire, nous commencerons par présenter une vulgarisation des principes de fonctionnement du *machine learning*. Ceci nous permettra de tester notre première hypothèse et de prendre en compte l'implémentation de la justice analytique lors de l'analyse de ses conséquences. Ce faisant, nous aborderons également le phénomène *big data* indissociable de la montée en puissance du *machine learning*. Bien sûr, les conséquences de l'existence de ces nouveaux outils ne se borneront pas aux effets pouvant être déduits de leur fonctionnement et il nous faudra donc considérer aussi les changements causés par son utilisation. Nous procéderons donc à une analyse des conséquences potentielles sur le monde du droit de l'utilisation de ces nouveaux outils par les différents acteurs classiques du monde juridique. Nous aborderons finalement les liens potentiels entre la justice analytique et un autre grand courant de transformation du droit.

### **3.1. Analyse de l'implémentation technique**

L'objectif de cette section est de fournir les clefs de compréhension du *machine learning*,

nécessaires à la l'étude des conséquences que peut avoir l'implémentation de la justice analytique sur le droit. Toutefois, la présentation donnée ici ne pourra qu'esquisser les traits d'une vulgarisation du *machine learning* et ne fournira pas un aperçu détaillé, et encore moins complet, des différents aspects du domaine. En effet, il s'agit d'un ensemble de techniques algorithmiques, aux buts différents selon les applications, encore en évolution grâce aux intenses recherches menées par les mondes académique et commercial. Notamment, nous ne parlerons ici que d'une classe d'algorithmes, dénommée « *machine learning* supervisé », car c'est elle qui est utilisée en pratique dans l'écrasante majorité des applications juridiques<sup>117</sup>. En effet, ce que nous avons appelé jusqu'ici « *machine learning* », n'est en réalité qu'un sous-ensemble des techniques qui composent ce domaine<sup>118</sup>.

### 3.1.1. *Esquisse des principes du machine learning supervisé*

Comme cela fut mentionné section 2.1.3, le *machine learning* présente des liens très étroits avec le domaine de la statistique. Ces liens sont en réalité au cœur de ce qui constitue l'essence même du *machine learning* : la recherche d'une automatisation du processus de modélisation statistique<sup>119</sup>. Par modélisation statistique, nous entendons le processus permettant de réduire la complexité d'une distribution de données disparates par son remplacement fonctionnel par une description mathématique plus simple.

Comme toutes les techniques appartenant au domaine du *machine learning*, les algorithmes supervisés ont donc comme objectif de déterminer (ou découvrir) certaines « tendances » présentes dans des données à l'aide d'une modélisation statistique. Leur particularité provient du fait que, pour atteindre ce but, l'algorithme utilise des « données d'entraînement » lui permettant de déterminer les paramètres d'un certain modèle choisi par l'informaticien<sup>120</sup>. Une

---

<sup>117</sup> Le « *clustering* », une technique de *machine learning* non supervisé qui permet de rassembler des documents « similaires » en un certain nombre de groupes prédéfini. Il est également utilisé en droit, notamment dans la procédure américaine de « *discovery* » ou pour la *due diligence*. Voyez H., SURDEN, « Machine Learning and Law », *Washington Law Review*, Vol. 89, N°1, 2014, pp. 87-115, spec. pp. 113-114. Son utilisation étant néanmoins très nettement moins fréquente que celle des algorithmes supervisés, nous ne l'aborderons pas ici.

<sup>118</sup> Concernant la distinction paradigmatique entre apprentissage supervisé et non supervisé, voyez S., SUTHAHARAN, *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, Springer, New-York, 2016, spec. pp. 7-10.

<sup>119</sup> Voyez, pour cette formulation précise, T., REIGELUTH, *Comporter la norme : La normativité de l'apprentissage algorithmique à partir du problème du comportement*, thèse de doctorat, Université Libre de Bruxelles, 2018, spec. p. 14.

<sup>120</sup> L'exemple le plus connu de « modélisation statistique » est celui qui consiste à prendre un nuage de « points » (des données à deux dimensions représentées sur un graphique plan) et à l'approximer par sa « droite de régression linéaire » :  $y = ax + b$ . Dans cette formule,  $(x, y)$  sont les deux dimensions du point tandis que  $a$  et  $b$  sont les *paramètres du modèle*. Le but de ce genre de régression est de s'assurer que la droite « colle » le plus aux points (qu'elle les représente le plus fidèlement) en ajustant les valeurs de  $a$  et  $b$ . Les méthodes pour y arriver, dans le cas cet exemple simpliste, sont parfois anciennes de plusieurs centaines d'années. Voyez plus de détails :

fois le modèle paramétrisé, l'algorithme est généralement utilisé sur d'autres données, absentes des données d'entraînement, afin de « prédire » certaines de leurs propriétés.

En pratique, l'utilisation de ce type d'algorithme requiert les cinq étapes suivantes<sup>121</sup>.

### **1. Collecter des données.**

La sélection et la collecte des données sur lesquelles travailler est un prérequis à l'utilisation du *machine learning* (de tout type). Sans données, pas d'apprentissage.

### **2. Représenter informatiquement les données**

Obtenir des données sous n'importe quelle forme ne suffit pas, il faut convertir celles-ci dans un format qui peut donner lieu à des calculs. Lors de cette conversion, deux types de catégories sont distinguées : celles dont on connaîtra les valeurs pour les données « tests » et celles qui feront l'objet d'une prédiction ou d'une classification.

### **3. Choisir le modèle et l'algorithme d'apprentissage.**

Il existe une foule de modèles<sup>122</sup> permettant d'extraire du « sens » de données disparates. Le choix d'un d'entre eux, est réalisé sur base de critères relatifs à sa performance<sup>123</sup>. Il influence, d'une part, les détails de la paramétrisation de l'étape 4 et, d'autre part, le format à imposer aux données lors de l'étape 2.

### **4. Paramétriser le modèle : l'apprentissage**

Il s'agit d'utiliser les données pour sélectionner les paramètres du modèle. C'est l'implémentation algorithmique, c'est-à-dire, pour mémoire, la suite d'étapes mathématico-informatiques, qui permet de faire le lien entre les données et les paramètres du modèle. La paramétrisation est à la fois le cœur du processus et un mécanisme sans grand enseignement dans notre cas. En effet, les détails de celle-

---

J. DEAN, *op.cit.*, spec. pp. 75-90.

<sup>121</sup> La présentation qui suit est une adaptation de la présentation donnée K., ASHLEY, *op.cit.*, spec. pp. 223-225.

<sup>122</sup> Par exemple, nous pouvons citer: le modèle de régression cité ci-dessus, les modèles en arbres de décision, le modèle *Support Vector Machine*, les modèles d'apprentissage par couche (et le *deep learning*). Pour le détail des types de modèles possibles et des leurs avantages respectifs, voyez S., SUTHAHARAN, *op. cit.*

<sup>123</sup> C'est lors de cette étape qu'interviennent les considérations importantes ayant pour but d'éviter la « surinterprétation » et la « sous-interprétation ». Nous n'aurons pas la possibilité de détailler cet équilibre important dans ce travail. Toutefois, pour simplifier, il s'agit de s'assurer que l'apprentissage sur les données d'entraînement se généralisera bien aux données (« tests ») sur lesquelles l'algorithme est utilisé et que les « tendances » détectées et apprises ne sont pas trop spécifiques aux données d'entraînement. Ce problème est intimement liée à ce qui est dénommé dans la littérature le « *Bias-Variance Tradeoff* ». Pour plus d'informations sur ce point, nous renvoyons le lecteur intéressé vers les références suivantes : T. HASTIE *et al.*, *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*, 2<sup>ème</sup> édition, Springer-Verlag New York, 2009, spec. pp. 37-39, 219-257 et S., SHALEV-SHWARTZ et S., BEN-DAVID, *Understanding machine learning : From Theory to Algorithms*, Cambridge University Press, New York, 2014, spec. pp. 60-78, 144-156.

ci dépendent fortement du type de modèle utilisé. Il s'agit toutefois toujours d'une suite d'étapes permettant de déterminer de manière univoque<sup>124</sup> les paramètres du modèle à partir des données collectées. La plupart des algorithmes de *machine learning* (les modèles dits « statistiques ») utilisent des méthodes reposant sur l'utilisation d'un grand nombre de corrélations qui ne sont pas humainement compréhensibles<sup>125</sup>. Certains reposent sur des « raisonnements » pouvant techniquement être appréhendés<sup>126</sup> mais, même dans ces rares cas, il n'est pas évident de tirer une quelconque « raison » pour laquelle le modèle se retrouve dans telle ou telle configuration.

#### 5. Utiliser le modèle avec les paramètres obtenus pour prédire ou classifier<sup>127</sup>.

C'est la partie réellement utile du processus, celle qui utilise concrètement les étapes précédentes. Pour ce faire, les étapes 1 et 2 sont réitérées : de nouvelles données « tests » sont collectées et représentées de la même manière que les données d'apprentissage. Toutefois, un des éléments (une valeur numérique) de la représentation manque pour les données tests. C'est cette valeur qui va faire l'objet d'une « prédiction » à l'aide du modèle appris et des autres valeurs des données tests.

#### 3.1.2. Illustrations d'applications utilisant du machine learning

Pour clarifier ce que signifient concrètement ces étapes, il est utile de les illustrer à l'aide d'exemples concrets. Nous allons donc détailler deux exemples d'utilisation du *machine learning* : l'une dans le cadre de la prédiction de l'issue d'un litige judiciaire et l'autre dans le cadre de l'analyse de texte.

---

<sup>124</sup> Et déterministe : étant donné le même modèle et le même ensemble de données d'apprentissage, les paramètres obtenus seront toujours identiques, sauf dans des cas précis où des nombres pseudo-aléatoires sont spécifiquement utilisés.

<sup>125</sup> L'exemple le plus célèbre est, bien sûr, celui des réseaux neuronaux, dont la paramétrisation utilise typiquement une fonction « coût » et la technique de la descente par gradient (plus connu sous son nom anglais « *gradient descent* »). L'état final du modèle n'est pas fait pour être compris par un humain mais uniquement pour maximiser les performances de l'algorithme. Pour plus de détails, voyez S., SUTHAHARAN, *op.cit.*, spec. pp. 289-301. Toutefois, certaines recherches sont menées dans le but de pouvoir « expliquer » a posteriori le contenu des modèles paramétrés (tels les réseaux neuronaux) ayant aboutis à certains résultats intéressants pour des cas simples. Voyez D. WARNER, « A Neutral Network-Based Law Machine: The Problem of Legitimacy », *Law, Computer & Artificial Intelligence*, Vol. 2, N°2, 1993, pp. 135-147 et F., BORGES *et al.*, « Artificial neural networks and legal categorization », *Legal Knowledge and Information Systems. Jurix: The Sixteenth Annual Conference*. Amsterdam: IOS Press, 2003, pp. 11-20.

<sup>126</sup> Nous pensons ici aux modèles « simples » tels que ceux faisant usage des arbres de décisions ou aux régressions linéaires. Ceux-ci sont toutefois moins utilisés car leurs performances sont souvent inférieures.

<sup>127</sup> Il est important de comprendre que du point de vue technique, prédire et classifier l'inconnu constitue la même activité.

Le premier exemple concerne la prédiction des décisions de la Cour suprême des Etats-Unis et provient d'une étude scientifique réalisée en 2014<sup>128</sup>. Dans ce cas, la collecte des données se fait aisément via la publication des décisions passées de la Cour sur son site internet. Toutefois, ces décisions se trouvent sous forme textuelle et sont fort longues. La deuxième étape de représentation informatique de ces données dans un format que l'ordinateur pourra traiter est donc cruciale : comment représenter ces décisions dans le cadre de notre utilisation du *machine learning* ?

Plusieurs approches sont possibles. L'intuition voudrait que la représentation informatique soit la plus proche possible du contenu conceptuel des décisions, qu'elle reflète le plus possible les éléments de fait et de droit pouvant être retirés de celles-ci. Une telle représentation pourrait utilement faire appel aux travaux de la recherche en *AI & Law* sur la manière de modéliser informatiquement le droit. Une telle approche serait néanmoins particulièrement longue et difficile. Une approche moins chronophage consisterait à résumer les décisions par certains facteurs (factuels et juridiques) clefs, les plus pertinents possibles. Toutefois, l'approche retenue par les chercheurs à l'origine de cette étude a consisté à représenter les décisions de la Cour par les informations présentées sur son site dont, entre autres : le circuit judiciaire d'origine, le type de contentieux, les désaccords entre les Cours inférieures. La liste complète des paramètres pris en considération est illustrée sur la Figure 3. Bien sûr, chaque type d'information doit être représenté de manière numérique. Pour cela, un numéro est attribué aux données non-numériques. Par exemple, le numéro « 1 » est attribué au contentieux en droit pénal, « 2 » au contentieux en droit commercial, et ainsi de suite.

Le choix du modèle et de l'algorithme d'apprentissage, la troisième étape, s'est basée sur des considérations liées aux propriétés de leurs prédictions, détails dans lesquels nous ne rentrerons pas<sup>129</sup>.

---

<sup>128</sup> D., KATZ, *et al.*, *Predicting the Behavior of the United States Supreme Court: A general approach*, 2014, publication sur le site du SSRN à l'adresse: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2463244](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2463244).

<sup>129</sup> Dans notre exemple, l'algorithme utilisé fut la *Random Forests of Decision Trees* (RFDT), pour des raisons liées à l'équilibre entre simplicité d'implémentation, variance de la prédiction et réduction de la surinterprétation. Pour des explications concernant ce choix, nous renvoyons le lecteur vers K., ASHLEY, *op.cit.*, spec. p. 116 et le livre de S., SUTHAHARAN précité.



<b>Case Information</b>	<b>Justice and Court Background Information</b>
Admin Action [S]	Justice [S]
Case Origin [S]	Justice Gender [FE]
Case Origin Circuit [S]	Is Chief [FE]
Case Source [S]	Party President [FE]
Case Source Circuit [S]	Natural Court [S]
Law Type [S]	Segal Cover Score [SC]
Lower Court Disposition Direction [S]	Year of Birth [FE]
Lower Court Disposition [S]	
Lower Court Disagreement [S]	
Issue [S]	
Issue Area [S]	
Jurisdiction Manner [S]	<b>Trends</b>
Month Argument [FE]	Overall Historic Supreme Court [FE]
Month Decision [FE]	Lower Court Trends [FE]
Petitioner [S]	Current Supreme Court Trends [FE]
Petitioner Binned [FE]	Individual Supreme Court Justice [FE]
Respondent [S]	Differences in Trends [FE]
Respondent Binned [FE]	
Cert Reason [S]	

**Figure 3. Exemple des catégories prises en compte pour représenter informatiquement les décisions de la Cour suprême des États-Unis d'Amérique<sup>130</sup>**

Une fois les données collectées et transformées et le modèle choisi, l'algorithme paramétrise le modèle choisi en fonction des données d'apprentissage (en l'occurrence, toutes les décisions passées de la Cour présentes sur son site). Enfin, une fois les paramètres déterminés, la prédiction devient possible. Pour cela, il faut introduire dans le modèle paramétrisé les caractéristiques d'un litige encore pendant devant la Cour, à l'exception de la solution du litige (qui, par définition, n'est pas encore connue). Le modèle, sur base des données numériques de représentation du litige pendant, calculera alors la classification ce nouveau litige dans une des catégories « rejeté » ou « accueilli »<sup>131</sup>, ce qui constitue la prédiction.

Le deuxième exemple concerne l'analyse textuelle de décisions de justice dans le cadre de la recherche de décisions passées pertinentes pour un cas donné<sup>132</sup>. Notre exemple concerne la « *history task* » qui consiste à retrouver, dans une base de données, les décisions de justice liées à un même cas mais provenant de différents niveaux de la hiérarchie judiciaire<sup>133</sup>. Pour ce faire, il est utile de pouvoir extraire certaines informations d'une décision de justice. Cette information peut être conceptuelle (comme le type de contentieux), ou peut consister en un ensemble de métadonnées (comme les noms des parties, la date de la décision, etc). De manière générale, le même genre de technique permet d'extraire certains types d'informations sémantiques (généralement simples) de textes juridiques<sup>134</sup>. Une fois ces données extraites, il

<sup>130</sup> Figure tirée de K., ASHLEY, *op.cit.*, spec. p. 115.

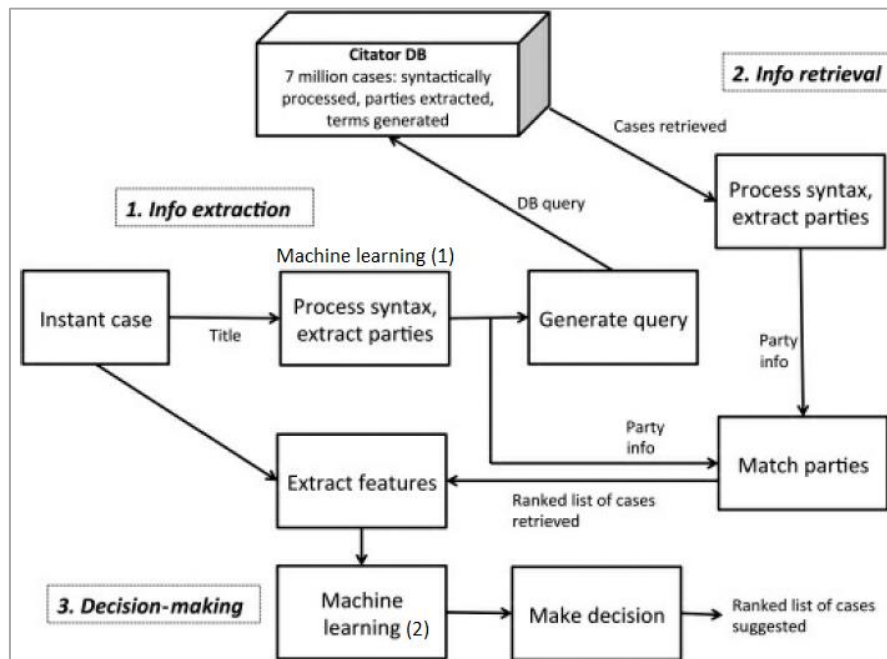
<sup>131</sup> Dans le papier de recherche initial, la prédiction sur base de cette méthode obtient un taux de réussite prédictif de 69,7%.

<sup>132</sup> Ce qui nous permet d'illustrer les faces « recherche » et « analyse » de la justice analytique en une seule fois.

<sup>133</sup> Par exemple, lorsqu'une décision partielle a fait l'objet d'un appel.

<sup>134</sup> Par exemple, des informations concernant les arguments, voyez K., ASHLEY, *op.cit.*, spec. p. 238.

faudra d'abord rechercher, dans la base de données, les décisions correspondant à ces informations avant de décider si les décisions trouvées sont pertinentes dans le cadre de la *history task*. Le schéma des étapes nécessaires est illustré Figure 4.



**Figure 4. Illustration des étapes d'extraction de décisions faisant partie de « l'histoire » du cas considéré. Les deux étapes de *machines learning* sont l'extraction d'information sémantique (1) et la prise de décision (2) concernant la pertinence des décisions trouvées dans la base de donnée (ici, « *Citor DB* »)<sup>135</sup>**

Ici aussi, la collecte des données se fera via les bases de données de jurisprudence d'éditeurs connus. La transformation de ces textes dans un format à la fois représentatif et pouvant être traité informatiquement constitue un bien plus grand défi. De plus, ici, il s'agit, d'une part, d'extraire les méta-informations de manière automatique<sup>136</sup> et, d'autre part, de prendre une décision quant à la pertinence des décisions trouvées dans la base de données. Il nous faut donc choisir un format qui permette de représenter le texte des décisions. La représentation de celles-ci sera en réalité différente dans les deux utilisations du *machine learning* mises en avant ci-dessus, bien que dans les deux cas, il s'agira de transformer le format textuel d'une décision juridique en « vecteur de propriétés » de cette décision. Puisque les deux utilisations transformeront le texte en vecteur, celles-ci utiliseront les mêmes modèles et algorithmes

<sup>135</sup> *Ibid.* spec. p. 235.

<sup>136</sup> En réalité, dans la plupart des bases de données des éditeurs connus, les métadonnées et les données conceptuelles simples (comme le type de contentieux) sont déjà présentes pour toutes les décisions, celles-ci étant bien souvent rentrées dans l'ordinateur à la main. L'emploi du *machine learning* pour extraire ce type de données est donc parfois réaliste (les éditeurs les plus avancés automatisant le processus) même si préalable au problème posé dans l'exemple et parfois inutile. L'explication ayant principalement un but illustratif, nous nous épargnons de développer ces arguties.

d'apprentissage. La méthode de prédiction une fois l'apprentissage effectué sera également conceptuellement la même que celle expliquée dans le cadre de notre premier exemple<sup>137</sup>. Nous nous concentrerons sur les choix de représentations et les différences des deux « vecteurs de propriétés ».

Un vecteur est un type d'objet mathématique qui possède certaines caractéristiques facilitant son traitement par un ordinateur et qui se présente sous la forme d'une juxtaposition d'éléments. Par exemple : (1, 2, 8) ou (4.45, 0, 56)<sup>138</sup> sont deux vecteurs de nombres. Le but de l'opération est donc de représenter chaque décision de justice sous cette forme, où les chiffres entre les virgules sont censés représenter des propriétés de la décision.

Pour ce faire, le texte subi d'abord un traitement linguistique<sup>139</sup> :

- Les mots inutiles (« le », « la », « un »,...) ainsi que la ponctuation sont éliminés.
- Les majuscules sont transformées en minuscules et les mots sont remplacés par leur racine (« nager » et « nageant » deviennent « nage »).
- Certaines informations peuvent être ajoutées (telles que le rôle grammatical des mots)

Une fois ce traitement effectué, le document est représenté sous la forme d'un vecteur de propriétés qui doit être choisi selon l'application. Pour la première utilisation du *machine learning* de l'exemple, le vecteur le plus fréquemment utilisé conviendra puisqu'il s'agit d'une tâche très générale d'extraction d'information. Il consiste dans un vecteur aussi long que le nombre de mots (traités comme décrit ci-dessus) existants dans le corpus entier de tous les textes de décisions utilisés pour la recherche. Pour chaque mot que le corpus contient, si ce document le contient, un « 1 » se trouve à la position de ce mot, et sinon un « 0 » s'y trouve<sup>140</sup>. En plus de ces éléments, le vecteur contient souvent de l'information supplémentaire, comme l'indicateur « Tf/Idf », défini section 2.2.4, de certains mots importants.

Pour évaluer la pertinence, lors de la deuxième utilisation du machine learning, certaines propriétés particulières, en lien avec l'application de l'exemple (donc le traçage historique d'un cas particulier) devront faire partie du vecteur de propriété. Dans le cadre du projet dont

---

<sup>137</sup> L'algorithme *Support Vector Machines* est très utilisé pour les vecteurs de propriété de textes et le fut lors de la création concrète de l'algorithme dont l'exemple présenté s'inspire. Pour plus d'explications sur cet algorithme voyez K., ASHLEY, *op.cit.*, spec. p. 236.

<sup>138</sup> Ces exemples de vecteurs peuvent s'interpréter comme des points de l'espace à trois dimensions, mais le chiffre entre les virgules peut avoir été obtenu par n'importe quel calcul. Le « sens » de ces chiffres est d'ailleurs totalement absent de sa représentation machine.

<sup>139</sup> La première étape est appelée en anglais la « *tokenisation* », la deuxième la « *normalisation* » et la troisième, facultative, l'« *annotation* ».

<sup>140</sup> La dimension de ce vecteur, c'est-à-dire le nombre d'éléments qu'il contient, est donc très grande.

l'exemple s'inspire<sup>141</sup>, un vecteur à huit dimensions fut utilisé, reflétant, par exemple, les propriétés suivantes des textes<sup>142</sup> : similarité des titres, citation de la décision initiale par la décision de la base de données, estimation de l'importance de l'information contenue dans le titre, probabilité de pertinence a priori, .... C'est donc sur base des vecteurs assemblés à partir de mesures construites et partagées par les deux textes (et donc n'en représentant aucun des deux) que l'algorithme de *machine learning* va paramétrer son modèle.

### 3.1.3. *Commentaires sur les données*

La première étape, « collecte des données », peut sembler triviale à première vue. S'il est clair que le *machine learning* ne peut fonctionner sans données, il est important de bien comprendre à quel point celles-ci sont cruciales. Car, en effet, si le modèle choisi et l'implémentation algorithmique sont les moyens qui permettent d'extraire des tendances des données, ces dernières sont celles qui contiennent l'information sur laquelle toute l'opération repose. Dès lors, si les données ne sont pas suffisamment abondantes et/ou suffisamment pertinentes, les prédictions obtenues peuvent être pires que mauvaises.

Concrètement, cela signifie que pour aboutir à une prédiction valide, l'ensemble des données d'entraînement collectées doit avoir les propriétés suivantes<sup>143</sup> :

- Contenir des données en quantité suffisante pour que la modélisation statistique ait un sens.
- Contenir des données dont on peut raisonnablement supposer qu'elles puissent être expliquées mieux que par le hasard<sup>144</sup>.
- Contenir des données pertinentes vis-à-vis du phénomène étudié.
  - ✓ Causalement pertinente.
  - ✓ Temporellement pertinente, donc à jour.
  - ✓ Factuellement pertinente, donc sans erreur.
- Contenir des données qui ne présentent pas, elles-mêmes, de biais statistiques<sup>145</sup>
  - ✓ Pas de surreprésentation d'un phénomène.

---

<sup>141</sup> Le « *History Project* », voyez K., ASHLEY, *op.cit.*, spec. p. 235.

<sup>142</sup> Pour plus de détails sur les quantités utilisées, voyez *ibid*, spec. p. 326. Rappelons également que dans le cadre de la deuxième utilisation du machine learning, il est question de l'évaluation de la pertinence d'un texte retrouvé dans une base de données à partir d'une décision initiale. Il y a donc bien deux décisions de justice.

<sup>143</sup> Voyez S., SUTHAHARAN, *op. cit.*, spec. pp. 62-66 et J. HURWITZ ET D. KIRSCH, *op.cit*, spec. p. 8.

<sup>144</sup> Les données doivent donc présenter une certaine « homogénéité causale », il ne sert par exemple à rien de rassembler des données provenant de phénomènes extrêmement différents et espérer obtenir une prédiction ou une classification fiable à l'aide du *machine learning*

<sup>145</sup> Pour plus d'informations sur ce point, voyez M. AIKENHEAD, « The Uses and Abuses of Neural Networks in Law », *Santa Clara High Technology Law Journal*, Vol. 12, N°1, 1996, pp. 31-70, spec. pp. 67-70.

- ✓ Pas de données partiellement incomplètes vis-à-vis d'un phénomène.

Bien sûr, ces propriétés peuvent être respectées par les données à des degrés divers, plutôt que simplement de manière binaire. Toutefois, cette liste permet déjà de voir certains des points clefs qui poseront question dans l'utilisation du *machine learning* sur des décisions de justice.

#### 3.1.4. *Portée concrète de l'implémentation sur le droit et le monde juridique*

Malgré la généralité de la présentation des sections 3.1.1 et 3.1.3, celle-ci est déjà suffisante pour en tirer certaines conclusions concernant l'impact de l'implémentation du *machine learning* supervisé d'un point de vue juridique, indépendamment de l'utilisation qui en est faite.

Tout d'abord, le caractère crucial des données et les différentes caractéristiques qu'elles doivent présenter va avoir un impact majeur.

En effet, le besoin de collecter des données en suffisance se traduit, en droit, par les débats déjà entamés sur l'« *open data* » (un des aspects du *big data*) et, plus particulièrement, sur la publication systématique des décisions de justice<sup>146</sup>. En effet, ces jugements équivaldraient à plus de données, donc une meilleure paramétrisation des modèles utilisés et donc, en théorie, des performances améliorées. De plus, les jugements non publiés constituent potentiellement des biais dans la représentativité des données. C'est pour cette raison que, si l'utilisation de la justice analytique se répand, elle sera précédée et accompagnée par une plus grande accessibilité des décisions de justice. Or à l'heure actuelle, aucun système centralisé de gestion des décisions de justice n'existe<sup>147</sup>. Une utilisation accrue de la justice analytique, particulièrement de ses facettes prédictives et statistiques, entraîne donc implicitement une amplification du débat sur une réforme de la gestion et de la publication des jugements<sup>148</sup>. Bien sûr, le débat sur la publication des décisions de justice est indissociable de celui concernant leur anonymisation<sup>149</sup>. Sur ce point, la présentation donnée ci-dessus ne nous informe que sur le fait qu'une plus grande anonymisation implique de plus faibles possibilités d'analyse<sup>150</sup>.

---

<sup>146</sup> Voyez, par exemple, le rapport de la commission de modernisation de l'ordre judiciaire - *Rapport consacré à la question de la publication des décisions judiciaires : La plume, le Pélikan et le nuage*, 2014, disponible en ligne à l'adresse <https://justice.belgium.be/sites/default/files/downloads/Rapport%20publication%20des%20d%C3%A9cisions%20judiciaires.pdf>.

<sup>147</sup> Voyez G., QUOISTIAUX, « L'intelligence artificielle peut-elle sauver la justice ? », *Trends tendance*, octobre 2016. Disponible en ligne à l'adresse <http://trends.levif.be/economie/lawyerz/l-intelligence-artificielle-peut-elle-sauver-la-justice/article-normal-603501.html>, dernière consultation le 12 avril 2018.

<sup>148</sup> Ou de tout autre type de documents d'intérêts non publiés à l'heure actuelle.

<sup>149</sup> Voyez le dossier paru dans le journal « Le Monde » : H., GUILLAUD, *La justice prédictive (3/3) : entre performance et optimisation*, septembre 2013. Disponible à l'adresse <http://internetactu.blog.lemonde.fr/2017/09/13/la-justice-predictive-23-predictions-et-regulations/>

<sup>150</sup> Elle n'est néanmoins pas très utile pour déterminer le juste équilibre entre les deux.

De plus, le fait que les données doivent être temporellement pertinentes, c'est-à-dire mise à jour, pose la question du traitement des changements du droit, que ce soit par revirements de jurisprudence ou changements législatifs. En effet, les décisions de justices datant d'avant ces changements conservent-elles leur pertinence ? Cette question ne se pose, encore une fois, que vis-à-vis de la face statistico-prédictive de la justice analytique. En effet, dans le cas d'un algorithme de traitement automatique des documents (par exemple des contrats), cette question n'est pas vraiment pertinente. Elle est par contre très problématique pour l'analyse statistique ou pour les prédictions algorithmiques. Actuellement, Predictice laisse à l'utilisateur le soin de faire le tri dans ses statistiques entre ce qui est pertinent d'un point de vue temporel et ce qui ne l'est plus, option qui n'est toutefois pas réellement viable pour les prédictions de l'issue de litiges. Cette considération dépend du volume de la production judiciaire annuelle et de la quantité de décisions nécessaires pour obtenir de bonnes performances prédictives. Si plus de décisions que ce qui est nécessaire à la qualité de la prédiction sont prises (et publiées) chaque année, l'influence des changements législatifs ou des revirements jurisprudentiels ne sera pas présent. Il suffira en effet, de retirer manuellement les décisions devenues obsolètes de l'ensemble des données d'entraînement. Dans le cas contraire néanmoins, celle-ci ne pourront être aussi simplement retirées, laissant entière la question de la gestion technique de cette situation.

Ensuite, l'impact de la deuxième étape de représentation informatique des données brutes est probablement encore plus important que celui liées aux considérations relatives à la collecte des données. Comme l'illustrent les deux exemples, les données à représenter ne sont généralement pas utilisables telles quelles et le choix du format de représentation va avoir une influence déterminante sur les tendances qui seront « apprises » par le modèle. Dans le cadre de la production de statistiques concernant les décisions juridiques ou de prédictions algorithmiques, le processus de représentation va mener à la prise en compte de « l'opinion » du créateur de l'algorithme. Cet impact est particulièrement bien illustré par l'aspect intuitivement « étranges » et « synthétiques » des objets mathématiques que les algorithmes du deuxième exemple de la section 3.1.2 prennent pour base. Cette étape, combinée avec le processus d'optimisation lié au choix du modèle et à sa paramétrisation<sup>151</sup>, aboutit inéluctablement à ce que les résultats produits par les algorithmes de *machine learning* reflètent

---

<sup>151</sup> Qui, comme cela fut décrit ci-dessus, résulte aussi d'un choix relatif à l'objectif de généralisation des tendances apprises sur les données d'apprentissage aux données « tests ». Ce choix aura également une influence sur l'apprentissage et sur la pondération de chaque décision dans le résultat final (et donc de son importance). Cet effet est toutefois, vraisemblablement, du deuxième ordre et plus difficile à déterminer précisément.

implicitement les choix des créateurs de l'algorithme quant à ce qui est pertinent<sup>152</sup>. Dans le cadre de Predictice ou de Case Law Analytics, par exemple, leur représentation des décisions de justice sur base des catégories de facteurs légalement et factuellement « pertinents ». Ceci sera, bien sûr, également le cas dans le cadre d'un traitement automatique de texte (extraction de métadonnées, d'arguments ou de types de clauses particulières). Toutefois, dans ce genre d'utilisation plus restreint dans leur objectif, cela ne semble pas être particulièrement important. Etant donné qu'ils incorporent mécaniquement les conceptions théoriques et idéologiques de leur(s) créateur(s)<sup>153</sup>, il est incohérent de penser que ces algorithmes possèdent une plus grande « objectivité » ou une plus grande « neutralité politique » que les êtres humains. Ce point est crucial, car on peut fréquemment lire dans la presse ou la doctrine des sentiments tels que « *[u]ne intelligence artificielle rendrait une meilleure justice qu'une personne influencée par sa propre idéologie et le contexte du jugement* »<sup>154</sup>. De même, rappelons que l'algorithme paramétrisant le modèle a comme objectif de « coller » au mieux aux données. Or dans le cadre de la face prédictive de la justice analytique, ces données sont des décisions de justice rendues par des humains. Comprendre cela est essentiel, car selon Antoine Garapon s'exprimant dans le cadre d'un article sur les enjeux de « justice prédictive »<sup>155</sup> : « *[il est un] fait avéré qu'aujourd'hui, le public est plus rassuré par une vérité établie mathématiquement et scientifiquement que par une décision humaine* ». Or les résultats des algorithmes de *machine learning* à visée prédictive sont précisément conçus pour refléter au plus près les décisions humaines ! La plus grande précision souvent vantée de la prédiction par algorithme vient précisément du fait qu'elle parvient à détecter dans les jugements *plus* que le droit positif et qu'elle reflète donc toutes aspérités *humaines* de la jurisprudence et donc tous ses défauts.

Comprendre que le créateur de l'algorithme imprime inéluctablement ses conceptions dans

---

<sup>152</sup> Voyez K., ASHLEY, *op.cit.*, spec. p. 222. Ce constat se combine au constat général concernant l'absence de neutralité des procédés techniques. La philosophie (notamment avec les réflexions Simondonienne sur les objets techniques) a déjà mis en avant que tous objets présentent en lui ses propres logiques et ses propres finalités. Il en va bien sûr de même de la justice analytique qui arbore les présupposés de ses inventeurs quant au besoin de rapidité du monde économique. Voyez T., REIGELUTH, *op. cit.*, spec. p. 41 et suivant. Ce constat dépasse toutefois celui-ci en ce qu'il inclut une deuxième couche d'influence de conceptions extérieurs à l'utilisateur, en plus de celle du concepteur de l'objet, celles du créateur de l'algorithme aura également son importance. En réalité ce constat généralise les critiques qui ont déjà été adressés aux systèmes de prédictions de la probabilité de récidives, voyez le dossier paru dans le journal « Le Monde » : H., GUILLAUD, *La justice prédictive (3/3): prédictions et régulations*, septembre 2013. Disponible à l'adresse <http://internetactu.blog.lemonde.fr/2017/09/13/la-justice-predictive-23-predictions-et-regulations/>

<sup>153</sup> Où de la synthèse des consultations réalisées auprès de la communauté juridique, comme c'est le cas pour Predictice et Case Law Analytics.

<sup>154</sup> Propos tenus par Jean-Pierre Buyle et rapportés dans G., QUOISTIAUX, *op. cit.*, spec. p. 3.

<sup>155</sup> A. GARAPON, « Les enjeux de la justice prédictives », *La semaine juridique*, hebdomadaire n°1-2, janvier 2017, spec. p. 5.

celui-ci permet de réaliser l'importance de l'aspect sociologique de l'évolution de la pratique du droit<sup>156</sup>. Cela permet également de remettre en question (voire de critiquer) le fait que tous ces algorithmes ont pour origine le secteur privé. Enfin, cela permet de démystifier un peu plus ces techniques et de réaliser qu'il s'agit de bien de *méthodes*, de procédés proches des statistiques, plus qu'il ne s'agit d'une « compréhension » ou « d'intelligence » qui s'émanciperait des défauts humains et, ce, malgré la désignation de celles-ci par l'expression « intelligence artificielle ». Réaliser tous cela ne remet, bien sûr, aucunement en question les performances de ces algorithmes.

Enfin, le détail de la combinaison des étapes 3 et 4 a également un autre impact important. Le fait que le résultat de l'apprentissage de la plupart des modèles de *machine learning* les plus efficaces ne soit pas réellement compréhensible rend la suggestion fréquente de « rendre les algorithmes publics » moins pertinente qu'il n'y paraît. S'il est important de pouvoir s'assurer que les algorithmes ne sont pas « truqués » dans un sens ou dans l'autre, peu de choses peuvent être réellement apprises de la publication de l'algorithme en tant que tel. Une clarification concernant la représentation des données, en revanche, serait utile et pourrait même faire l'objet d'arguments juridiques en soi.

### **3.2. Analyse de l'utilisation par les acteurs du droit**

Dans cette section, nous procédons à l'analyse des conséquences de l'utilisation de la justice analytique. Pour ce faire, nous aborderons l'étude de son utilisation par différents acteurs selon une structure en trois « sphères ». La première sphère abordée sera la sphère politique, constituée par les gouvernements, parlements et politiciens nationaux ainsi que par les organisations internationales. La deuxième sphère que nous aborderons sera celle des milieux du droit au sens restrictif. Celle-ci se décompose en trois sous-entités d'acteurs ayant un impact direct sur l'application du droit : les juges, les avocats et le monde académique. Enfin, la dernière sphère sera celle des usagers, utilisateurs ou consommateurs du droit : les particuliers et les entreprises.

Pour chacune de ces sphères et chacun de ces acteurs, nous mettrons notre hypothèse de reconfiguration des pratiques à l'épreuve en tentant de déduire des pratiques actuelles l'impact de l'existence des nouvelles possibilités offertes par l'utilisation des nouveaux outils. Nous considérerons aussi bien les conséquences liées à leur usage direct que les conséquences

---

<sup>156</sup> A l'initiative de laquelle ne se trouvent pas des juristes mais des ingénieurs, des informaticiens et des hommes d'affaire, voyez *ibid*, spec. p. 1.



résultant des attentes collectives vis-à-vis de ces technologies.

Pour ce faire, une hypothèse de travail est incontournable, sous peine de rendre l'entreprise brève et inutile. Il nous est, en effet, nécessaire de postuler un usage répandu des techniques sous-jacentes aux outils commercialement offerts à l'heure actuelle par les *legaltechs* (que celui-ci advienne par leur intermédiaire ou non). Dans le cas contraire, soit il ne nous est pas possible de penser les conséquences d'un usage inexistant soit il nous est trivial de conclure à l'impact insignifiant d'une utilisation négligeable. Il est, en effet, clair que l'impact pratique, si pas psychologique, actuel des *legaltechs* utilisant l'intelligence artificielle en droit telle que nous l'avons définie est insignifiant<sup>157</sup>. Sauf peut-être à chercher les conséquences sur l'imaginaire des juristes résultant du battage médiatique entourant ces technologies, postuler que l'usage de celles-ci ne se répandra pas mène à la conclusion que leur impact restera constant et donc négligeable.

Bien qu'elle semble inévitable, il est important de noter dès à présent que cette hypothèse est critiquable. Postuler l'usage répandu de ces technologies n'est-il pas équivalent à s'écarter de notre objectif de démystification ? Notre réponse à cette critique est un « non » prudent. Tout d'abord, nous avons limité considérablement l'objet de l'hypothèse en question et nous nous sommes assuré que celui-ci constitue l'utilisation réaliste de l'intelligence artificielle en droit aussi bien d'un point de vue théorique que pratique. Dès lors, même s'il n'est pas certain que ces technologies vont se répandre dans tous les milieux que nous considérerons ci-dessous<sup>158</sup>, nous ne nous pencherons que sur un objet dont la réalité pratique est clairement établie. Nombre des conséquences déduites ci-dessous ne sont d'ailleurs pas dépendantes d'une utilisation généralisée de ces nouvelles technologies. Dès lors, plutôt que de prendre l'analyse ci-dessous pour une prédiction du futur, il est plus prudent de la considérer comme un catalogue des impacts *potentiels* de l'utilisation de ces technologies<sup>159</sup> par les différents groupes d'acteurs considérés. Nous laissons donc au lecteur l'appréciation de la probabilité que lesdites technologies soient, à court, moyen ou long terme, utilisées par ceux-ci.

### 3.2.1. *Le point de vue de la sphère politique*

L'utilisation de la justice analytique par les acteurs de la sphère politique peut entraîner des

---

<sup>157</sup> Ne serait-ce que par leur quasi-absence en Belgique !

<sup>158</sup> Il n'est, par exemple, pas évident de déterminer quelle force, entre la réticence des milieux juridiques à la nouveauté et les tendances à l'investissement dans ces nouvelles technologies, va prendre le dessus dans les différentes sphères considérées, du moins à court terme.

<sup>159</sup> Notez que nous ne considérons pas non plus les interactions de ces nouvelles techniques avec les représentations informatiques du droit développées par la recherche en *AI & Law*, ou avec les autres technologies utilisées par les autres *legaltechs*.

effets de quatre ordres différents, ceux-ci étant liés entre eux de la manière suivante. L'effet structurant concerne la création d'une nouvelle panoplie d'indicateurs statistiques sur la justice belge dont découlera probablement, d'une part, une modification des interactions entre pouvoir législatif, exécutif et judiciaire et, d'autre part, une intensification de certaines pratiques de mise en compétition entre les pays par des organismes internationaux. La création de ces indicateurs statistiques se fera en parallèle d'une volonté d'un plus grand partage des données. Comme discuté section 3.1.4, l'utilisation de la justice analytique renforcera la volonté que les décisions de justice constituent un *open data* et inversement. Enfin, la sphère politique ne pourra manquer de réagir aux changements qu'induiront l'utilisation de la justice analytique dans les milieux des professionnels du droit. Ce dernier point sera toutefois abordé lors de la discussion desdits changements.

#### 3.2.1.1. Une réification de « La Jurisprudence »

La quantification d'un domaine qui, jusqu'alors, n'était décrit et appréhendé que par les mots n'est jamais un phénomène anodin<sup>160</sup>. Cette affirmation peut être illustrée, par exemple, par le concept de « l'intelligence » qui se développa en parallèle de notion de quotient intellectuel ou de « l'opinion publique » qui prit son essor avec les sondages (par exemple, ceux de l'institut *Gallup*)<sup>161</sup> ou encore à la notion de « réputation » des universités, qui prit un essor considérable avec l'apparition du classement de Shanghai. L'idée que la justice analytique ne fera *que* « refléter » ou « mesurer » une réalité sociale préétablie oublie donc que la quantification crée une nouvelle façon de concevoir et d'interagir avec le monde social et, cela, indépendamment de la fidélité avec laquelle cette dernière reflète son objet d'étude<sup>162</sup>.

Puisque la justice analytique offre de nouvelles possibilités de quantification statistiques, elle ne pourra pas laisser l'imaginaire collectif indemne et, ce, particulièrement au niveau politique. De plus, si « quantifier, c'est convenir puis mesurer »<sup>163</sup>, force est de constater que la définition de la convention de ce qui est à mesurer est ici laissée, en tout cas initialement, à l'appréciation du secteur privé des *legaltechs*. Or, une fois les méthodes quantificatives établies et devenues routinières, leurs produits sont naturalisés : devenus « réalité » par l'effet de l'habitude et de

---

<sup>160</sup>A., DESROSIÈRES. *Pour une sociologie historique de la quantification : L'Argument statistique I*. Nouvelle édition, Paris, Presses des Mines, 2008, spec. n°7.

<sup>161</sup>*Ibid*, spec. n°8.

<sup>162</sup>*Ibid*, spec. n°9.

<sup>163</sup> Avant de pouvoir « mesurer » des quantités numériques liées à un phénomène social, il convient d'abord de s'accorder sur la grandeur pertinente à étudier, voire de les créer de toute pièce en commun. Voyez A., DESROSIÈRES, *op. cit.*, spec. N° 9-13.

leur utilité. La période de transition menant vers l'utilisation des techniques de justice analytique à des fins de productions statistiques intensives de la jurisprudence marquera donc un tournant dans la conception publique<sup>164</sup> de ce qu'est la jurisprudence. Si les conséquences d'un tel changement ne sont pas prévisibles dans leur ensemble, on peut déduire le prolongement qu'auraient certaines pratiques existantes des gouvernants, une fois les statistiques produites par la justice analytique à disposition.

### 3.2.1.2. De nouvelles interactions entre pouvoirs

La première pratique dont le prolongement s'envisage sans difficulté est celle du « management des juges » par le, ou à l'instigation du, pouvoir exécutif. Cette pratique, qui a déjà fait l'objet d'une analyse doctrinale attentive<sup>165</sup>, constitue une branche des diverses réformes entreprises sous l'impulsion du « *New Public Management* » avec pour objectif une « efficacité » accrue de l'organisation du pouvoir judiciaire et de l'administration de la justice. Concrètement, il s'agit d'un ensemble de dispositifs d'évaluation, de contrôle et de gestion de la production judiciaire et de ses producteurs.

Tout d'abord, et trivialement, en supposant que les outils de la justice analytique soient mis à la disposition des juges, ceux-ci pourront s'attendre à voir leur productivité mise sous une pression supplémentaire.

De plus, et crucialement, si la justice analytique ne perturbe en rien les normes managériales de gestion, force est de constater qu'elle possède un potentiel de démultiplication des capacités d'évaluation et de contrôle. Plus précisément, elle ouvre la porte vers une révolution des critères d'évaluation de la qualité des décisions de justice. Certaines des possibilités qu'offre son utilisation viennent presque instinctivement à l'esprit : vérifier la conformité des décisions rendues aux prédictions de l'algorithme, comptabiliser le nombre de décisions réformées en appel ou cassées en cassation<sup>166</sup>, établir des statistiques liées à la « productivité d'un juge »

---

<sup>164</sup> C. EYRAUD, « Réflexions pour une sociologie de la quantification statistique et comptable », *Entreprises et histoire*, Vol. 79, N°2, 2015, pp. 74-87, spec. p. 74.

<sup>165</sup> Voyez, par exemple, le livre de B., FRYDMAN et E. JEULAND, *Le nouveau management de la justice et l'indépendance des juges*, Bruxelles, Dalloz, 2011 contenant plusieurs contributions sur le sujet ou encore les articles suivants : J. ALLARD, « L'indépendance de la justice et le management judiciaire : contre le « pouvoir » ? La ruse de la raison économique » *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014, pp. 33-50, K., MUNUNGU LUNGUNGU, « Le management judiciaire ou le glas de l'indépendance du juge ? Un éclairage de droit constitutionnel », *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014, pp. 126-167, F., SCHOENAERS, « Lorsque le management entre au tribunal : évolution ou révolution ? », *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014, pp. 171-209 ou le livre T. MARCHANDISE *et al.*, *Quel management pour quelle justice?*, Larcier, 2013.

<sup>166</sup> Une pratique déjà envisagée, mais rendue infiniment plus simple à l'aide de la justice analytique. Voyez B., FRYDMAN, *L'évolution des critères et des modes de contrôle de la qualité des décisions de justice*, Working Papers du Centre Perelman de Philosophie du Droit, Vol. 4, 2007. Publié en ligne à l'adresse :

exprimée en nombre de pages et/ou via la complexité de ses jugements, etc. La justification de ces outils pour créer ce type de critères ne serait d'ailleurs pas difficile. En effet, la justice analytique se basant sur les décisions de justice passées, il serait simple d'utiliser l'argument du besoin de sécurité juridique pour justifier d'évaluer la production juridique présente sur base de celle du passé<sup>167</sup>. Toutefois, et bien que cette utilisation de la technologie ne soit encore présente qu'en germe, l'utilisation de la justice analytique peut également s'appliquer *aux arguments* présents dans les jugements<sup>168</sup>. Il devient alors, en théorie, possible d'établir combien de fois un juge utilise l'article 159 de la Constitution pour écarter l'application d'un règlement illégal, la fréquence de ses références aux principes généraux du droit, etc. Ces possibilités ouvrent donc le champ d'un contrôle potentiel du juge, et des jugements, nettement supérieur à tout ce qui a été mis en place jusque maintenant.

Une autre pratique, mise en œuvre par le « pouvoir politique »<sup>169</sup>, qui risque de se voir transformée par l'utilisation de la justice analytique est celle qui consiste à répondre aux développements récents de la jurisprudence par une production législative. Ceci est, bien sûr, parfaitement légitime : si l'interprétation que le pouvoir judiciaire fait de ses lois heurte le législateur celui-ci est habilité à les modifier, abroger ou remplacer ou à adopter des lois interprétatives<sup>170</sup>. Encore un fois, l'utilisation de la justice analytique ne fait que modifier l'échelle potentielle de ce phénomène. Plutôt que de réagir aux « grandes » décisions des Cours suprêmes de notre pays, la quantification statistique de *toutes* les décisions de justice que rend possible la justice analytique donne la possibilité de réagir à « La Jurisprudence » dans son ensemble<sup>171</sup>. Une pratique intensive de ce genre reviendrait à une sorte de « pilotage législatif » de la jurisprudence. Si rien n'atteste que de telles extrémités soient probables, elles deviennent

---

[www.philodroit.be/IMG/pdf/levolution\\_des\\_criteres\\_et\\_des\\_modes\\_de.pdf](http://www.philodroit.be/IMG/pdf/levolution_des_criteres_et_des_modes_de.pdf)

<sup>167</sup> L'argument est simplement énoncé, et est donc fort politiquement, mais pas forcément pertinent. En effet, la procédure judiciaire représente un équilibre entre des objectifs potentiellement contradictoires : la sécurité juridique et le caractère « juste » de la décision.

<sup>168</sup> Comme déjà mentionnée section 2.1.4, la technologie d'extraction d'arguments à partir de textes juridiques en est encore à ses débuts, mais elle reste techniquement faisable et a déjà été utilisée dans un cadre non-juridique, par exemple via le programme Debater d'IBM. Pour un exemple de ce qui se rapproche de l'état de l'art dans le domaine, voyez R., MOCHALES et M-F., MOENS, « Argumentation mining », *Artificial Intelligence and Law*, Vol. 19, N°1, pp. 11-22. Plus d'informations sur le sujet voyez : K., ASHLEY, *op. cit.*, spec. p. 21 et K., ASHLEY et V., WALKER, « From information retrieval (IR) to argument retrieval (AR) for legal cases: report on a baseline study ». *Proceedings of the 26th Annual Conference on Legal Knowledge and Information Systems 2013*, IOS Press: Amsterdam, pp. 29–38.

<sup>169</sup> Expression qui se réfère ici aux groupes politiques qui forment une majorité dans les différents parlements et constituent donc les gouvernements.

<sup>170</sup> Un exemple célèbre est celui de la réaction parlementaire à l'arrêt Le Compte, voyez M. UYTENDAELE, *Trente leçons de droit constitutionnel*, 2e éd., Bruxelles, Bruylant, 2014, spec. pp. 165-168.

<sup>171</sup> Bien sûr, les statistiques sur les décisions judiciaires existent déjà. Notre argument est toutefois que la différence d'échelle que permet la justice prédictive se transforme, in fine, en différence qualitative.

en tout cas possibles<sup>172</sup>.

Enfin, le niveau international possède ses propres mécanismes de « management », non pas des juges, mais des juridictions. L'exemple le plus caractéristique se trouve dans les indicateurs produits par la Banque Mondiale, « *Doing Business* »<sup>173</sup> (qui contient des catégories comme « facilité de l'exécution forcée d'un contrat ») et « *Rule of Law* »<sup>174</sup>. Ceux-ci classent les Etats en fonction, d'une part, de leur attractivité (du point de vue du capital) et, d'autre part, du respect de l'Etat de droit<sup>175</sup>. Sur ce plan également, une utilisation de la quantification statistique dans un but d'approfondissement de ces pratiques déjà établies semble probable. Or ce type d'indicateur influence les pouvoirs politiques de nombreux pays, renforçant probablement les conclusions présentées sur le pilotage de la jurisprudence et le management des juges.

L'aspect perturbant des pratiques potentielles présentées ci-dessus est qu'elles posent la question du respect des valeurs judiciaires cardinales de notre ordre juridique<sup>176</sup> : l'indépendance du juge, l'indépendance du pouvoir judiciaire, la légalité de la procédure et l'exigence de procès équitable. Nous ne sommes pas les premiers à nous inquiéter d'une diminution de l'indépendance du juge à cause des techniques managériales<sup>177</sup>, mais la généralisation de l'ingérence des pouvoirs exécutif et législatif dans le pouvoir judiciaire rendu possible par ces outils semble donner une dimension nouvelle à ces inquiétudes.

### 3.2.2. *Le point de vue des milieux du droit*

Dans ce point, nous abordons dans l'ordre, l'utilisation de la justice analytique par les juges, les avocats puis par le monde académique.

#### 3.2.2.1. *Juges, justice et le poids de la comparaison.*

L'analyse de l'influence de l'utilisation de la justice analytique par les juges va dépendre du modèle considéré sur la manière dont les juges prennent des décisions. En effet, pour simplifier,

---

<sup>172</sup> Imaginez, par exemple, que les statistiques produites sur « La Jurisprudence » mettent en évidence que dans un contentieux politiquement sensible, un « rôle » (disons, l'employeur) gagne nettement plus souvent que l'autre (disons, le travailleur). Ne deviendrait-il pas alors tentant pour le pouvoir politique d'intervenir pour « rééquilibrer » la jurisprudence ?

<sup>173</sup> Voyez le site internet <http://www.doingbusiness.org/>, qui présente cet indicateur et sa construction.

<sup>174</sup> Voyez le site internet de la Banque mondiale à l'adresse [http://databank.banquemondiale.org/data/reports.aspx?Report\\_Name=WGI&Id=1\\_aa1cf5a](http://databank.banquemondiale.org/data/reports.aspx?Report_Name=WGI&Id=1_aa1cf5a). Dernière consultation : 18 avril 2018.

<sup>175</sup> Pour plus d'informations sur en quoi ces indicateurs constituent une forme de « management », voyez B., FRYDMAN et A. VAN WAEYENBERGE, *Gouverner par les standards et les indicateurs: De Hume au rankings*, Primento, 2013.

<sup>176</sup> Telles que formulées par B., FRYDMAN « Concilier le management avec les valeurs du judiciaire », *Quel management pour quelle justice ?*, T. Marchandise (dir.), Bruxelles, Larcier, 2012, pp. 81-88

<sup>177</sup> *Ibid*, spec. p. 85.

deux écoles s'affrontent : les réalistes et les formalistes<sup>178</sup>. Les premiers considèrent que le juge prend sa décision de manière intuitive et justifie celle-ci *a posteriori* à l'aide d'arguments juridiques tandis que les seconds considèrent que le mode de raisonnement précède la prise de décision et constitue la méthode par laquelle le juge arrive à la solution du litige. De plus, les trois faces de la justice analytique doivent ici être distinguées parce qu'elles interagissent différemment avec les deux modèles du jugement.

L'amélioration des capacités de recherche ne devrait pas faire de différence significative selon les deux modèles. Si certains affirment que l'accélération de la recherche permettra aux juges de se concentrer sur des questions juridiques nouvelles et complexes<sup>179</sup>, les considérations relatives à la productivité et au management des juges développées section 3.2.1.2 rendent cet optimisme quelque peu douteux en Belgique<sup>180</sup>. Le traitement (partiel mais automatique) de dossiers volumineux de documents textuels devrait avoir le même genre d'effet, une augmentation de la productivité des magistrats, sans que le modèle de prise de décision n'entre en compte.

Par contre, en ce qui concerne la face statistico-prédictive de la justice analytique, les détails de son utilisation auront une si grande influence sur les conséquences pouvant en être déduites et les scénarios possibles sont si divers qu'une analyse n'est pas possible sans une hypothèse supplémentaire. En effet, considérons les deux scénarios théoriques suivants. Le premier est celui d'un juge ayant simplement la possibilité d'utiliser un outil tel que Prédicte, qui lui offre une meilleure connaissance statistique des décisions de justice passées. Le deuxième est celui d'un juge ayant l'obligation de justifier avec plus de soin un écart par rapport à la prédiction donnée par un algorithme prédictif<sup>181</sup>. Il est évident que les conséquences pratiques de ces deux scénarios sont incomparables. Afin de ne pas faire de la science-fiction, nous considérerons uniquement l'impact de la consultation de ces outils par le juge. Sous cette hypothèse, et selon le modèle réaliste, la connaissance de la prédiction ou des statistiques concernant les décisions

---

<sup>178</sup> Voyez, pour plus d'explications sur ces deux écoles: T., CAPURSO, « How Judges Judge: Theories on Judicial Decision Making », *University of Baltimore Law Forum*, Vol. 29, N°1, 1998, pp. 5-16 et V. TUMONIS, « Judicial decision-making from an empirical perspective », *Baltic Journal of Law & Politics*, Vol. 6, N°1, 2013, pp. 140-162.

<sup>179</sup> Propos tenus par Jean-Marc Sauvé, Vice-président du Conseil d'État français, lors du du bicentenaire de l'Ordre des avocats français le 12 février 2018. Son intervention fut reproduite sur le site internet à l'adresse : <https://www.eff.fr/actualites/affaires/themes-divers/details.html?ref=r-b8ab7ec2-06eb-4609-9605-b13c78f8c675>

<sup>180</sup> Mais peut-être, autant l'espérer, observera-t-on une diminution du « coût de la justice ».

<sup>181</sup> Ce scénario fut, par exemple, mentionné dans le dossier paru dans le journal « Le Monde » : H., GUILLAUD, *La justice prédictive (1/3): l'enjeu de l'ouverture des données*, septembre 2013. Disponible à l'adresse <http://internetactu.blog.lemonde.fr/2017/09/09/la-justice-predictive-13-lenjeu-de-louverture-des-donnees/>

de justice antérieures devrait avoir pour effet potentiel de modifier<sup>182</sup> l'intuition du juge et, donc, la solution du litige. En effet, comment ne pas se laisser influencer, même partiellement et inconsciemment, par la connaissance de ce que feraient ses pairs dans une situation donnée ? La motivation, elle, devrait rester largement intacte. Dans ce cas, la justice analytique se reposant sur les décisions de justice passées, on peut s'attendre à un renforcement (même léger) de la normativité du précédent et à une homogénéisation<sup>183</sup> de la jurisprudence. On peut également dire que les conceptions incluses implicitement dans les algorithmes de *machine learning* auront ainsi une légère expression dans les décisions de justice puisque ces outils auront influencé le juge. Par contre, selon le modèle formaliste, la consultation de statistiques concernant d'anciennes décisions ou de prédictions algorithmiques ne devrait rien changer à la méthode de résolution d'un litige pour le juge.

L'exposition des deux modèles étant des simplifications permettant de rapidement présenter les deux extrémités du débat, il est possible que certaines parties du jugement (tel que les balances d'intérêts, les estimations de montants de dommages moraux, ...) ou certains contentieux, soient plus susceptibles d'être correctement modélisées par les thèses réalistes, tandis que d'autres le seraient plus par des thèses formalistes. Si tel est le cas, les conclusions valides pour chaque modèle seraient applicables aux parties du jugement *ad hoc*.

Il semble également probable que l'influence des outils prédictifs et statistiques basés sur la justice analytique seront d'autant plus importants dans les affaires « risquées » tel que le jugement de terroristes ou les affaires très médiatisées. Sous pression, il semblerait logique qu'un juge cherche à éviter la critique en s'assurant de juger comme ses pairs l'auraient fait.

#### 3.2.2.2. De nouvelles pratiques chez les avocats

Les avocats sont probablement ceux qui se devront, à cause de la concurrence entre eux, de faire l'utilisation la plus intensive et plus recherchée de la justice analytique<sup>184</sup>. Cette utilisation visera l'affinage des prédictions de l'issue des litiges de leurs clients, l'accélération du traitement de dossiers volumineux dont une partie basique sera automatique et dans une recherche plus rapide et plus ergonomique. Leur utilisation de ces technologies semble d'ailleurs une condition *sine qua non* de leur usage sérieux par d'autres acteurs. Cette utilisation aura trois types d'effets : des conséquences socio-économiques, des questions soulevées d'un

---

<sup>182</sup> Légèrement ou fortement, cela impossible de le prédire.

<sup>183</sup> Cela pourrait être particulièrement le cas des barèmes, tel que ceux liées à l'indemnisation pour licenciement abusif, pour lesquelles les magistrats disposent déjà de grilles officieuses.

<sup>184</sup> Ce sont d'ailleurs eux que les *legaltechs* visent en priorité d'un point de vue marketing.

point de vue déontologique et deux impacts de type juridique.

Du point de vue socio-économique, l'utilisation de la justice analytique semble devoir mener soit vers une diminution du coût moyen d'une consultation, soit vers une escalade dans la « finesse » des arguments juridiques. Dans la première branche de l'alternative, le travail de « faible » valeur ajoutée que sont la recherche juridique et le traitement « sommaire » de dossiers volumineux, qui constituent une partie importante de la fonction des avocats-stagiaires<sup>185</sup>, des avocats en début de carrière et des assistants juridiques<sup>186</sup>, viendra à diminuer. Ce raisonnement semblerait donc conduire à la conclusion, non pas que certains métiers « disparaîtraient » mais plutôt que, si la productivité de ces professions augmente grâce aux nouvelles technologies et que la quantité de travail reste constante, le nombre de poste total diminuera. De tels changements socio-économiques dans la profession d'avocat ne manqueraient d'ailleurs pas d'attirer l'attention du législateur<sup>187</sup>. Dans la deuxième branche de l'alternative, la face « traitement de documents » de la justice analytique, notamment sa composante d'extraction d'information pourrait également être appliquée aux articles doctrinaux, aux travaux préparatoires et à d'autres types de textes juridiques permettant de soutenir des arguments. Les arguments liées à la qualité des outils eux-mêmes et à leurs productions prédictives ou statistiques (notamment relatives aux questions de représentation des données analysées section 3.1.4) pourraient aussi, probablement, faire partie de ces arguments supplémentaires. Dès lors, l'automatisation d'une partie du travail serait potentiellement compensée par le plus grand effort investi dans l'argumentation juridique. Le critère qui fera la distinction entre ces deux alternatives sera la sensibilité des juges au surcroît d'argumentation. En effet, en dehors d'un lien causal entre argumentation accrue et chances de gain, seule la première branche de l'alternative sera viable économiquement.

D'un point de vue déontologique, la face prédictive de la justice analytique peut mettre l'avocat dans deux nouvelles situations de conflit d'intérêts<sup>188</sup>. La première, en réalité, se produit déjà à l'heure actuelle. Il s'agit de la situation dans laquelle les chances de succès d'une action sont

---

<sup>185</sup> Voyez H., SURDEN, *op. cit.*, spec. p. 113.

<sup>186</sup> M.A., NEARY et S., XIN CHEN, « Artificial Intelligence: Legal Research and Law Librarians », *AALL spectrum*, Vol. 21, N° 5, 2017, pp. 16-20, spec. p 17.

<sup>187</sup> Le projet de réforme commandité par l'actuel ministre de la Justice, actuellement en discussion dans les différents barreaux s'affiche déjà comme (entre autres) une réponse préemptive à l'arrivée de l'intelligence artificielle. Voyez la commande sur le site internet du ministre et le projet de réforme, respectivement aux adresses <https://www.koengeens.be/fr/news/2018/02/26/rapport-sur-l-avenir-de-la-profession-d-avocat> et [https://justice.belgium.be/sites/default/files/rapporttoekomstadvocatuurrapportavenirprofessiond\\_avocatfinal.pdf](https://justice.belgium.be/sites/default/files/rapporttoekomstadvocatuurrapportavenirprofessiond_avocatfinal.pdf). Pour la partie sur l'intelligence artificielle du projet de réforme, voyez notamment les pages 93 à 106.

<sup>188</sup> Situation où l'intérêt de l'avocat et celui de son client divergent fortement.



faibles. Du point de vue de l'avocat, un client est un client, et il ne sera pas moins bien payé si l'affaire est perdue d'avance. Toutefois, l'arrivée de la prédiction algorithmique apportera une couche de complexité supplémentaire à ce genre de situation : que faire lorsque l'avis de l'avocat et la probabilité calculé au moyen de l'algorithme divergent ? Le deuxième type de situation concerne l'interaction entre la prédiction algorithmique et un autre phénomène de mode : la notation des avocats<sup>189</sup>. Ici, le conflit d'intérêt se trouve alors dans la prise en charge d'affaires « difficiles ». En effet, si les finances ne sont pas trop mauvaises, ne vaudrait-il pas mieux pour l'avocat prendre prioritairement des affaires « faciles », celles dont le score prédictif est suffisamment élevé pour ne pas prendre le risque ? Ceci pourrait être encore exacerbé dans les associations d'avocats si celles-ci décidaient d'utiliser la face statistico-prédictive de la justice analytique dans le but de « mesurer » les performances de leurs membres<sup>190</sup>. Ces pratiques pourraient-elle être autorisées d'un point de vue déontologique ? Des réponses à ces questions devront vraisemblablement être apportées par les barreaux. Or, les réponses données auront une influence importante sur le nombre de cas où la justice analytique aura un effet performatif sur le client, effet analysé section 3.2.3.1.

Enfin, deux phénomènes potentiels, d'ordre juridique, peuvent résulter d'une utilisation intensive de la justice analytique. Le premier consiste dans un ralentissement de l'évolution conceptuelle du droit en pratique. En effet, rappelons que le développement des technologies de justice analytique est coûteux en temps et en argent. Or, comme expliqué section 3.1.4, les algorithmes fonctionnent sur des bases qui sont le reflet de ce que le créateur de l'outil considère comme pertinent. Ceux-ci se fondent donc, sur des concepts, des catégories et des objets juridiques que les *humains* ont tenus pour importants et pertinents à un *moment donné*<sup>191</sup>. Si l'utilisation de ces outils se répand parmi les praticiens du droit, ceux-ci feront donc implicitement usage de ces catégories, objets et concepts. Toutefois, il est vrai que la plus grande partie du travail de la grande majorité des avocats n'est pas de réinventer la roue et de faire progresser la théorie du droit. Ce genre d'évolution se conçoit d'ailleurs souvent en

---

<sup>189</sup> Voyez en France l'impact d'un arrêt de la Cour de cassation [https://www.lesechos.fr/10/07/2017/lesechos.fr/030439006682\\_evaluation-des-avocats-sur-le-web---la-donne-a-change.htm](https://www.lesechos.fr/10/07/2017/lesechos.fr/030439006682_evaluation-des-avocats-sur-le-web---la-donne-a-change.htm). D'un point de vue technique, il n'y aurait d'ailleurs aucune difficulté à prédire, non pas les chances de succès en justice *in abstracto*, mais les chances de justice avec tel ou tel avocat. De même, il ne serait pas techniquement problématique de faire des analyses statistiques sur les conclusions des avocats. Ces différentes pratiques étant néanmoins (si pas techniquement, du moins conceptuellement) fortement hypothétique, nous ne les analyserons pas plus en détails.

<sup>190</sup> Cette mesure pourrait être conçue, par exemple, à l'aide du pourcentage de réussite de ces derniers ou de coût moyen de leur travail pour le client. Voyez, pour une analyse des scénarios plausibles par une analogie avec le monde de la finance, D. KATZ, *op. cit.*, spec. pp. 929-936, 948-949.

<sup>191</sup> Cela serait d'autant plus vrai dans le cas d'une combinaison des techniques de la justice analytique avec les modèles informatiques du droit développés par la recherche en *AI & Law*. Comme mentionné plus haut, ces combinaisons restant à la fois plus théoriques et hypothétiques, elles ne seront pas développées plus en détails.

premier lieu lors de la réflexion doctrinale plutôt que directement dans la pratique. Seulement, il reste nécessaire que ces avancées fassent le pas de la théorie à la pratique et c'est cette étape qui risque d'être potentiellement ralentie par une utilisation intensive des techniques de justice analytique. Le deuxième effet concerne un autre type de performativité des algorithmes statistico-prédictifs qui surgit dans le cadre de négociations entre avocats. Il semble, en effet, probable que les résultats de ces outils soient utilisés comme argument lors de négociation de renoncement à une action contre compensation financière. Ceci ferait, une fois de plus, intervenir l'influence des créateurs de ces outils dans cette facette de la pratique du droit.

### 3.2.2.3. Nouveaux objets et nouveaux outils pour le monde académique

L'impact de la justice analytique sur le monde de la recherche sera de deux types. Premièrement, les nouvelles pratiques se devront d'être étudiées, donnant ainsi de nouveaux objets d'étude à la doctrine et aux chercheurs en droit. Deuxièmement, à condition de les comprendre et d'apprendre à les utiliser, les outils de la justice analytique constituent de véritables trésors méthodologiques. Ces outils permettront potentiellement de détecter des tendances encore jamais analysées (La justice est-elle plus clémente à certains moments de l'année ? A quel point les juges sont-ils sensibles aux pressions médiatiques ?), de mettre à jour des pratiques difficilement étudiées auparavant (Quid du sexisme en droit ? Du racisme ? Du mépris de classe ?) ou d'étudier des objets encore jamais considérés (L'analyse de masse de conclusions des avocats devient possible).

En ce qui concerne l'enseignement, les réformes hypothétiques du curriculum dépendraient trop fort de la pratique des autres acteurs pour tenter une analyse des conséquences de la justice analytique sur celui-ci<sup>192</sup>.

### 3.2.3. *Le point de vue des utilisateurs du droit, particuliers et entreprises*

Etant donné que la justice analytique *en pratique* se présente actuellement comme un outil à destination du professionnel, il nous faut préciser ce qui est entendu par l'utilisation de la justice analytique par les entreprises et les particuliers. Pour nous, toute modification du comportement de ces derniers causée par la présence de la justice analytique est considérée comme une utilisation (certes indirecte) de ces technologies. Définie ainsi, des trois faces de la justice analytique, seule la face prédictive semble pouvoir engendrer ce type d'utilisation.

---

<sup>192</sup> Voyez pour quelques réflexions sur la difficulté de déterminer quoi enseigner des technologies aux étudiants en droit aux Etats-Unis : K., ASHLEY, *op. cit.*, spec. pp. 48-49.

### 3.2.3.1. La performativité et l'impact de l'implémentation

Par l'entremise de leurs avocats ou juristes, il n'est pas improbable que les entreprises et les particuliers utilisent (à tout le moins parmi d'autres facteurs) la prédiction algorithmique pour les aider dans leurs prises de décisions<sup>193</sup>. Cette dernière peut bien sûr être influencée dans deux sens opposés : introduire une action en justice, ou s'abstenir.

Cette dernière version, négative, semble révéler tout particulièrement la puissance des technologies<sup>194</sup>. La probabilité résultant d'un algorithme prédictif pourrait, en effet, jouer un rôle, potentiellement significatif, dans l'abandon d'un désir d'action en justice. Vu d'un autre angle, dans cette version, l'utilisateur du droit « accepte » le « verdict » de la machine (qui n'est alors plus du tout « prédictif » mais bien concrétisé), puisque celui-ci « admet » d'une certaine manière sa défaite en n'introduisant pas son action. Ce phénomène arrive, bien entendu, déjà à l'heure actuelle par l'entremise des conseils d'un avocat<sup>195</sup>. Toutefois, la prédiction algorithmique et la prédiction d'un juriste ont des propriétés méthodologiques fondamentalement différentes. En effet, l'évaluation d'une situation juridique par un avocat, une fois couchée sur papier, ressemblera, méthodologiquement, fortement à celle d'un juge. Ceci est on ne peut plus différent dans le cas d'une prédiction algorithmique. Dans ce cas, l'implémentation de l'algorithme ayant mené à la décision devient le « dispositif » du jugement algorithmique. Or, cette technique ne se repose pas sur des *raisons* mais bien sur des *corrélations*, ce qui implique une transformation du mode discursif des « jugements ». Puisque l'algorithme se fonde également sur les conceptions juridiques du créateur de l'outil, décider de ne pas introduire l'action sur base de la prédiction de l'algorithme revient donc implicitement à accepter cette conception du droit<sup>196</sup>. Cela revient aussi à accepter de voir son cas *individuel* transformé en une série de caractéristiques catégorisées et fixes, les particularités ayant dû céder devant le besoin de généralisation. Enfin, l'algorithme est basé sur un ensemble de décisions

---

<sup>193</sup> L'influence des technologies sur les décisions sera probablement d'autant plus grande que leurs performances et leur fonctionnement conserveront un côté « mystique » pour le particulier ou l'entreprise en question. Il n'est, de plus, pas certain que tous les particuliers soient « bien » informés par les probabilités qui sont générées par les prédictions algorithmiques. En effet, les humains (et cela inclut les avocats, du moins actuellement) sont notoirement mauvais dans leur compréhension de la signification concrète d'une probabilité donnée. Pour plus d'information sur les mauvaises aptitudes de jugement statistique du cerveau humain, voyez D., KAHNEMAN, *Thinking, Fast and Slow*, New-York, Farrar, Straus and Giroux, 2011, spec. pp. 109-185.

<sup>194</sup> La version positive nous force, en effet, à passer le test de l'action en justice devant un juge. Dès lors, l'influence des technologies dans la chaîne des événements est atténuée.

<sup>195</sup> Celui-ci est d'ailleurs souvent considéré comme le premier juge du dossier de son client. Voyez, par exemple, la présentation de ce qu'est un avocat par le célèbre site d'avocat.be à l'adresse [https://avocats.be/sites/default/files/texte\\_avocat3.pdf](https://avocats.be/sites/default/files/texte_avocat3.pdf) ou encore des présentations par les avocats eux-mêmes comme sur le site internet de ce cabinet <https://www.renson-avocats.be/role-et-les-missions-avocat.html>.

<sup>196</sup> Aussi bien du droit *positif* que de l'ontologie juridique.

passées dont l'extension chronologique n'est pas forcément connue à l'avance. Or, celles-ci ont potentiellement été influencées par des considérations sociales devenues obsolètes entre temps mais qui ont tout de même influencé le modèle statistique de l'algorithme prédictif<sup>197</sup>.

D'aucuns affirmeront que ces considérations ne sont qu'une sorte de nostalgie préemptive et mal placée. Seule compte l'efficacité prédictive de l'algorithme, les *méthodes* utilisées pour aboutir à la prédiction ne sont pas pertinentes. Encore une fois, ce genre d'arguments reflète une certaine conception du droit sur laquelle chacun peut se faire sa propre conviction. Les constats méthodologiques énoncés ci-dessous restent la contrepartie de l'utilisation d'algorithmes prédictifs basés sur le *machine learning* dans le cadre de la justice.

Une autre série d'effets se produisent si l'utilisation de la face prédictive de la justice analytique se répand de manière très intensive, au point d'en affecter, même partiellement, la production judiciaire. Dans ce cas, une rétroaction se met en place dans l'algorithme prédictif lui-même, puisqu'il repose en partie sur les décisions de justice les plus récentes qu'il aura influencées.

Pour illustrer le genre de rétroaction possible, prenons un cas extrême<sup>198</sup> où seuls les litiges allant encore devant les tribunaux sont ceux dont la probabilité de succès produit par l'algorithme prédictif est entre 40% et 60%. Tous les autres étant réglés par des procédures transactionnelles<sup>199</sup>. Dans ce cas, si la base de données d'entraînement de l'algorithme de *machine learning* est mis à jour avec les décisions les plus récentes, le nombre et la proportion des « causes difficiles » parmi l'ensemble des décisions seront nettement plus élevés, affectant les prédictions de l'algorithme au point de remettre potentiellement en question la validité du modèle pour des causes « simples ».

### 3.2.3.2. Une vision actuarielle du droit

Au-delà de ce risque de performativité, une autre tendance, certes plus hypothétique, peut être envisagée. Celle de l'impact de la justice prédictive sur la mentalité avec laquelle un particulier se rend chez son avocat. Aux termes mêmes du fondateur de Case Law Analytics, la composante

---

<sup>197</sup> En effet, la société change. Comme pour la question des changements législatifs et jurisprudentiels, cette considération dépend du volume de la production judiciaire annuelle et de la quantité de décisions nécessaires pour obtenir de bonnes performances prédictives. Si le premier est plus important que la seconde, l'influence de l'évolution de la société peut facilement être évitée en n'utilisant que des décisions récentes. Toutefois, dans le cas inverse, des décisions plus anciennes devront être utilisées, amenant potentiellement à une influence de l'évolution sociale.

<sup>198</sup> Notons que si cet exemple est particulièrement extrême, dans l'hypothèse où les algorithmes de justice analytiques se répandent, il n'est pas incohérent d'envisager effectivement qu'une plus grande proportion des litiges favorisant une partie soit réglée hors des tribunaux.

<sup>199</sup> Une telle extrémité engendrerait une raréfaction des décisions de justice qui n'est pas prise en compte ici car a priori trop irréaliste.

prédictive de la justice analytique mériterait d'être dénommée « justice actuarielle ». Utilisée cette terminologie reflète déjà une approche étroitement économique, même financière, du processus judiciaire. Celui-ci est envisagé comme un phénomène naturel impactant les gains et pertes d'une entreprise et dont il conviendrait de limiter les risques.

Il n'est pas notre intention ici de critiquer les termes employés par M. Levy-Vehel, tout révélateurs qu'ils soient. Celui-ci ne fait qu'écho à une vision économique du droit<sup>200</sup> qui s'est répandue bien plus largement que ne le montre cet exemple particulier et qui répond à la préoccupation principale (voir exclusive) des entreprises : la maximisation du profit. Toutefois, si cette comptabilité actuarielle de la justice se répand dans les cabinets d'avocats, il ne serait pas étonnant qu'elle renforce une modification des mentalités, non pas des entreprises<sup>201</sup>, mais des particuliers.

### 3.3. La justice analytique et le futur du droit

Un des présupposés de la section précédente était que les conséquences de la justice analytique étaient limitées aux conséquences de ses utilisations par les acteurs classiques du droit. Cette hypothèse implicite se justifiait par le fait que l'utilisation de ces technologies par d'autres acteurs reste incertaine. Toutefois, en s'autorisant à envisager des scénarios plausibles, bien que plus hypothétiques, il est possible de faire le lien entre la justice analytique et d'autres grands courants de transformation du droit qui ne sont pas toujours le fruit de ses acteurs traditionnels. Plus précisément, comment ne pas voir le parallèle entre la justice analytique, et plus particulièrement sa face statistico-prédictive, avec le phénomène de migration des normes juridiques dans des dispositifs techniques<sup>202</sup> ? Les exemples<sup>203</sup> les plus classiques d'algorithmes ayant un effet normatif, tel que Bitcoin et les autres crypto-monnaies, les contrats intelligents (« *smart contracts* ») d'Ethereum, la voiture autonome de Google ou encore la surveillance automatisée des violations du droit d'auteur par le système ContentID élaboré par Youtube sont désormais devenus si connus qu'ils pourraient presque être considérés comme

---

<sup>200</sup> Pour plus d'information sur cette évolution, voyez B., FRYDMAN, *Les transformations du droit moderne*, Bruxelles, Story-Scientia, 1999, spec. pp. 47-59.

<sup>201</sup> Qui sont déjà, du moins dans le cas des compagnies d'assurance, intéressées par cette quantification actuarielle du risque juridique. Voyez : H., GUILLAUD, *La justice prédictive (1/3)*, op. cit., spec. §16.

<sup>202</sup> Voyez sur ce point la remarquable conférence de G., LEWKOWICZ, *Le droit dans les choses, le futur du droit ?*, 2016, dont les diapositives sont disponibles sur le site internet du centre Perelman de philosophie du droit de l'ULB à l'adresse : [http://www.philodroit.be/IMG/pdf/gl\\_-\\_droit\\_dans\\_les\\_choses\\_-\\_futur\\_du\\_droit-2.pdf](http://www.philodroit.be/IMG/pdf/gl_-_droit_dans_les_choses_-_futur_du_droit-2.pdf).

<sup>203</sup> Les exemples qui suivent sont tirées de la conférence citée ci-dessus.

faisant partie de la culture générale.

Certes, comme déjà expliqué, rien dans la justice analytique n'a de visée normative directe. Ni ses principaux acteurs (les fondateurs et directeurs des *legaltechs*), ni ses utilisateurs, ni ses penseurs<sup>204</sup> ne songent sérieusement à utiliser directement ces outils à des fins normatives ou pour « automatiser la justice ». Toutefois, on ne peut s'empêcher de voir un parallèle frappant entre, par exemple, la conception du système entourant ContentID<sup>205</sup>, à savoir reproduire la détection de violation de *copyright* et l'évaluation de ce qui constitue le « *fair use* » et la prédiction algorithmique qui tâche de reproduire la parole du juge<sup>206</sup>. Celui-ci n'est-il, conceptuellement, un cas particulier de celle-là ? La ressemblance est encore plus frappante quand on sait que Youtube utilise également du *machine learning* dans l'implémentation des algorithmes de ContentID<sup>207</sup>.

De plus, les acteurs traditionnels du droit ne semblent pas avoir été à l'initiative de la création ou de l'utilisation des nouveaux objets juridiques mentionnées en exemples, tout comme ils réagissent à l'arrivée de la justice analytique, sans en être les instigateurs. Ceux-ci n'auront donc peut-être pas l'initiative (et encore moins le contrôle) de la manière dont celle-ci sera combinée avec le courant de migration des normes juridiques dans les objets techniques, particulièrement une fois les bases de données constituées. En effet, il est tout de même question, malgré tous leurs défauts et les insuffisances actuelles, de technologies permettant d'automatiser (même partiellement) l'acquisition de meilleures connaissances de textes juridiques et l'émission de jugements. Il serait donc intuitivement logique d'en faire usage dans l'objectif de créer des dispositifs techniques à visée normative, mais plus versatiles dans les types de « décisions » pouvant être prises<sup>208</sup>. Les *smarts contracts* dont l'exécution est automatique ne pourraient-ils pas tirer profit d'un mécanisme de contestations tranchées sur base de classifications algorithmiques<sup>209</sup> ?

---

<sup>204</sup> Sauf à de rares exceptions, comme Jean-Pierre Buyle qui proposait d'utiliser la prédiction algorithmique pour automatiser les contentieux de moindre valeur économique. Voyez G., QUOISTIAUX, *op. cit.*, spec. p. 3.

<sup>205</sup> Qui, *per se*, ne concerne que l'identification parmi les vidéos Youtube de contenu répertorié dans une base de données.

<sup>206</sup> C. SPRIGMAN, « Will Algorithms Take the Fairness out of Fair Use », *Journals of Things We Like*, Vol. 1, 2018, pp. 1-2, spec. p. 1.

<sup>207</sup> *Ibid*, p. 3

<sup>208</sup> Ce type de technologie est nécessaire à la société envisagé, analysée et critiquée, par exemple, dans M. HILDEBRANDT et B., KOOPS, « The Challenges of Ambient Law and Legal Protection in the Profiling Era », *The Modern Law Review*, Vol. 73, N° 3, 2010, pp. 428-460.

<sup>209</sup> On pourrait même envisager, une fois ce type de technologie suffisamment établie, l'établissement d'une clause d'arbitral faisant dépendre l'issue du litige d'une prédiction algorithmique. Aboutissement finale de la logique économique du droit, la décision serait pratiquement instantanée et peu coûteuse.

S'avancer plus loin, affirmer plus que l'existence d'une convergence possible de l'utilisation des technologies de la justice analytique avec d'autres dispositifs techniques ayant (eux) une visée normative, relève toutefois de la science-fiction.

#### 4. Conclusion

En bilan de ce travail, il semble que nous ayons apporté de nombreux éléments de réponse à notre question de base : quel impact aura l'intelligence artificielle en (et sur le) droit ?

Après un approfondissement du sens à donner à l'objet de notre interrogation, aussi bien dans les champs de la recherche académique que sur le marché des *legaltechs*, nous avons défini le phénomène actuel de l'intelligence artificielle en droit. Renommé « justice analytique », il s'agit d'un phénomène ayant pour base le *big data* et les techniques du *machine learning* et prenant corps selon trois pratiques : l'amélioration de la recherche, l'automatisation du traitement de documents juridiques et la production de statistiques ou de prédictions sur les litiges judiciaires.

Nous avons ensuite procédé à une analyse technique du *machine learning* qui, bien que sommaire, nous a permis de déduire les conséquences que l'implémentation de la justice analytique aurait sur le monde juridique. Nous avons ainsi conclu, d'une part, à une exacerbation du débat relatif à la publication des décisions de justice et, d'autre part, à l'impossibilité de concevoir un algorithme de *machine learning* qui ne soit pas le reflet implicite de ce que ses créateurs ont considéré comme pertinent. Notre première hypothèse concernant l'absence de neutralité des nouvelles technologies a donc été vérifiée.

Nous avons ensuite analysé les conséquences qu'aurait l'utilisation de la justice analytique par les différents acteurs classiques du droit. Ce faisant, nous avons dû postuler que les techniques et pratiques à la base de la justice analytique se répandraient à toutes les sphères considérées, malgré le caractère contestable (car incertain) de cette hypothèse au vu de nos conclusions antérieures sur l'évolution du marché des *legaltech*. Nous avons également choisi de ne pas considérer les hypothétiques interactions futures avec d'autres types de technologies ou modèles informatiques développés par la recherche en *AI & Law*. Les conséquences déduites en termes politiques consistaient en une plus grande ingérence potentielle du pouvoir politique dans le pouvoir judiciaire et dans une probable réification discursive de la jurisprudence. Celles déduites par rapport aux milieux du droit (et aux particuliers) concernaient, d'abord, des risques de performativité des prédictions ou statistiques produites (au niveau des juges, des avocats et des particuliers) qui interagissaient, entre autres, avec les réponses apportées aux nouvelles

questions déontologiques se posant aux avocats. Nous avons également conclu à un ralentissement potentiel de l'évolution de la théorie du droit et un bouleversement probable de certaines catégories socio-professionnelles comme les assistants juridiques et les avocats stagiaires. Enfin, nous avons abordé brièvement les liens entre la justice analytique et le phénomène de migration des normes juridiques dans les dispositifs techniques (« le droit dans les choses »), concluant à une combinaison probable de ces phénomènes. Sur cette base, notre seconde hypothèse (qui postulait que l'impact des nouvelles technologies se restreindrait à une reconfiguration de pratiques existantes autour de celles-ci) n'est que partiellement vérifiée. En effet, si la majorité des effets susmentionnés consistent dans la reconfiguration attendue, certains effets probables (comme le ralentissement potentiel de l'évolution de la théorie du droit ou la combinaison de la justice analytique avec le « droit dans les choses ») ne s'y réduisent pas.

De nombreuses pistes d'approfondissement ont été ouvertes tout au long de ce travail. Quel sera l'impact des *legaltechs* n'utilisant pas l'intelligence artificielle ? Quid des systèmes-experts ? De quelles manières les conséquences déduites ici sont-elles modifiées ou approfondies lorsque l'interaction entre la justice analytique et les modèles informatiques du droit est prise en compte ? Toutes ces questions constituent une base propice pour de nouvelles recherches qui auront leur importance pour la compréhension de l'évolution future du droit.

Enfin, notre travail a également ses limites, dont au moins deux peuvent constituer le point de départ de nouvelles recherches. Premièrement, notre analyse du *machine learning* fut relativement sommaire et une analyse plus approfondie, ainsi qu'une description plus quantitative, permettrait peut-être d'approfondir nos conclusions et de découvrir d'autres conséquences de l'implémentation de la justice analytique. Deuxièmement, notre étude n'aborde que très peu l'aspect psychologique des technologies, se bornant à critiquer la mystification médiatique. Dès lors, notre travail pourrait utilement être approfondi par l'étude des effets psychologiques de l'intelligence artificielle (et des autres technologies) sur le monde juridique.



## 5. Bibliographie

La grande majorité des sources ci-dessous ne suivant pas le triptyque classique des sources juridiques (législation, jurisprudence, doctrine), nous avons choisi de les classer selon la première section dans laquelle elles furent citées. Pour chaque partie, l'ordre alphabétique par auteur est ensuite utilisé. Toutefois, dans un but de clarté, les articles de presse cités à titre illustratif et les sites internet ne faisant pas référence à des documents précis ne sont pas repris ci-dessous.

### 5.1. Introduction

- BAKER, S., *Final Jeopardy: The Story of Watson, the Computer that Will Transform Our World*, Boston, Mariner Books, 2012.
- FRYDMAN, B., *Comment penser le droit global?*, Working Papers du Centre Perelman de Philosophie du Droit, Vol.1, 2012. Publié en ligne à l'adresse : [http://www.philodroit.be/IMG/pdf/comment\\_penser\\_le\\_droit\\_global\\_2011.pdf](http://www.philodroit.be/IMG/pdf/comment_penser_le_droit_global_2011.pdf)
- KNIGHT, W., « This Robot Could Transform Manufacturing », *MIT Technology Review*, 2012. Publication, en ligne, à l'adresse <https://www.technologyreview.com/s/429248/this-robot-could-transform-manufacturing/#comments>. Dernière consultation le 11 avril 2017.
- MISHRA, P., *et al.*, « Of Art and Algorithms: Rethinking Technology & Creativity in the 21st Century », *TechTrends*, 2013, Vol. 57, N°3, pp. 10-14.
- NARAYANAN, A., *Law, Computer Science, and Artificial Intelligence*, Chicago, Intellect, 1998.
- SARTOR, G., et BRANTING, L., « Judicial Applications of Artificial Intelligence », Netherlands, Springer, 1998.
- SILVERMAN, A., *Mind, machine, & metaphor: an essay on artificial intelligence and legal reasoning*, Colorado, Westview Press, 1993.
- WALDROP, M., « Autonomous vehicles: No drivers required », *Nature*, Vol. 518, N°7537, 2015, pp. 20-23.

### 5.2. Inventer l'intelligence artificielle

- ASHLEY, K., *Artificial intelligence and legal analytics: new tools for law practice in the digital age*, Cambridge, Cambridge University Press, 2017.

- BENCH-CAPON, T., *et al.*, « A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law », *Artificial Intelligence and Law Journal*, Vol. 20, N°3, 2012, pp. 1-101.
- BEXHENRY, F., *et al.*, « Introduction to the special issue on Artificial Intelligence for Justice (AI4J) », *Artificial Intelligence and Law*, 2017, Vol. 25, N°1, pp. 1–3.
- BOELLA, G., *et al.*, « A Critical Analysis of Legal Requirements Engineering from the Perspective of Legal Practice », *7th IEEE International Workshop on Requirements Engineering and Law (RELAW)*, 2014, pp. 14-21.
- DEAN, J., *Big Data, Data Mining and Machine Learning: Value Creation For Business Leaders and Practitioners*, New Jersey, John Wiley & Sons, Inc., 2014.
- EZRATTY, O., *Les avancées de l'intelligence artificielle - 3*, 2016. Articles publiés en ligne à l'adresse <http://www.oezratty.net/wordpress/2016/avancees-intelligence-artificielle-3/>. Dernière consultation le 3 avril 2018.
- FERNÁNDEZ-BARRERA, et M., SARTOR, G., « The Legal Theory Perspective: Doctrinal Conceptual Systems vs. Computational Ontologies », *Approaches to legal ontologies: theories, domains, methodologies*, P., Casanovas, G., Sartor, dir., Londres, Springer, 2010, pp. 15-23.
- HASHMI, M., et GOVERNATORI, G., « Normative Requirements for Business Process Compliance », *Australasian Symposium on Service Research and Innovation*, 2013.
- HURWITZ, J., et KIRSCH, D., *Machine learning for dummies: IBM limited edition*, New Jersey, John Wiley & Sons, Inc., 2018.
- JORDAN, M., et MITCHELL, T., « Machine learning: Trends, perspectives, and prospects », *Science*, Vol. 349, N°6245, 2015, pp. 255-260.
- KATZ, D., « Quantitative legal prediction—or—how I learned to stop worrying and start preparing for the data-driven future of the legal services industry », *Emory law journal*, Vol. 62, pp. 909-966.
- LANGLEY, P., « The changing science of machine learning », *Machine Learning*, Vol. 82, N°3, 2011, pp. 275–279.
- MEHL, L., « Automation in the Legal World », *Mechanisation of Thought Processes*, Vol. 2, 1959, pp. 755-787.
- MICHALSKI, R., *et al.*, *Machine Learning: An Artificial Intelligence Approach*, Berlin,

Springer Berlin Heidelberg, 2013.

- MOZINA, M., *et al.*, « Argument Based Machine Learning Applied to Law », *Artificial Intelligence and Law*, Vol. 13, N°1, pp. 53-73.
- PALIWALA, A., « Rediscovering artificial intelligence and law: an inadequate jurisprudence? », *International Review of Law, Computers & Technology*, Vol. 30, N°3, 2016, pp.107-114.
- RISSLAND, E., *et al.*, « AI and Law: A fruitful synergy », *Artificial intelligence*, Vol. 150, N°1, 2003, pp. 1-15.
- SAMUEL, A., « Some Studies in Machine Learning Using the Game of Checkers », *IBM Journal of Research and Development*, Vol.3, N°3, 1959, pp. 210-229.
- SARTOR, G., et ROTOLO, A., « AI and Law », *Agreement technologies*, P. Casanovas, G., Sartor, dir., Londres, Springer, 2013, pp. 199-207.
- SUSSKIND, R., *The End of Lawyers?: Rethinking the Nature of Legal Services*, Oxford, Oxford University Press, 2010.

### 5.3.Utiliser l'intelligence artificielle

- CHARPENTIER CONSULTING, *French Legaltech List*, 2017, publiée, en ligne, à l'adresse:[https://media.wix.com/ugd/c21db1\\_14b04c49ba7f46bf9a5d88581cbda172.pdf](https://media.wix.com/ugd/c21db1_14b04c49ba7f46bf9a5d88581cbda172.pdf).
- DAY ONE, *Droit & digital : réalité et prospectives*, 2017, publiée en ligne à l'adresse : [http://www.actuel-direction-juridique.fr/sites/default/files/droit\\_digital\\_realites\\_prospectives\\_vf.pdf](http://www.actuel-direction-juridique.fr/sites/default/files/droit_digital_realites_prospectives_vf.pdf).
- LE MONDE, *La justice prédictive*, rapport publié en trois parties, en ligne. Première partie à l'adresse <http://internetactu.blog.lemonde.fr/2017/09/09/la-justice-predictive-13-lenjeu-de-louverture-des-donnees/>.
- PALMER, M., *et al.*, « Semantic Role Labelling », *Synthesis Lectures on Human Language Technologies*, Vol. 3, N°1, 2010, pp. 1-103.
- TRACXN, *Legal Tech Report*, 2016, publiée, en ligne, à l'adresse : <https://s3.amazonaws.com/tracxnblog/wp-content/uploads/2016/12/Tracxn-Research-Legal-Tech-Landscape-December-2016-1.pdf>,

## 5.4. Analyse de l'implémentation

- AIKENHEAD, M., « The Uses and Abuses of Neural Networks in Law », *Santa Clara High Technology Law Journal*, Vol. 12, N°1, 1996, pp. 31-70.
- BORGES, F., *et al.*, « Artificial neural networks and legal categorization », *Legal Knowledge and Information Systems. Jurix: The Sixteenth Annual Conference*. Amsterdam: IOS Press, 2003, pp. 11-20.
- COMMISSION DE MODERNISATION DE L'ORDRE JUDICIAIRE - *Rapport consacré à la question de la publication des décisions judiciaires : La plume, le Pélikan et le nuage*, 2014. Disponible en ligne à l'adresse <https://justice.belgium.be/sites/default/files/downloads/Rapport%0publication%20des%20d%C3%A9cisions%20judiciaires.pdf>. Dernière consultation le 12 avril 2018.
- GARAPON, A., « Les enjeux de la justice prédictives », *La semaine juridique*, N°1-2, 9 janvier 2017.
- HASTIE, T., *et al.*, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2<sup>ème</sup> édition, New-York, Springer-Verlag New York, 2009.
- KATZ, D., *et al.*, *Predicting the Behavior of the United States Supreme Court: A general approach*, 2014. Publication sur le site du SSRN à l'adresse: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2463244](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2463244).
- QUOISTIAUX, G., « L'intelligence artificielle peut-elle sauver la justice ? », *Trends tendance*, octobre 2016. Disponible, en ligne, à l'adresse <http://trends.levif.be/economie/lawyerz/l-intelligence-artificielle-peut-elle-sauver-la-justice/article-normal-603501.html>. Dernière consultation le 12 avril 2018.
- REIGELUTH, T., *Comporter la norme : La normativité de l'apprentissage algorithmique à partir du problème du comportement*, thèse de doctorat, Université Libre de Bruxelles, 2018.
- SHALEV-SHWARTZ, S., et BEN-DAVID, S., *Understanding machine learning: From Theory to Algorithms*, Cambridge University Press, New York, 2014.
- SURDEN, H., « Machine Learning and Law », *Washington Law Review*, Vol. 89, N°1, 2014, pp. 87-115.
- SUTHAHARAN, S., *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, New-York, Springer,

2016.

- WARNER, D., « A Neutral Network-Based Law Machine: The Problem of Legitimacy », *Law, Computer & Artificial Intelligence*, Vol. 2, N°2, 1993, pp. 135-147.

## 5.5. Analyse de l'utilisation

- ASHLEY, K., et WALKER, V., « From information retrieval (IR) to argument retrieval (AR) for legal cases: report on a baseline study ». *Proceedings of the 26th Annual Conference on Legal Knowledge and Information Systems 2013*, IOS Press: Amsterdam, pp. 29–38.
- ALLARD, J., « L'indépendance de la justice et le management judiciaire : contre le "pouvoir" ? La ruse de la raison économique » *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014 pp. 33-50.
- CAPURSO, T., « How Judges Judge: Theories on Judicial Decision Making », *University of Baltimore Law Forum*, Vol. 29, N°1, 1998, pp. 5-16.
- DESROSIÈRES, A., *Pour une sociologie historique de la quantification : L'Argument statistique I*, Nouvelle édition, Paris, Presses des Mines, 2008.
- EYRAUD, C., « Réflexions pour une sociologie de la quantification statistique et comptable », *Entreprises et histoire*, Vol. 79, N°2, 2015, pp. 74-87.
- FRYDMAN, B., « Concilier le management avec les valeurs du judiciaire », *Quel management pour quelle justice ?*, T., Marchandise (dir.), Bruxelles, Larcier, 2012, pp. 81-88.
- FRYDMAN, B., *Les transformations du droit moderne*, Bruxelles, Story-Scientia, 1999.
- FRYDMAN, B., *L'évolution des critères et des modes de contrôle de la qualité des décisions de justice*, Working Papers du Centre Perelman de Philosophie du Droit, Vol. 4, 2007. Publié en ligne à l'adresse : [http://www.philodroit.be/IMG/pdf/levolution\\_des\\_criteres\\_et\\_des\\_modes\\_de.pdf](http://www.philodroit.be/IMG/pdf/levolution_des_criteres_et_des_modes_de.pdf)
- FRYDMAN, B., et JEULAND, E., *Le nouveau management de la justice et l'indépendance des juges*, Bruxelles, Dalloz, 2011.
- FRYDMAN, B., et VAN WAEYENBERGE, A., *Gouverner par les standards et les indicateurs: De Hume au rankings*, Primento, 2013.
- HENRY, P., et HOFSTRÖSSLER, P., *L'avenir de la profession d'avocat : Rapport au Ministre de la Justice K. Geens / De toekomst van het advocatenberoep : Rapport aan*

*de Minister van Justitie K. Geens*, 2018.

- KAHNEMAN, D., *Thinking, Fast and Slow*, New-York, Farrar, Straus et Giroux, 2011.
- MOCHALES, R. et MOENS, M-F., « Argumentation mining », *Artificial Intelligence and Law*, Vol. 19, N°1, pp. 11-22.
- MUNUNGU LUNGUNGU, K., « Le management judiciaire ou le glas de l'indépendance du juge ? Un éclairage de droit constitutionnel », *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014, pp. 126-167.
- MARCHANDISE, T., *et al.*, *Quel management pour quelle justice?*, Larcier, 2013.
- NEARY, M.A., et XIN CHEN, S., « Artificial Intelligence: Legal Research and Law Librarians », *AALL spectrum*, Vol. 21, N°5, 2017, pp. 16-20.
- SCHOENAERS, F., « Lorsque le management entre au tribunal : évolution ou révolution ? », *Rev. dr. U.L.B.*, Vol. 41, N°1-2, 2014, pp. 171-209.
- TUMONIS, V., « Judicial decision-making from an empirical perspective », *Baltic Journal of Law & Politics*, Vol. 6, N°1, 2013, pp. 140-162.
- UYTENDAELE, M., *Trente leçons de droit constitutionnel*, 2e éd., Bruxelles, Bruylant, 2014.

## 5.6.Futur du droit

- HILDEBRANDT, M., et KOOPS, B., « The Challenges of Ambient Law and Legal Protection in the Profiling Era », *The Modern Law Review*, Vol. 73, N° 3, 2010, pp. 428-460.
- LEWKOWICZ, G., *Le droit dans les choses, le futur du droit ?*, 2016. Publié, en ligne, à l'adresse: [http://www.philodroit.be/IMG/pdf/gl - droit dans les choses - futur du droit-2.pdf](http://www.philodroit.be/IMG/pdf/gl_-_droit_dans_les_choses_-_futur_du_droit-2.pdf)
- SPRIGMAN, C., « Will Algorithms Take the Fairness out of Fair Use », *Journals of Things We Like*, Vol. 1, 2018, pp. 1-2.