

**Faculté de droit et de criminologie**

# **De l'intelligence artificielle dans la pratique du droit**

Réception en droit européen, incidence sur la profession d'avocat et éthique

Auteur : Raphaël CANVAT  
Promoteur : Christophe LAZARO  
Année académique 2019-2020  
Master en droit HD

## **Plagiat et erreur méthodologique grave**

Le plagiat, fût-il de texte non soumis à droit d'auteur, entraîne l'application de la section 7 des articles 87 à 90 du règlement général des études et des examens.

Le plagiat consiste à utiliser des idées, un texte ou une œuvre, même partiellement, sans en mentionner précisément le nom de l'auteur et la source au moment et à l'endroit exact de chaque utilisation\*.

En outre, la reproduction littérale de passages d'une œuvre sans les placer entre guillemets, quand bien même l'auteur et la source de cette œuvre seraient mentionnés, constitue une erreur méthodologique grave pouvant entraîner l'échec.

\* À ce sujet, voy. notamment <http://www.uclouvain.be/plagiat>.

*À mon Père,  
Qui repose dans le calme des dieux  
Parmi les cerisiers  
Roses, infusés d'astres lactescents*

*À Sophie,  
Qui sait déjà ce qu'elle comprendra plus tard*

« Toute question relative à notre avenir immédiat est désormais une question relative aux causes premières et aux causes dernières – une question métaphysique »<sup>1</sup>.

## AVANT-PROPOS

La science et les technologies ont connu, ces dernières décennies, une évolution quasi-exponentielle. La révolution digitale affecte d'ores et déjà toutes les strates socio-économiques des sociétés modernes et industrialisées bouleversant par là même les us et coutumes dans de nombreuses professions<sup>2</sup>. Qu'en est-il du monde du droit, de la justice et plus spécifiquement des cabinets d'avocats en Europe ? Ont-ils pris acte de cet « électrochoc numérique »<sup>3</sup> ; prennent-ils le train en marche ou risquent-ils de rater le coche ?

Pour J.-P. BUYLE, président d'Avocats.be de 2016 à 2019, il y lieu de tirer le constat suivant : « [l]e monde juridique est en papier et la révolution numérique est en cours »<sup>4</sup>. Pourtant cette révolution ne s'arrête pas à la digitalisation, à la dématérialisation ou au fait de *going paperless* et *cloud-based*<sup>5</sup>. En fait, on observe deux lames de fond : d'un côté les entreprises doivent se réinventer et se moderniser en se digitalisant sans quoi elles risquent de disparaître et, de l'autre, les avocats sont de plus en plus perçus par les justiciables comme des prestataires de services ou des entrepreneurs dont les prestations, par exemple, naguère taxées par des honoraires, sont aujourd'hui soumises à la TVA<sup>6</sup>. Bien plus, à l'ère du digital et avec le développement du *World Wide Web* et des politiques dites d'*open data* et d'*open access*, l'information juridique est de plus en plus directement et aisément accessible. En conséquence, les avocats et les juristes ont en partie perdu le monopole du savoir qui leur appartenait jusqu'alors sur cette science obscure *prima facie* qu'est le droit.

<sup>1</sup> J.-L. CURTIS, « Préface », K. AMIS, *L'univers de la science-fiction*, Paris, Payot, n°32, 1960, p. 8.

<sup>2</sup> Selon le cabinet de conseil MCKINSEY & COMPANY la plupart des entreprises se trouvent, avec la révolution digitale, face à une véritable disruption technologique qui implique la nécessité de devoir se réinventer. Voy. i.a., <https://www.mckinsey.com/>, consulté en avril 2019.

<sup>3</sup> B. MARGARITELLI, « L'électrochoc numérique vs. les avocats », *Journ. spé. soc.*, 2018, n°38, pp. 1 et s.

<sup>4</sup> Voy. <http://numerique.avocats.be/>, consulté en avril 2019. Cf. J.-P. BUYLE, A. VAN DEN BRANDEN, « La robotisation de la justice », H. JACQUEMIN, A. DE STREEL, *L'intelligence artificielle et le droit*, Bruxelles, Larcier, CRIDS, 2017, pp. 259-317.

<sup>5</sup> C'est-à-dire de transitionner vers un environnement sans support papier et en réseau via Internet.

<sup>6</sup> En Belgique, les avocats ne sont plus exemptés de TVA depuis le 1/1/2014.

Voy. <https://finances.belgium.be/>, consulté en avril 2019.

L'ère de la digitalisation demande donc un changement systémique : les cabinets d'avocats doivent se moderniser ; ils doivent à leurs clients plus de rapidité et d'efficacité dans ce qui devient un véritable marché dans lequel, pour être compétitif, il convient d'être connectés et travailler de concert avec le numérique et les *legaltechs*<sup>7</sup>. À la fois cheval de Troie d'une « *common lawisation* » du droit européen continental et boîte de Pandore contenant l'espoir d'une justice meilleure, plus efficace et moins coûteuse, comment s'assurer que la digitalisation et l'intégration d'outils liés à l'intelligence artificielle (ci-après IA) dans les cabinets et les prétoires ne se fera pas sur l'autel de principes généraux de droit ou d'impératifs éthiques ou déontologiques ?

## INTRODUCTION

### I. PROBLÉMATIQUE

#### § 1 – Contexte

À peine le droit, au début du XXI<sup>ème</sup> siècle, commence-t-il à légiférer sur les nouvelles technologies, comme par exemple les questions de cyber-droit, d'e-contrats, de signature électronique et de protection des données à caractère personnel que se profile déjà à l'horizon la quatrième révolution industrielle, dans la droite ligne de la révolution numérique, à savoir celle des robots, de l'IA et des sociétés privées toutes puissantes, plus riches que certains États, *i.e.*, les géants de Web<sup>8</sup>, qui investissent massivement dans la recherche en robotique, les biotechnologies, la génétique et l'IA<sup>9</sup>. L'impératif technoscientifique dispose que ce qui est techniquement possible doit être réalisé<sup>10</sup>. Poussée par la logique capitaliste du profit, inspirée du libéralisme économique et alimentée par la consommation de masse, la science se

<sup>7</sup> Les *legaltechs* désignent originellement les entreprises exploitant les technologies de l'information dans le domaine du droit afin de proposer des services juridiques innovants ; ces entreprises sont le plus souvent des start-ups spécialisées dans le domaine juridique.

<sup>8</sup> Ou les *Big Five Tech Companies* ou GAFAM, soit GOOGLE, AMAZON, FACEBOOK et MICROSOFT, auxquels on ajoute parfois IBM, TWITTER, NVIDIA, TESLA ou encore NETFLIX.

<sup>9</sup> Le succès des technosciences combiné au triomphe du réductionnisme scientifique réveillent d'anciens fantômes : les tenants du transhumanisme rêvent de prolonger indéfiniment la durée de la vie.

Voy. à cet égard les travaux du *Google X Lab* et de *CALICO (California Life Company)* qui combinent recherche en nanotechnologies, biotechnologies, informatique, sciences cognitives et intelligence artificielle le tout dans une perspective transhumaniste assumée.

<sup>10</sup> Cfr la loi de GABOR dite du “déterminisme technologique” : « what can be made, will be made ». G. HOTTOIS privilégie quant à lui l'expression “impératif technoscientifique” : « actualiser, réaliser tout ce qui est possible ».

développe de nos jours à une vitesse effrénée. L'éthique et le droit ont peine à suivre et à s'adapter à un monde qui évolue à toute vitesse. Combien de décennies faudra-t-il attendre avant de voir s'opérer la « mutation technologique » de l'être humain par les thérapies géniques, la cybernétique et le clonage ? À quand les premières nanopuces dans le cerveau pour permettre d'augmenter artificiellement la mémoire et les capacités mentales<sup>11</sup> ? À quand la première singularité : la première IA à la fois *complète* et *consciente* ?<sup>12</sup> Autant d'eau pour alimenter le moulin de la recherche académique sur les questions éthiques et de droit liées aux technologies émergentes... L'industrie cinématographique hollywoodienne et sa pléthore de films et de séries de science-fiction (ci-après SF) nous fait assez miroiter l'avenir glorieux d'une post-humanité. Et pourtant, comment ne pas voir que la prise en main par l'homme de sa propre évolution entraînera des discriminations ? Le législateur ne devrait-il pas préconiser, quand il s'agit de biotechnologies, d'IA, et, plus largement, de la prolongation de la vie, du contrôle du comportement et de la manipulation génétique, l'argument de la pente glissante et une heuristique de la peur, voire un certain catastrophisme<sup>13</sup> ?

## § 2 – Formulation de la question de recherche et plan général du mémoire

Dans le cadre restreint de ce mémoire qui s'intéresse, de façon générale, au droit et à l'IA, il ne sera pas question de robotique ou d'*hardware-based AI* – de C3PO, d'HAL 9000, de TARS ou encore de Rick DECKARD ou de Dolores ABERNATHY – mais bien plutôt de *software-based AI* telle que mise en œuvre dans certains ERP<sup>14</sup> de nouvelle génération destinés aux juristes et qui supplantent les anciens systèmes experts légaux<sup>15</sup>. Nous parlerons donc de ce qu'il est convenu de désigner sur le vocable « *legaltech* » et d'IA de type *narrow* ou *weak* implémentées dans des solutions informatiques professionnelles par opposition à

<sup>11</sup> Voy. à ce sujet G. HOTTOIS, *Essais de philosophie bioéthique et biopolitique*, Paris, Librairie philosophique J. Vrin, 1999 ; *Dignité et diversité des hommes*, Paris, Librairie philosophique J. Vrin, 2009 ; J.-N. MISSA, « L'homme recombinaire : les enjeux éthiques et philosophiques de la modification du génome humain », J.-Y. GOFFI (dir.), *Regards sur les technosciences*, Paris, Vrin, 2006, pp. 111 et s.

<sup>12</sup> À ce sujet, voy. J.-G. GANASCIA, *Le mythe de la singularité : Faut-il craindre l'intelligence artificielle ?* Paris, Seuil, 2017.

<sup>13</sup> Sur ces questions voy. H. JONAS, « L'homme en tant qu'objet de la technique », *Le principe responsabilité, une éthique pour la civilisation technologique*, Paris, Flammarion, Champs essais, 2008, pp. 50-57 et J.-P. DUPUY, *Pour un catastrophisme éclairé : Quand l'impossible est certain*, Paris, Seuil, 2004.

<sup>14</sup> *Enterprise Resource Planning* ou, en français, progiciels de gestion intégrée.

<sup>15</sup> Il s'agit de logiciels implémentant *i.a.*, *natural language processing*, *machine learning*, *data mining*, *documents automation*, *automated legal research*, *practice management* ou encore *predictive analytics tools*.

l'IA dite générale ou *strong*. Sont également hors du champ de recherche les questions se rapportant aux *lethal autonomous weapons* (LAWs) et au droit de la guerre ou encore les questions de droit ayant trait aux véhicules autonomes et aux questions de responsabilité et d'assurance. Dans ce mémoire, nous nous proposons plutôt d'étudier les possibilités d'intégration et d'implémentation d'outils liés à l'IA (*IA-powered tools*) dans les cabinets d'avocats et les éventuels problèmes et bénéfices que suscite cette évolution.

Pour ce faire, nous commençons d'abord par évoquer l'histoire intellectuelle de la notion d'IA à travers l'histoire afin de planter le décor. Avec cette brève tentative d'archéologie foucauldienne de l'IA, nous soulignons que le désir de créer des machines automates, androïdes ou non, relève, d'un fantasme archaïque sous-jacent profondément ancré dans l'inconscient collectif et l'histoire des civilisations (**partie I – section I**).

Dans un deuxième temps, nous esquissons une définition de ce que l'on entend par IA quand il est question d'outils d'automation et de systèmes d'apprentissage machine dans le contexte du droit. Ce faisant, nous soulignons le passage du paradigme des systèmes experts légaux des années 1990 à celui, plus récent, du *machine learning* et des *legaltechs*. Par ailleurs, nous nous demandons s'il existe, actuellement, en Europe ou ailleurs, du droit positif encadrant l'IA et s'il y en a, quelle en est la portée normative. Enfin, nous prenons acte de l'avis de différents experts du monde académique (CRIDS@UNAMUR et CODEX@STANFORD) ainsi que des logiciels et solutions développés par différentes sociétés et start-ups actives dans le domaine (**partie I – section II**).

Dans un troisième temps, nous confrontons l'IA aux droits humains fondamentaux (ci-après DHF). En effet, s'il est clair que les outils liés à l'IA modifient la façon de travailler des cabinets d'avocats, il faut encore se demander si l'IA, dans ce contexte, est un moteur de progrès social et technique. Si l'on doit s'attendre à une amélioration significative des conditions de travail grâce aux outils d'automatisation ; ces améliorations peuvent-elles, d'un autre côté, nuire aux justiciables, voire aux juristes eux-mêmes, et mettre en porte à faux certains grands principes protégés par des conventions européennes et internationales ? Du fait du caractère eurocentré, spéculaire et relatif des DHF, nous proposons de privilégier les notions axiologiquement neutres d'autonomie (de la personne et de la justice), de dignité humaine et de prudence afin de pallier au risque de cristallisation de la justice par la logique reproductive des algorithmes ainsi qu'au risque de substitution du calcul probabiliste à la délibération du juge – ou de l'avocat – sur le cas d'espèce (**partie I – section III**).

Dans la seconde partie de ce mémoire, nous nous intéressons, plus fondamentalement et dans une perspective à la fois comparative, éthique et critique, à l'usage qui est fait de l'IA et des systèmes experts légaux de nouvelle génération aux États-Unis, au Japon et en Europe. Nous nous interrogeons aussi sur les enjeux éthiques que soulève leur usage dans la sphère du droit : en particulier, nous envisageons la question des données personnelles et les risques liés à leur exploitation, la question de la transparence et de la gouvernance algorithmique ainsi que les limites intrinsèques de l'IA appliquée au droit. En effet, si un logiciel sans conscience peut sans conteste venir en aide au praticien, l'analyse et l'interprétation du droit peuvent-elles à leur tour être formalisées dans un langage de programmation et réduites dans un système strictement computationnel ? (**partie II – section I**).

Nous clôturons ce mémoire par de brèves considérations spéculatives sur ce que nous appelons “*common lawisation*” du droit continental (**troisième partie**) et des perspectives pour l'avenir accompagnées de considérations d'ordre éthique et philosophique (**conclusion**).

### § 3 – Exposé des enjeux

Comme on peut déjà le lire dans un article de 2014 qui étudie l'impact potentiel de techniques liées au *machine learning* sur la pratique du droit, « [t]here may be a limited, but not insignificant, subset of legal tasks that are cable of being partially automated using current AI techniques »<sup>16</sup>. Qu'en est-il des questions liées à l'office du juge et à la justice prédictive ? Verra-t-on bientôt, en droit européen et en droit belge, l'introduction et puis l'utilisation d'outils associés au *machine learning*, au *data mining* et au *predictive analytics* pour instruire des dossiers, rédiger des conclusions et compiler des pièces de procédure, voire pour rendre des jugements ? L'introduction des IA dans les prétoires est-elle la solution toute désignée à ce problème récurrent, dans certains pays d'Europe, qu'est l'arriéré judiciaire ? Pourrait-elle également permettre de réduire l'aléa judiciaire, c'est-à-dire l'imprévisibilité des décisions judiciaires ? L'IA, cette boîte de Pandore, contient-elle l'espoir d'un remède efficace et bon marché à l'engorgement des tribunaux et à l'imprévisibilité des décisions qui, respectivement, étouffe la justice et frustre le justiciable ?

<sup>16</sup> H. SURDEN, « Machine Learning and Law », *Washington Law Review*, Vol. 89, N°1, 2014, pp. 87-115.

## § 4 – Exposé du cadre juridique

Existe-t-il, actuellement, en Europe ou ailleurs, du droit positif encadrant ce qu'il y a lieu d'entendre par IA ? S'il y en a, quelle en est la portée normative ? Notons d'ores et déjà, qu'il apparaît, à première vue, qu'à l'heure actuelle, il existe plus de *soft law* que de droit positif encadrant les questions de droit liées à l'IA. La Commission européenne a, par exemple, récemment mis en place un « INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE »<sup>17</sup>. Ce groupe a publié deux rapports en avril 2019: (1) *Ethics Guidelines for Trustworthy AI* et (2) *Definition of AI: Main capabilities and scientific disciplines*. Nous aurons également l'occasion d'étudier de près une Charte éthique et un Livre blanc traitant la question des rapports entre droit et IA publiés, respectivement, en 2018 et en 2020, par la Commission européenne. Nous verrons que l'Europe et le Japon sont plus avancés en termes de *policymaking* et de régulation en ce qui concerne l'IA que ne le sont les États-Unis dans la mesure où on ne trouve pas encore, outre-Atlantique, de commission officielle d'experts étudiant l'impact et les risques de l'introduction de l'IA tant dans la société que dans les entreprises<sup>18</sup>.

## § 5 – Hypothèse de recherche

En bref, ce mémoire consiste en une étude approfondie de :

(1) l'histoire intellectuelle de l'IA et de l'automatisation ainsi que son impact spécifique sur la profession d'avocat et la pratique du droit (recherche appliquée) et ;

(2) questions de droit, de politique et d'éthique (*policymaking* et régulation) liées à l'interaction entre la technologie et l'humain (recherche fondamentale) et, plus spécifiquement, les possibilités et limites ainsi que les problèmes et bénéfices possibles de l'intégration de l'IA dans des cabinets d'avocats.

Nous ne nous proposons donc pas d'étudier de façon excessivement générale l'IA du point de vue du droit, mais bien plutôt, nous visons l'analyse des conséquences de l'interaction entre

<sup>17</sup> Voy. I. LIETZEN, *Robots: Legal Affairs Committee Calls for EU-wide Rules*, Eur. Parliament News, 2017, disponible sur <http://www.europarl.europa.eu/news/>, consulté en avril 2019.

<sup>18</sup> R. CALO, « Artificial Intelligence Policy: A Primer and Roadmap », UC Davis Law Review, N°51, 2017, pp. 399-435, disponible au lien <https://ssrn.com/abstract=3015350>, consulté en avril 2019.

l'IA – les outils liés au *machine learning*, au *natural language processing*, au *data mining* et au *predictive analytics* – et les avocats, juristes et parajuristes travaillant au sein des cabinets.

Plus spécifiquement, nous nous intéresserons aux questions liées au changement paradigmatique qu'engendre l'introduction, dans la sphère du droit, d'outils liés à ces nouvelles technologies : « [s]oftware tools are beginning to affect various aspects of lawyers' work, including those tasks that historically relied upon expert human judgment, such as predicting court outcomes [...] These developments will predictably transform both how lawyers do legal work and resolve disputes on behalf of their clients. In the longer term, it is difficult to predict the impact of artificially intelligent tools will be, as lawyers incorporate them into their practice and expand their range of services on behalf of clients »<sup>19</sup>.

Dit plus simplement, notre question de recherche se formule de la façon suivante : comment et dans quelle mesure les cabinets d'avocats, notamment les petites et moyennes structures, en Europe, pourront-elles bénéficier d'outils d'automation liés à des logiciels mettant en œuvre de puissants algorithmes ? Quel en sera l'impact sur la pratique du droit ? Quels sont les tenants et aboutissants de ce phénomène ? Quelles possibilités et limites, problèmes et bénéfices possibles de l'intégration de l'IA dans des cabinets d'avocats ?

Nous aurons par ailleurs l'occasion de développer la thèse suivant laquelle l'introduction du *machine learning*, du *data mining*, du *predictive analytics* et plus largement de l'IA dans la sphère du droit risque bien de constituer un véritable changement de paradigme et de révolutionner la pratique du droit<sup>20</sup>. La réception en droit européen continental de tradition civiliste de ces nouveaux outils de comparaisons et, partant, de prédictions analytiques et casuistiques (*computational model of casuistic reasoning*) peut également être perçue comme symptomatique de ce que l'on pourrait appeler une “*common lawisation*” du droit. En effet, à l'instar du développement des modes alternatifs de résolution des conflits qui nous viennent notamment des États-Unis et du Canada et qui privatisent, dans une certaine mesure, la justice et viennent faire concurrence aux cabinets d'avocats et à ce que l'on pourrait appeler le

<sup>19</sup> B. ALARIE, A. NIBLETT, H. YOON, « How Artificial Intelligence Will Affect the Practice of Law », *University of Toronto Law Journal*, N°68 (supplement 1), 2018, pp. 106-124.

<sup>20</sup> Par les notions de “changement de paradigme” ou de “révolution” nous entendons, au sens de T. KUHN, une véritable « transformation dans la vision du monde » de cette science – certes sociale – qu'est le droit.

Cfr T.S. KUHN, *La structure des révolutions scientifiques*, Paris, Flammarion, 1983, p. 157 et s.

“circuit judiciaire classique”, il semble légitime de se demander si et dans quelle mesure l’introduction des IA dans les cabinets et les prétoires ainsi que la généralisation de son utilisation risquent de renforcer ce phénomène de privatisation du droit continental.

Bien plus, les progiciels mettant en œuvre de l’IA et des outils liés au *machine learning* intégrés dans la pratique des avocats et des juristes de tradition continentale et civiliste ne risquent-ils pas d’hisser au rang de principe général de droit la règle du précédent (*stare decisis*), pourtant caractéristique de l’ordonnement juridique des pays de *common law* ?

Enfin, il y a lieu de se demander si et dans quelle mesure l’utilisation de logiciels mettant en œuvre de l’IA pourrait accentuer davantage la distorsion de concurrence entre les petites et les moyennes ou grandes structures en créant une inégalité des armes entre les cabinets d’avocats ayant ou non accès à des outils technologiques de pointe tel que les logiciels d’analyse prédictive du patrimoine jurisprudentiel.

## II. MÉTHODOLOGIE

Au niveau méthodologique, nous nous proposons d'adopter une démarche essentiellement comparative. Plus précisément nous aurons égard au droit américain dit de *common law* et au droit européen. À ce titre, nous nous intéressons particulièrement aux sociétés et aux start-ups actives dans le domaine du droit et de l'IA (*AI and Law*) et précurseurs en la matière au travers de l'analyse discursive de communications d'entreprise. Nous tâchons par ailleurs de nous référer, autant que faire se peut, à la jurisprudence et à la législation européenne et internationale, en particulier à la *soft law*, afin d'en tirer des enseignements et par là même étayer nos propos.

Cela étant dit, ce mémoire a une dimension résolument historiciste et philosophique en ce qu'il retrace le développement de l'idée même de la création d'une intelligence synthétique et, plus spécifiquement, de l'association entre IA et droit – du paradigme des anciens systèmes experts légaux à celui de l'apprentissage machine – grâce aux différentes publications et ouvrages doctrinaux que nous avons pu rassembler.

Ce mémoire envisage par ailleurs les questions éthiques que soulève l'intégration d'outils liés à l'IA (en particulier du *machine learning*) dans le monde du droit et, plus particulièrement, dans les cabinets d'avocats. Ce faisant, nous posons la question hautement spéculative du futur des cabinets d'avocat et de l'évolution de la pratique du métier de juriste.

Le caractère original de la présente démarche tient donc en ce qu'elle se veut principalement juridique mais également, accessoirement, comparative, sociologique et philosophique. Ce mémoire se veut donc pluridisciplinaire, voire transdisciplinaire, en ce qu'il a pour objectif la compréhension d'un phénomène sociétal complexe *i.e.*, l'intégration d'outils informatiques mettant en œuvre de l'IA dans le monde du travail qui ne s'arrête pas, il va sans dire, à la sphère du droit et aux cabinets d'avocats.

## PREMIÈRE PARTIE

### CADRE HISTORIQUE ET PRATICO-NORMATIF

#### I. HISTOIRE INTELLECTUELLE DE L'IA À TRAVERS LES ÂGES

« Si chaque instrument, en effet, pouvait, sur un ordre reçu, ou même deviné, travailler de lui-même, comme les statues de DÉDALE, ou les trépieds d'HÉPHAÏSTOS, “qui se rendaient seuls – dit le poète – aux réunions des dieux” ; si les navettes tissaient toutes seules ; si l'archet jouait tout seul de la cithare, les entrepreneurs se passeraient d'ouvriers, et les maîtres, d'esclaves »<sup>21</sup>.

« Any sufficiently advanced technology is indistinguishable from magic »<sup>22</sup>.

Aujourd'hui plus que jamais encore, la réalité dépasse la (science) fiction. Pour s'en convaincre, examinons la « Résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique », dans laquelle le Parlement européen considère que (A) « [...] depuis la créature de FRANKENSTEIN imaginée par Mary SHELLEY au mythe antique de PYGMALION, en passant par le golem de Prague et le robot de Karel ČAPEK, inventeur du terme [*i.e.*, le terme “robot”], les humains ont rêvé de construire des machines intelligentes, le plus souvent des androïdes à figure humaine ; (B) [...] maintenant que l'humanité se trouve à l'aube d'une ère où les robots, les algorithmes intelligents, les androïdes et les autres formes d'intelligence artificielle, de plus en plus sophistiqués, semblent être sur le point de déclencher une *nouvelle révolution industrielle* qui touchera probablement toutes les couches de la société, il est d'une importance fondamentale pour le législateur d'examiner les *conséquences et les effets juridiques et éthiques* d'une telle révolution, sans pour autant étouffer l'innovation »<sup>23</sup>.

Bien que l'IA soit *per se* une notion éminemment moderne, nous allons voir, dans les paragraphes qui suivent, que l'on peut effectivement en retracer l'idée – disons l'histoire conceptuelle<sup>24</sup> – dans l'Antiquité et le Moyen-Âge alors que la science, dans son acceptation moderne, n'en est encore qu'à ses balbutiements. L'intérêt de l'homme pour la création de

<sup>21</sup> ARISTOTE, *Politique*, Livre I, Chapitre IV, 1253b3, § 5, trad. B. SAINT-HILAIRE, Paris, Ladrance, 1874.

<sup>22</sup> A.C. CLARKES, « Clarke's Third Law », *Profiles of the Future: an inquiry into the limits of the possible*, New York, Bantam Books, 1967.

<sup>23</sup> Parlement européen, *Résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, 2015/2103(INL), 2017. Nos italiques et annotations.

<sup>24</sup> Sur les notions d'histoire conceptuelle et d'histoire intellectuelle des idées voy. *i.a.*, les travaux R. KOSELLECK et la notion de *Begriffsgeschichte* chez HEGEL. Cf. A. LOVEJOY, *The Great Chain of Being: A Study of the History of an Idea*, New York, Harper, 1960 ; M. FOUCAULT, *L'archéologie du savoir*, Paris, Gallimard, 1969.

machines “intelligentes” est en effet fort ancien : « [s]ur un plan scientifique, les développements théoriques et pratiques sur les automates d’une part, et les travaux en logique mathématique d’autre part, en constituent les fondements »<sup>25</sup>. Bien plus, on pourrait dire que, d’un point de vue psychanalytique, ce désir de créer des machines automates, androïdes ou non, relève, sans nul doute, d’un fantasme archaïque profondément ancré dans l’inconscient collectif et l’histoire des civilisations. D’aucuns parlent d’ailleurs à ce titre de « force de séduction du numérique », ou plus largement, de ce que l’on pourrait appeler « l’attrait libidinal des technologies » agissant comme « amplificateurs de notre vie psychique »<sup>26</sup>. Par cette tentative de généalogie de l’IA, nous entendons souligner l’ancienneté de ce phantasme archaïque de (re)création artificielle du vivant, du sentient et de l’intelligent qui, aujourd’hui, si l’on en croit les transhumanistes, pourrait bien être en passe de transformer l’homme en « Deus ex Machina », au sens littéral, *i.e.*, en un dieu issu de la machine.

## § 1 – Antiquité

De façon générale, on peut dire que l’idée même d’êtres intelligents synthétiquement créés remonte loin dans le passé. S’agirait-il d’un fantasme archaïque éminemment masculin, à savoir celui d’être capable de créer ou plutôt de construire la vie quand cette aptitude, qui relève proprement du miracle<sup>27</sup>, semble exclusivement réservée à ce que l’on pourrait appeler, dans le sillage de S. DE BEAUVOIR, le “deuxième sexe” ? À moins que l’idée de pouvoir trouver le moyen d’échapper à certaines tâches serviles, répétitives, voire aliénantes, soit la raison véritable qui se cache derrière ce souhait d’avoir, à sa disposition, un esclave automate qui ne connaît nul maux, ni fatigue ni même la mort ; qui ne se plaint et ne se révolte jamais ? Après tout, ne dit-on pas que fainéantise est mère d’intelligence ?

Dans la culture occidentale dont les origines sont classiquement associées au monde hellénistique et à la mythologie grecque, la première instance mythique d’un être synthétique,

<sup>25</sup> J.-P. et M.-C. HATON, *L’intelligence artificielle*, coll. Que sais-je ?, Paris, PUF, 1989, p. 5.

<sup>26</sup> Voy. M. HUNYADI, « Du sujet de droit au sujet libidinal. L’emprise du numérique sur nos sociétés », *Esprit*, vol. mars, n°3, 2019, p. 118, disponible au lien <https://www.cairn.info/>, consulté en mai 2020, cité dans A. GARAPON, J. LASSÈGUE, « Mythe de la délégation aux machines », *Des véhicules autonomes à l’intelligence artificielle*, Bruxelles, Éditions Larcier, 2020, p. 217.

<sup>27</sup> N.L. HERZFELD, « In Our Image: The Desire for Artificial Intelligence », *In Our Image – Artificial Intelligence and the Human Spirit*, Minneapolis, Fortress Press, 2002, pp. 1-9. Ne parle-t-on pas d’ailleurs du “miracle de la vie” ?

automate et doué de réflexion émane d'HÉPHAÏSTOS (Ἡφαίστος), le dieu du feu et de la métallurgie. Certaines de ses créations comme ses automates dorés<sup>28</sup> ou encore son fameux géant de bronze, TALOS (Τάλως)<sup>29</sup>, protecteur et gardien de l'île de Crète qui effectuait, à un intervalle régulier, le tour du royaume du Roi MINOS (Μίνως)<sup>30</sup> peuvent en effet faire penser à ce que l'on pourrait désigner anachroniquement sous le vocable "robot"<sup>31</sup>. Il est pourtant singulier, en effet, tant pour les Anciens que pour les Modernes, que ces êtres faits d'airain et de métal puissent être capables de se mouvoir par eux-mêmes comme s'ils étaient doués de volition et de réflexion. Le mythe de PYGMALION constitue également, en ce sens, un autre exemple d'un être fait non d'airain mais d'ivoire cette fois qui prend vie grâce à l'intervention divine d'APHRODITE et s'anime tant et si bien que son créateur, le sculpteur GALATÉE, en tombe éperdument amoureux<sup>32</sup>. Enfin, c'est à DÉDALE (Δαίδαλος), célèbre sculpteur et architecte athénien, que revient l'honneur d'être le premier mortel à avoir pu rendre l'inerte animé ou, du moins, d'en avoir donné l'illusion, grâce à ses automates et à ses innovations techniques en matière de sculpture et d'ingénierie relevant, il va sans dire, plus du simulacre que de la robotique<sup>33</sup>.

## §2 – Moyen-Âge

Au Moyen-Âge, on trouve à nouveau trace de ce désir archaïque de création synthétique du vivant – du sentient et de l'intelligent – et, plus largement, de cette volonté d'animer ce qui est, à l'origine, inanimé et inerte, chez les mystiques et les adeptes des

<sup>28</sup> HOMER, *Iliade*, trad. E. LASSERRE, Paris, Garnier Frère, coll. Flammarion, 1965, livre XVIII, 376, 418-419 et 468-473, pp. 307-333 ; voy. également HÉSIODE, *Théogonie* ; P. WERNER, « Le forgeron boiteux », *Légendes de la Grèce Antique*, Genève, Éditions Minerva, 1980, pp. 51-52.

<sup>29</sup> APOLLONIOS DE RHODES, *L'Expédition des Argonautes ou la Conquête de la Toison d'Or – Poème en quatre Chant*, Chant IV, trad. française, J.-J.-A. CAUSSIN, Paris, Desrez, 1838 ; P. WERNER, « JASON et les Argonautes », *Op. cit.*, pp. 113-117.

<sup>30</sup> A. MARCINKOWSKI et J. WILGAUX, « Automates et créatures artificielles d'HÉPHAÏSTOS : entre science et fiction », *Techniques & Culture*, pp. 43-44, 2004, disponible au lien <https://doi.org/10.4000/tc.1164>, consulté en février 2020.

<sup>31</sup> Le terme « robot » est en effet bien postérieur ; il date du début du XX<sup>ème</sup> siècle et d'une pièce de théâtre de SF de l'auteur tchèque Karel ČAPEK, à savoir *R.U.R - Rossum's Universal Robots*. Il est par ailleurs intéressant de relever ici que *robota* signifie « corvée » en tchèque et que *rob* veut dire « esclave » en slave ancien.

<sup>32</sup> OVIDE, *Les métamorphoses*, trad. J. CHAMONARD, Paris, Garnier Frère, coll. Flammarion, 1966, livre X, pp. 260-261 ; P. WERNER, « La statue du Pygmalion », *Op. cit.*, p. 201.

<sup>33</sup> A. MARCINKOWSKI et J. WILGAUX, *Ibid. idem*

sciences et techniques occultes. Ainsi, Jābir IBN HAYYĀN dit GEBER, un alchimiste persan du VIII<sup>ème</sup> siècle, concentrait ses travaux alchimiques sur ce qu'il nommait *Takwin* (تكوين) i.e., la génération artificielle d'êtres-vivants par l'Alchimie<sup>34</sup>. Semblablement, au XIII<sup>ème</sup> siècle, Albertus MAGNUS avait, dit-on<sup>35</sup>, réussi à créer un "androïde" bien que, selon la légende, « cet automate fut anéanti par saint Thomas D'AQUIN qui le brisa à coups de bâton »<sup>36</sup>. Theophrastus VON HOHENHEIM dit PARACELSE aurait également réussi, au XVI<sup>ème</sup> siècle, à concevoir un homoncule, c'est-à-dire un être-humain miniature artificiellement créé en laboratoire<sup>37</sup>. Il en va de même, toujours au XVI<sup>ème</sup> siècle, de Rabbi Yehouda LEVAÏ BEN BETZALEL dit le MAHARAL, un rabbin, talmudiste, mystique et philosophe, que l'on associe avec le Golem (גולם) de Prague i.e., un humain artificiel fait d'argile et qui aurait pris vie par l'inscription, sur son front, de l'un des noms de Dieu et l'introduction, dans sa bouche, d'un parchemin sur lequel était inscrit le nom Ineffable<sup>38</sup>. On le voit, tant à l'Antiquité qu'au Moyen-Âge et jusqu'au début de l'époque moderne avec PARACELSE et le MAHARAL, les récits à la fois mythiques et, vraisemblablement, fantasmagoriques de la génération synthétique d'êtres vivants – à la fois sentients et intelligents – sont légions.

### §3 – Époque moderne

C'est pourtant à l'époque moderne et dans le sillage des travaux précurseurs de René DESCARTES (voy. la dichotomie corps-esprit, l'idée de l'animal-machine et, par extension, le

<sup>34</sup> Voy. C.G. JUNG, *Alchemical Studies*, New York, Princeton University Press, 1983 ; W.R., NEWMAN et A., GRAFTON (ed.), *Astrology and Alchemy in Early Modern Europe*, Cambridge, MIT Press, 2006.

<sup>35</sup> D. DIDEROT et J. LE ROND D'ALEMBERT (eds), *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*, etc., University of Chicago, ARTFL Encyclopédie Project, Fall 2017 Edition, R. MORRISSEY et G. ROE (eds.), disponible au lien <http://encyclopedia.uchicago.edu/>, consulté en février 2020. Voy. les entrées « Androïde », « Automates » par D'ALEMBERT et « Alchimie » par MALOUIN. D'après ce dernier, les principaux auteurs d'Alchimie sont GEBER, LE MOINE, BACON, RIPLEY, LULLE, Jean et Isaac LE HOLLANDOIS, Basile VALENTIN, PARACELSE, VAN ZUCHTEN et SENDIGOVIVUS. Plusieurs d'entre eux auraient tenté, en laboratoire, de façon opérative ou symbolique, de recréer du vivant. Voy. également à ce sujet U. ECO, « Le discours alchimique et le secret différé » dans *Les limites de l'interprétation*, Paris, Grasset, coll. essais, 1992, pp. 87-105.

<sup>36</sup> C. DE PLANCY, *Dictionnaire infernal*, Paris, Henri Plon, 1863, pp. 13-15.

<sup>37</sup> C.G. JUNG, *Op. cit.*, pp. 157-160 ; P. MCCORDUCK, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*, New York, W.H. Freeman and Company, 1979, p. 11-12.

<sup>38</sup> N.L. HERZFELD, *Op. cit.*, p. 2.

concept de l'homme-machine<sup>39</sup> repris et développé plus tard, en 1748, par Julien OFFRAY DE LA METTRIE<sup>40</sup>) qu'apparaissent, au cours du XVII<sup>ème</sup> siècle, les premières machines à calculer<sup>41</sup>. La première machine d'arithmétique est l'œuvre du mathématicien et philosophe Blaise PASCAL. Elle sera ensuite améliorée par LEIBNIZ et obtiendra, finalement, ses lettres de noblesse avec Charles BABBAGE et son *analytical engine* au cours du XIX<sup>ème</sup> siècle<sup>42</sup>. En effet, cette suite d'innovations en matière de mécanique, de logique et de mathématique constitue les fondations sur lesquelles sera bâti ni plus ni moins que l'ordinateur moderne et, par extension, ce que l'on pourrait appeler, par abus de langage, les premières "machines pensantes" dont la première réalisation *bona fide* est sans conteste la machine universelle d'Alan TURING (1936) et l'architecture de John VON NEUMANN (1945).

Enfin, du côté de la SF et des récits qui appartiennent plutôt au genre du fantastique, à l'époque moderne, on peut notamment citer *Frankenstein; or, The Modern Prometheus* de Mary SHELLEY en 1818 ou encore *Darwin among the Machines* de Samuel BUTLER en 1863. En ce qui concerne l'ouvrage de SHELLEY, il est intéressant de relever que c'est après avoir cherché vainement la solution dans « la magie de PARACELSE » que le Docteur FRANKENSTEIN se résout à se tourner vers la science moderne pour donner vie à sa création<sup>43</sup>. C'est d'ailleurs ce qui explique qu'il est considéré que cet ouvrage, qui précède ceux de Jules VERNES de plusieurs décennies, se démarque du fantastique dit gothique et constitue la première véritable œuvre de SF dans la mesure où, à proprement parler, il n'y a pas de SF tant qu'il n'y a pas de science<sup>44</sup> ; la magie et l'hermétisme doivent avoir été relégués au second plan. SHELLEY et BUTLER inaugurent ainsi un thème qui connaîtra un grand succès dans la

<sup>39</sup> R. DESCARTES, *Discours de la méthode suivi d'extraits de la dioptrique, des météores, des météores, de la vie de Descartes par Baillet, du Monde, de l'Homme et de lettres*, Paris, Garnier-Flammarion, 1966. Voy. *Traité de l'homme*, extraits sur la « machine » du corps pp.231-234 et la V<sup>ème</sup> partie du *Discours de la méthode* pp. 67-95

<sup>40</sup> J.O. DE LA METTRIE, *L'homme-machine*, Paris, Gallimard, coll. folio essais, 1999.

<sup>41</sup> Voy. G. CHAZAL, *À quoi rêvent les machines ?* Dijon, EUD, 2016 et, du même auteur, *Philosophie de la machine : néo-mécanisme et post-humanisme*, Dijon, EUD, 2013.

<sup>42</sup> Pour une histoire détaillée du développement des premières machines à calculer et de l'IA en général, voy. i.a., P. MCCORDUCK, *Op. cit.* et V. PRATT, *Machines à penser : une histoire de l'intelligence artificielle*, Paris, PUF, 1995. À noter que l'*Ars Magna* de l'écrivain mystique, philosophe et alchimiste Raymond LULLE (c. 1300) pourrait, à certains égards, constituer la toute première entreprise de conception d'une machine à calculer.

<sup>43</sup> K. CANVAT, *La science-fiction*, Bruxelles, Didier Hatier, 1991, p. 26.

<sup>44</sup> J. GATTÉGNO, *La science-fiction*, Paris, PUF, coll. Que sais-je ? 1971, pp. 7-36 ; J.-L. STEINMETZ, *La littérature fantastique*, Paris, PUF, coll. Que sais-je ? 1990, pp. 27-29. Voy. également C. AZIA et J. GOIMARD, *Encyclopédie de poche de la SF*, Paris, Presses Pocket, 1986.

littérature de SF : le thème de la création se retournant contre son créateur – thème, au demeurant, éminemment œdipien et qui, du reste, est déjà présent dans le mythe du GOLEM à ceci près que la création dont il est question a été artificiellement créée par la science moderne et non pas la magie. Ce thème est d'ailleurs souvent repris au sujet du développement fulgurant qu'a connu l'IA à la fin du XX<sup>ème</sup> siècle : il s'agit du mythe de la singularité technologique à savoir ce moment critique où l'intelligence artificielle dépasse celle des humains et devient dès lors à la fois imprévisible et incontrôlable<sup>45</sup>.

#### §4 – Époque contemporaine

Bien que le concept soit bien plus ancien et que l'on puisse le faire remonter au mythe de TALOS, de PYGMALION, on encore du GOLEM et, dans la littérature de SF, à FRANKENSTEIN et R.U.R<sup>46</sup>, l'IA commence à prendre forme de façon concrète, on l'a vu, à l'époque moderne, avant de prendre son envol dans les années 1950 qui marquent, en quelque sorte, sa naissance officielle. L'IA au sens contemporain du terme voit en effet le jour en août 1956 lors d'une conférence restée célèbre : *The Dartmouth Summer Research Project on Artificial Intelligence* à Hanover, New Hampshire. Quatre noms sont associés avec l'invention de l'expression « *artificial intelligence* ». Il s'agit de John MCCARTHY (jeune assistant adjoint de Dartmouth), Marvin MINSKY (junior *research fellow* en mathématique et en neurologie à Harvard), Nathaniel ROCHESTER (manager d'un centre de recherche d'IBM) et Claude SHANNON (mathématicien des laboratoires BELL déjà bien connu à l'époque pour sa théorie probabiliste de l'information) ; « [a]lthough they came from different backgrounds – there were men trained as mathematicians, as psychologists, as electrical engineers – and although some worked for industries and others were at universities, they had in common a *belief* (more like a *faith* at that point) that what we call *thinking* could be understood in a formal and

<sup>45</sup> Voy. à ce sujet, J.-G. GANASCIA, *Op. cit.*, Le thème de la « singularité technologique » a d'ailleurs été illustré dans divers romans de SF qui ont fait date et marqué la postérité du genre : en particulier *Do Androids Dream of Electric Sheep?* de Philip K. DICK en 1966 et *Robot Dreams* d'Isaac ASIMOV paru en 1986. On peut encore citer des films de SF plus récents tels que *I, Robot* réalisé par A. PROYAS en 2004, *Transcendance* réalisé par W. PFISTER en 2014 ou encore la série *Westworld* de J. NOLAN et L. JOY. À ce mythe de la singularité, on peut encore ajouter celui de la « délégation aux machines » que l'on pourrait aussi appeler celui du « grand remplacement ». Voy. à ce sujet A. GARAPON, J. LASSÈGUE, « Mythe de la délégation aux machines », *Des véhicules autonomes à l'intelligence artificielle*, Bruxelles, Éditions Larquier, 2020, pp. 207-221.

<sup>46</sup> D. TURCQ, « L'intelligence artificielle appliquée », *Travailler à l'ère post-digitale ; quel travail pour 2030*, Malakoff, Dunod, 2019, p. 31.

scientific way, and that the best nonhuman instrument for doing it was the digital computer »<sup>47</sup>. C'est en effet cette croyance, cette *foi*, pour reprendre l'expression de MCCORDUCK, en la possibilité de formaliser les processus mentaux de réflexion et de pensée (*thinking*) – ce que nous pourrions appeler de façon large “intelligence” – qui fonde les premières recherches scientifiques concrètes sur l'IA. Pourtant, d'un point de vue épistémologique, il paraît d'emblée clair que l'on ne saurait réduire l'intelligence humaine à une capacité de calcul ou à un algorithme, aussi complexe soit-il<sup>48</sup>. C'est sans aucun doute sur ce point, sur lequel nous aurons l'occasion de revenir *infra*, que l'IA montre toutes ses limites ; il manquera toujours à un programme informatique, aussi complexe et élaboré soit-il, cette plus-value, cet “on-ne-sait-quoi” que l'on désigne abstraitement sous le vocable “esprit”, ou “conscience” (ou encore *mind* en anglais) ; *i.e.*, il lui manquera toujours cette “visée intentionnelle” pour parler comme les philosophes adeptes de la phénoménologie<sup>49</sup> – visée intentionnelle qui s'incarne avant tout dans un *corps propre*, dans la chair, et non un corps purement matériel fut-il animé de façon artificielle<sup>50</sup>.

Cela étant, il n'empêche qu'il convient de prendre acte de l'importance qu'a pris l'IA ces dernières décennies, nonobstant les périodes d'hibernation que son développement a pu connaître au cours de la deuxième moitié du XX<sup>ème</sup> siècle, surtout qu'elle apparaît, à bien des égards, comme une technologie fondamentalement disruptive (au même titre que l'électricité, le numérique/digital, l'internet ou le Web 2.0) ; l'IA change la donne et a, d'ores et déjà, un impact concret au niveau social, sociétal, médical ainsi que dans la sphère du travail comme nous le verrons en ce qui concerne le domaine du droit.

<sup>47</sup> P. MCCORDUCK, *Op. cit.*, p. 93 et s. nos italiques. À cette liste non-exhaustive, il convient d'ajouter A. SAMUEL, T. MOORE, R. SOLOMONOFF, O. SELFRIDGE, A. NEWELL et H. SIMON qui assistèrent et apportèrent leur contribution à la fameuse conférence. Voy. S. BRINGSJORD et N.S. GOVINDARAJULU, « Artificial Intelligence », The Stanford Encyclopedia of Philosophy, Winter 2019 Edition, E. N. ZALTA (ed.), disponible au lien <https://plato.stanford.edu/>, consulté en février 2020.

<sup>48</sup> Voy. à ce sujet J. SEARLE, « Minds, Brains and programs », *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980 et en particulier l'argument de la *Chinese Room* ; cf. W.V.O. QUINE, « Mind versus Body », *Quiddities : An Intermittently Philosophical Dictionary*, London, Belknap Press of Harvard University Press, 1987.

<sup>49</sup> J.-P. DUPUY, « L'obstacle de l'intentionnalité », *Aux origines des sciences cognitives*, Paris, La découverte, 1999, pp. 96-102. Voy. également les travaux fondateurs de F. BRENTANO et de E. HUSSERL en ce qui concerne le concept de l'intentionnalité et la phénoménologie en tant que discipline philosophique.

<sup>50</sup> Voy. les travaux de M. HEIDEGGER et, en particulier, sa notion de *Dasein*. Cf. avec *La phénoménologie de la perception* de M. MERLEAU-PONTY et sa notion de corps propre/phénoménal ou encore de chaire ontologique.

D'aucuns pensent à ce titre que :« [l']intelligence artificielle va nous “augmenter” en nous permettant de disposer de savoir-faire que notre cerveau seul ne peut maîtriser »<sup>51</sup>. On voit poindre ici en filigrane les fantasmes transhumanistes associés au développement de l'IA et au mythe de la singularité que nous avons déjà évoqués *supra*<sup>52</sup>.

## II. AI-POWERED TOOLS DANS LA PRATIQUE DU DROIT: DU PARADIGME DES LEGAL EXPERT SYSTEMS À CELUI DU MACHINE LEARNING

Après avoir posé quelques jalons historiques et philosophiques dans les paragraphes qui précèdent, nous essayons à présent de définir précisément ce que l'on entend par IA quand il est question d'outils d'automation mis en œuvre dans des systèmes d'apprentissage machine et des progiciels de gestion de cabinets d'avocats. Comment définit-on l'IA dans ce contexte ? Existe-t-il, actuellement, en Europe ou ailleurs, du droit positif encadrant ce qu'il y a lieu d'entendre par IA ? S'il y en a, quelle en est la portée normative ?

### §1 – Définitions et enjeux

De façon large, on peut dire que « l'[IA] est un ensemble de règles et d'algorithmes mathématiques et informatiques qui analysent un ensemble de données et proposent un ensemble d'utilisation à partir de ces données »<sup>53</sup>. Plus précisément, l'IA « englobe l'ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine [...]. Elle désigne donc les systèmes qui font preuve d'un comportement intelligent en analysant leur environnement et en prenant des mesures – avec un certain degré d'autonomie – pour atteindre des objectifs spécifiques. Ces systèmes peuvent être purement logiciels, agissant dans le monde virtuel [*e.g.*, ALEXA de AMAZON ou SIRI de APPLE], logiciels d'analyse d'images, moteurs de recherche ou systèmes de reconnaissance vocale et faciale »<sup>54</sup>. D'après l'INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE mis en place par la

<sup>51</sup> D. TURCQ, *Op. cit.*, p. 25. Voy. également G. HOTTOIS, *Philosophie et idéologies trans/posthumanistes*, Paris, coll. Librairie philosophique J. Vrin, 2017 ; a contrario et en ce qui concerne plus spécifiquement le droit J.-B. HUBIN, H. JACQUEMIN, B. MICHAUX, *Le juge et l'algorithme : juges augmentés ou justice diminuée ?* Bruxelles, Larcier, coll. CRIDS, 2019.

<sup>52</sup> Voy. à ce sujet, J.-G. GANASCIA, *Op. cit.*

<sup>53</sup> D. TURCQ, *Op. cit.*, p. 34.

<sup>54</sup> AI4BELGIUM, *Sondage sur la perception des belges quant à l'IA*, 2019, disponible au lien <https://www.ai4belgium.be/>, consulté en mai 2019.

Commission Européenne et dont le but est d'établir une base commune sur cette technologie émergente qu'est l'IA ; d'en définir les contours, les différents aspects et surtout de fournir des recommandations pour les régulations européennes (*policies recommendations*) ainsi que des lignes directrices éthiques (*ethics guidelines*) concernant l'IA<sup>55</sup>, celle-ci se définit comme suit :

« Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take *to achieve the given goal*. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions »<sup>56</sup>.

Ajoutons encore ceci: l'IA se définit de façon plus adéquate en considérant d'abord et avant tout le *but poursuivi* dans le développement de ladite IA. Au-delà de la dichotomie discutée *supra* entre *strong* et *weak AI*, il y a lieu de considérer non plus deux mais quatre types d'IA possibles organisés selon les buts poursuivis au cours de leur développement : à savoir, d'une part, de se rapprocher le plus possible d'une capacité de réflexion et de raisonnement semblable à celle des humains (*human-based*) ou plutôt d'un idéal de rationalité (*ideal rationality*) et, d'autre part, de concevoir une IA censée raisonner et penser (*reasoning-based*) ou plutôt censée agir dans le monde (*behavior-based*). Ces quatre possibilités organisées selon le but poursuivi couvriraient l'ensemble des recherches – littératures scientifiques et fictionnelles comprises – existant sur le sujet de l'IA :

|                         | Human-Based                                                                                                                    | Ideal Rationality                                                                                                  |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| <b>Reasoning-Based:</b> | Systems that “think” like humans.<br><i>e.g., “machines with minds in the fullest sense” .. Westworld TV serie</i>             | Systems that “think” rationally.<br><i>e.g., legal expert systems, ERP, law/chatbots, machine learning systems</i> |
| <b>Behavior-Based:</b>  | Systems that act like humans.<br><i>e.g., TURING test and VOIGHT-KAMPPF test in Blade Runner and Philip K. DICK’s SF novel</i> | Systems that act rationally.<br><i>e.g., “automation of intelligent behavior”; autonomous cars, LAWs</i>           |

*Four Possible Goals for AI According to AIMA<sup>57</sup>*

<sup>55</sup> INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE SET UP BY THE EUROPEAN COMMISSION (AI HLEG), *A definition of AI: Main capabilities and scientific disciplines*, 2019, disponible au lien <https://ec.europa.eu/newsroom/>, consulté en mai 2019.

<sup>56</sup> AI HLEG, *Ibid.*, p.6. Nos italiques.

<sup>57</sup> S. RUSSELL, P. NORVIG, *Artificial Intelligence: A Modern Approach 3rd edition*, Saddle River, NJ, Prentice Hall, 2009 cité dans S. BRINGSJORD et N.S. GOVINDARAJULU, *Op. cit.*, Nos italiques avec nos exemples.

Plus spécifiquement, dans le domaine du droit, l'IA trouve à s'appliquer dans ce que l'on appelle les "*legal expert systems*" (ci-après SEL) qui relève de la catégorie *ideal/reasoning* du tableau ci-dessus dans la mesure où : « [u]n système expert est un [...] logiciel capable d'atteindre les performances d'un expert humain pour une tâche donnée, ceci dans un domaine restreint et bien délimité, grâce à l'exploitation d'un ensemble de connaissances [ou données] fournies explicitement et acquises pour l'essentiel auprès d'experts du domaine considéré »<sup>58</sup>. Autrement dit, les SEL sont des outils numériques artificiels, des logiciels, capables de reproduire partiellement les mécanismes cognitifs rationnels d'un expert humain, dans un domaine donné et notamment de répondre à des questions plus ou moins précises à partir de données connues ou de règles déterminées et ce, en implémentant de l'IA<sup>59</sup>. Pour ce qui concerne le droit, un SEL peut être utile pour :

- « L'études de normes ;
- La liquidation de dossiers de retraite ;
- La gestion de réglementation complexe ;
- L'aide à la rédaction d'actes (contrats, actes notariés, etc.) ;
- L'aide à l'utilisation de textes légaux »<sup>60</sup>.

Au début des années 1990, en Belgique, « l'emploi de techniques dérivées de l'[IA] dans l'informatique juridique semble quasiment inexistante »<sup>61</sup>. En effet, il y a 30 ans, cette technologie émergente n'en est encore qu'à ses débuts ; on comprend dès lors combien ces nouvelles technologies implémentant de l'IA appliquée au droit ont pu évoluer en l'espace de trois décennies. Les logiciels et SEL présentés par la délégation belge au Conseil de l'Europe lors du « 9<sup>e</sup> Symposium sur l'informatique juridique » paraissent aujourd'hui pour le moins désuets. Il est frappant de remarquer que la Belgique, en particulier le réseau universitaire

<sup>58</sup> J.-P. et M.-C. HATON, *Op. cit.*, p. 69.

<sup>59</sup> C'est à Stanford, Californie, qu'au cours des années 70, apparaissent les premiers SE avec notamment HEURISTIC DENDRAL un SE permettant d'analyser et d'identifier les constituants chimiques d'un matériau ou encore PUFF qui diagnostiquait les maladies respiratoires et PROSPECTOR dans le domaine de la recherche de minerai. Voy. J.-F. FOGEL, « La percée de l'intelligence artificielle », *Le Point*, n°500, Paris, 3/10/1983, p. 34.

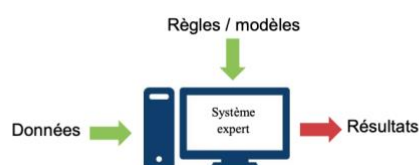
<sup>60</sup> J.-P. et M.-C. HATON, *Op. cit.*, pp. 74-75. Cette liste est bien entendu loin d'être exhaustive.

<sup>61</sup> D. BROUWERS et M.F. ROBBEN, « Rapport sur les systèmes experts juridiques en Belgique », délégation de la Belgique, *Actes du 9<sup>e</sup> Symposium sur l'informatique juridique en Europe*, 10-12 octobre 1989, Bonn ; « Les systèmes basées sur l'intelligence artificielle dans le domaine juridique », Strasbourg, Conseil de l'Europe, Affaires juridiques, 1991, p. 64.

belge, n'a rien à envier, à l'époque du moins, aux allemands ou aux français : elle est en fait le seul pays à présenter une série de solutions informatiques qui ne sont pas, contrairement à ses voisins européens, de simples projets pilotes<sup>62</sup>. Les SEL sont donc essentiellement des IA faibles censées penser de façon rationnelle et incarnées, si l'on peut dire, dans des logiciels développés dans un but précis. Ces SEL sont capables de reproduire certains mécanismes cognitifs d'un expert humain, tel un juriste ; en effet, tout bien considéré, qu'est-ce qu'un avocat ou un paralégal si ce n'est un expert du droit éventuellement spécialisé dans l'une ou l'autre des branches (les différents domaines) de cet arbre aux multiples racines (les sources et les principes) qu'est le droit et qui, par le truchement de raisonnements logiques et rationnels, trouve des solutions à des conflits et des problèmes relevant du droit ?

Dans le domaine du droit, le *natural language processing* (NLP), le *machine learning* (ML) et l'analyse automatisée de la jurisprudence à l'aide du *data mining* (DM) sont véritablement ce qui constitue le cœur des logiciels développés actuellement. En ce sens, nous l'avons vu, ces logiciels ne sont pas autre chose que de l'IA de type *weak* bien que ceux-ci sont aujourd'hui tout à fait capables d'extraire des tendances complexes (ou *patterns*) et d'apprendre par eux-mêmes grâce au NLP, au ML et au DM<sup>63</sup>. Enfin, il est important de souligner que les SEL tels que décrits ci-avant constituent, en quelque sorte, les ancêtres des nouvelles IA mettant en œuvre du ML comme l'illustrent les deux schémas<sup>64</sup> ci-dessous :

Graphique.2 : Les anciens systèmes experts étaient programmés avec des règles reproduisant la logique du raisonnement juridique

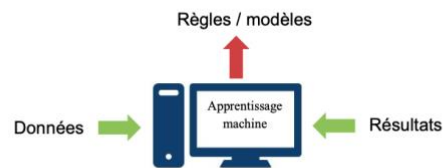


<sup>62</sup> Sont cités et présentés les logiciels PRISME (SE permettant d'identifier les aides publiques auxquelles les entreprises établies en Belgique peuvent prétendre), SPECTRE (SE étudiant l'impact des aides publiques sur les résultats de l'entreprise), M.I.J.A (SE d'aide à la décision juridique en matière de sécurité sociale) ainsi qu'une série de systèmes documentaires tels que JUSTEL, CREDOC, JUDIT et JURIDIS.

<sup>63</sup> Commission européenne pour l'efficacité de la justice, « Caractéristiques de fonctionnement de l'intelligence artificielle (apprentissage machine) appliquée aux décisions judiciaires », *Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement*, adoptée par la CEPEJ lors de sa 31<sup>ème</sup> réunion plénière à Strasbourg, CEPEJ(2018)14, 2018, pp. 24-28.

<sup>64</sup> CEPEJ, *Charte éthique*, *Ibid.*, p. 26.

Graphique 3 : L'apprentissage machine produit seul des modèles en recherchant automatiquement des corrélations entre données et résultats.



Comme indiqué dans les légendes des schémas repris ci-dessus, les anciens SE étaient programmés avec des règles reproduisant la logique du raisonnement juridique (*e.g.*, le fameux syllogisme judiciaire). De ce fait, ceux-ci étaient fortement limités à cause de leur incapacité intrinsèque à interpréter « l’infinité des situations et des contextes »<sup>65</sup> et, en conséquence, devenaient peu efficaces au-delà d’une centaine de règles d’encodage, tant en termes de performance d’exécution que de maintenance ; le raisonnement suivi par le système devenait pour ainsi dire impossible à appréhender pour ses concepteurs, ce qui pose évidemment problème en termes de transparence<sup>66</sup>. En revanche, les nouveaux systèmes d’apprentissage machine qui viennent donc supplanter les anciens SE des années 1990 sont capables de procéder eux-mêmes à l’identification de modèles statistiques existant dans les données. Ils peuvent ensuite les mettre en correspondance avec des résultats spécifiques, ce qui leur permet, *in fine*, d’adapter leur raisonnement aux situations qui se présentent à eux et, ce faisant, apporter une aide précieuse aux avocats, juristes et parajuristes pour des tâches nettement plus complexes concernant, par exemple :

- La modélisation formelle et/ou computationnelle de raisonnements juridiques (argumentation, raisonnement “en fait” et “en droit”) ;
- L’analyse prédictive (*predictive analytics*), l’aide à la décision (*decision making*), à l’évaluation du risque juridique et de l’aléa judiciaire (*risk assessment pricing and timeline predictions of litigation*) et à la construction de barèmes de sommes allouées plus précis dans différents types de contentieux (*e.g.*, *LEGAL INSIGHTS* voy. *infra*);

<sup>65</sup> H. DREYFUS, *What Computers Still Can’t Do. A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992 cité par D. CARDON, *À quoi rêvent les algorithmes, nos vies à l’heure des Big Data*, Paris, Seuil, 2015 p. 59 et CEPEJ, *Charte éthique*, *Ibid.*, *idem*. À noter que la critique de l’IA par DREYFUS paraît en partie dépassée aujourd’hui. Voy. A. BONNET, J.-P. HATON, J.-M. TRUONG-NGOC, R. LOTTÈS, *Systèmes-experts : vers la maîtrise technique*, Paris, Inter Éditions, coll. Informatique intelligence artificielle, 1986

<sup>66</sup> CEPEJ, *Charte éthique*, *Ibid.*, *idem*. Voy. *Infra*, partie II, §3, enjeux éthiques et risques, b) *Black-box AI, explainability* et transparence.

- La catégorisation des contrats selon des critères multiples et la détection de clauses contractuelles divergentes, contradictoires ou incompatibles ;
- L'étude et la comparaison analytiques des normes, des principes et de la jurisprudence dans des bases de données avec un ou des cas factuels (moteur de recherche et de comparaison de jurisprudence avancé avec, le cas échéant, analyse prédictive) ;
- “Chatbots” de renseignement du justiciable ou de support de celui-ci dans le cadre d'une démarche litigieuse ;
- “Lawbots” pour automatiser des tâches légales peu complexes et répétitives ;
  - *Automatic legal text classification and summarization* ;
  - *Automated information extraction from legal databases and texts* ;
  - *Conceptual or model-based legal information retrieval* ;
- La résolution de litiges en ligne<sup>67</sup>.

## §2 – Avantages, inconvénients et limites

L'IA telle qu'implémentée dans les nouveaux SEL mettant en œuvre du ML au travers de technologies comme la *Robotic Process Automation* avec laquelle le système apprend les tâches à automatiser en observant le comportement d'experts humains ou encore procède lui-même à l'identification de modèles statistiques existants dans les données comporte nombre d'avantages : en particulier, des gains significatifs en termes de temps, d'argent et de productivité. En effet, actuellement, « [c]ompanies are achieving productivity gains by using software robots to perform routine, rules-based service processes. If implemented well, such automation can result in *high-performing human-robot teams, in which software robots and human employees complement one another* [...] Although the term “robot” brings to mind visions of electromechanical machines that perform human tasks, the term as it relates to service automation refers to something less threatening: software that performs certain repetitive and dreary service tasks previously performed by humans, *so that humans can focus on more unstructured and interesting tasks* »<sup>68</sup>. Autrement dit et pour tordre le cou une fois pour toutes à ce stéréotype qu'est le mythe de la singularité et celui, souvent relayé par les médias et le septième art, du “grand remplacement” de l'humain par les machines ; il y a lieu de considérer que l'IA, dans la sphère du travail et notamment du droit, est avant tout un outil

<sup>67</sup> Voy. « État des lieux de l'utilisation des algorithmes d'intelligence artificielle dans les systèmes judiciaires des États membres du Conseil de l'Europe », *Charte éthique*, Annexe 1, *Ibid.*, p. 14. Nos annotations.

<sup>68</sup> Voy. <https://sloanreview.mit.edu/article/>, consulté en mars 2020. Nos italiques.

fondamentalement « additif »<sup>69</sup>, voire *supplétif*, c'est-à-dire qui s'ajoute et vient en aide à l'intelligence humaine en lui épargnant le côté proprement laborieux de certaines tâches répétitives, voire serviles, de la profession la rendant par la même occasion moins pénible et, partant, permettant aux travailleurs de chair et d'os de se concentrer sur d'autres tâches plus stimulantes sur le plan intellectuel ou faisant davantage appel aux *soft* qu'aux *hardskills*.

Néanmoins, force est de rappeler, qu'à l'heure actuelle, l'IA n'est pas capable de :

- Inventer, imaginer en connectant des éléments disparates, faire preuve de créativité ;
- Ressentir des émotions (bien qu'il soit possible d'en simuler) ;
- Avoir une conscience de soi et de ses actes ; ainsi le logiciel ALPHAGO développé par GOOGLE DEEPMIND a battu, en 2017, le champion de monde de Go (围棋)<sup>70</sup> ... sans le savoir. Il en va de même de DEEPBLUE qui a battu Garry KASPAROV aux échecs en 1997 ou encore de WATSON, l'IA d'IBM, contre les champions de *Jeopardy!* en 2011 ;
- Avoir un jugement proprement humain c'est-à-dire non seulement rationnel mais aussi *émotionnel* ou encore faire preuve d'empathie, de regrets, etc.<sup>71</sup>.

En effet, « les IA dites “fortes” [*strong AI*] de la littérature de science-fiction n'existent pas : ce type d'IA, qui seraient douées non seulement d'intelligence mais aussi de conscience, demeurent purement fictionnelles. Les systèmes d'apprentissage machine [*machine learning system*] actuellement en cours de développement sont décrits comme des IA “faibles” et sont capables d'extraire des tendances complexes et d'apprendre à partir de grandes quantités de données de manière efficace et souvent avec des niveaux élevés de précision prédictive »<sup>72</sup>. Si les *legaltechs* « admettent elles-mêmes qu'elles n'intègrent pas encore [l'IA] dans leur programme conduisant à organiser des procédures en ligne [...] [e]n revanche, des modes de calculs sophistiqués dépassant la capacité humaine sont aujourd'hui bien réels, ce qui permet aux machines d'intégrer par elles-mêmes des informations »<sup>73</sup>. Enfin, si l'idée selon laquelle on pourrait remplacer le juge ou l'avocat relève bel et bien encore du mythe et du fantasme –

<sup>69</sup> D. TURCQ, *Op cit.*, p. 49.

<sup>70</sup> Il s'agit du plus ancien jeu de stratégie combinatoire abstrait connu comparable, dans une certaine mesure, au jeu d'échecs ou de dames.

<sup>71</sup> Voy. D. TURCQ, *Op. cit.*, pp. 39-40.

<sup>72</sup> CEPEJ, *Charte éthique*, *Op. cit.*, p. 25. Voy. également A. BENSAMOUN et G. LOISEAU, (dir.) *Droit de l'intelligence artificielle*, Les Intégrales 15, Paris, Librairie générale de droit et de jurisprudence, 2019.

<sup>73</sup> A. BENSAMOUN et G. LOISEAU, *Op. cit.*, p. 187.

du moins en ce qui concerne l'Europe<sup>74</sup> – il semble néanmoins établi que, de nos jours, les logiciels implémentant de l'IA et mettant en œuvre du *machine learning* tels que développés par nombre d'entreprises spécialisées dans le domaine juridique s'avèrent être des outils à la fois utiles, précieux, voire même incontournables, pour tout qui cherche à travailler de façon efficace et de concert avec les nouvelles technologies. Bien plus, il est probable que dans un avenir proche, « des applications à forte valeur ajoutée, et donc forcément onéreuses, de *machine learning* puissent avoir une efficacité supérieure à l'expérience et au « bon sens » contentieux des avocats travaillant leurs dossiers de manière traditionnelle. L'utilisation de ces applications pourrait ainsi encore plus accentuer la distorsion de concurrence et une inégalité des armes entre des cabinets d'avocats ayant eu ou non recours de manière onéreuse à de tels logiciels d'analyse "prédictive" des jurisprudences »<sup>75</sup>. En d'autres termes, le fait qu'un cabinet d'avocats puisse avoir à sa disposition des SEL performants risque bien de constituer un avantage considérable vis-à-vis de la concurrence. Dit de façon plus lapidaire encore, il est probable que les meilleurs cabinets de demain, ceux avec le "*success rate*" le plus élevé, seront ceux-là même qui seront équipés des meilleures technologies en ce compris, au premier chef, des logiciels mettant en œuvre du NLP, du DM, et du ML ainsi que des outils de *predictive analytics*.

La question qui semble se dégager de prime abord des considérations qui précèdent est la suivante : peut-on réduire le travail du juge, de l'avocat et plus généralement du juriste à l'exercice de la qualification juridique, à l'application du syllogisme judiciaire et à la résolution de problèmes d'interprétation du droit ? N'y a-t-il pas, par ailleurs, quelque chose d'irréductible et d'inquantifiable qu'aucun SEL, aussi complexe soit-il, ne pourra jamais saisir – quelque chose qui relève, par exemple, de l'empathie ou de l'intuition ? Si les SEL ne doivent être compris que comme une intelligence additive et supplétive au travail des juristes, il devient évident que cette chaîne de questionnement ne constitue pas, en soi, un véritable obstacle : l'IA ne fait que venir en aide au juriste ; en aucun cas elle ne le remplace.

Cette première objection balayée d'un revers de main, adressons plutôt la question faitière à laquelle nous avons déjà fait allusion *supra* et sur laquelle nous devons encore revenir, à savoir celle des limites de l'IA. En effet, pour créer une IA qui serait à la fois complète et consciente (*strong AI*, à la fois *human reasoning* et *behavior-based AI*) et non pas simplement

<sup>74</sup> Nous verrons *infra* que la situation est tout autre aux États-Unis, au Japon ou encore en Chine.

<sup>75</sup> CEPEJ, *Charte éthique*, *Op. cit.*, pts. 45, p. 22.

faible (*goal-oriented, ideal et reasoning-based AI*), il faudrait encore pouvoir comprendre parfaitement la nature de la conscience humaine qui demeure encore et toujours – pour l’essentiel du moins – une boîte noire pleine de mystères ; la conscience, l’expérience subjective, voire même l’activité cérébrale elle-même, demeurent aujourd’hui encore d’absolus irréductibles<sup>76</sup>. Notre connaissance du fonctionnement de cet organe gras qu’est le cerveau est, pour reprendre l’expression du philosophe J. SEARLE, tout au plus, « maigre »<sup>77</sup>.

D’un point de vue psychologique cette fois, privée de conscience de soi, d’intentionnalité (ou, pour parler comme M. MERLEAU-PONTY d’un « corps propre »), et, a fortiori, faute de connaissance adéquate du cerveau humain *id quod est*, il manquera toujours à l’IA l’intelligence émotionnelle et, au premier chef, l’empathie *i.e.*, cette émotion proprement humaine et nécessaire à la pratique du droit qu’a priori aucun programme ne saurait reproduire fidèlement et authentiquement. Pour le surplus, selon SEARLE, comme l’illustre bien sa fameuse expérience de pensée dite de la « *Chinese Room* », le projet de réalisation d’IA forte s’avère impossible à réaliser dans la mesure où les états mentaux sont avant tout des phénomènes biologiques : « Aucun programme informatique n’est en lui-même suffisant pour donner un esprit à un système. Bref, le programme n’est pas un esprit, et ne suffit pas en lui-même pour posséder un esprit [...] Conscience, intentionnalité, subjectivité, relation de causalité mentale, toutes relèvent de l’histoire de notre vie biologique, tout comme la croissance, la reproduction, la sécrétion de la bile, ou la digestion »<sup>78</sup>. En effet, les programmes d’ordinateur fonctionnent sur une base purement formelle (en particulier, avec la logique modale et computationnelle) et dès lors *syntactique*, alors que l’esprit procède de contenus mentaux *i.e.*, de contenus *sémantiques*. Ainsi, pour R. SCHANK, sommité en matière d’IA et de sciences cognitives, « no computer can understand when you tell it something », et

<sup>76</sup> T. NAGEL, « What Is It Like to Be a Bat? », *The Philosophical Review*, Vol. 83, N°4, 1974, pp. 435-450.

<sup>77</sup> J. SEARLE, *Du cerveau au savoir*, Paris, Hermann, 1985, p. 9. Voy. aussi, du même auteur : *Intentionality: An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press, 1983.

<sup>78</sup> J. SEARLE, « Les ordinateurs peuvent-ils penser ? », *Ibid.* pp. 37-56. Sur la « *Chinese Room* » voy. *idem* pp. 42-43 et, du même auteur, « Minds, Brains and programs », *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980. Cf. G. M. EDELMAN, « Ordinateurs numériques: la fausse analogie », *Biologie de la conscience*, Paris, Éditions Odile Jacob, 1992, pp. 288-300 ; H. PUTNAM, *Représentations et Réalité*, Paris, Gallimard, 1990.

WATSON, la célèbre IA d'IBM sur laquelle ROSS – sans doute l'un des SEL les plus performants disponible actuellement – s'appuie « doesn't know what it is saying »<sup>79</sup>.

Nonobstant les limites inhérentes au projet de réalisation d'une *strong AI* et le fait que, pour reprendre, en la paraphrasant, la célèbre maxime rabelaisienne, « [intelligence] sans conscience n'est que ruine de l'âme », il demeure qu'un problème se pose du fait de l'inégalité des armes entre, d'une part, les cabinets disposant de technologies de pointe, telle que des SEL ou des ERP de nouvelle génération implémentant NLP, ML et DM, et, d'autre part, les cabinets n'ayant pas les ressources opérationnelles et/ou financières nécessaires pour en bénéficier.

On peut cependant conjecturer qu'à l'avenir, pour les avocats et les juristes, les *softskills* tels que le sens commun, l'intelligence émotionnelle, l'éthique professionnelle ou les compétences interpersonnelles, pour n'en citer que quelques-unes, pourraient devenir aussi, voire même plus importants encore, que les *hardskills* tels que la connaissance du droit au sens large que des systèmes ML sont *in fine* mieux à même de maîtriser de par leurs capacités de calcul, de mémorisation et de comparaison analytique et casuistique proprement surhumaines<sup>80</sup>.

### §3 – *Soft law* européen

Existe-t-il, actuellement, en Europe ou ailleurs, du droit positif encadrant ce qu'il y a lieu d'entendre par IA ? S'il y en a, quelle en est la portée normative ? À l'heure actuelle, en Europe, il existe plus de *soft law* que de droit positif *stricto sensu* encadrant les questions de droit liées à l'IA. La Commission européenne a d'ailleurs récemment mis en place un « INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE »<sup>81</sup>. Ce groupe a publié deux rapports en avril 2019: (1) *Ethics Guidelines for Trustworthy AI* et (2) *Definition of AI: Main capabilities and scientific disciplines* que nous avons utilisés *supra* pour dégager

<sup>79</sup> R. SCHANK, « Machines that Think are in the Movies », J. BROCKMAN (ed.), *What to Think About Machines that Think*, New York, Harper Collins, 2015.

Voy. D. Cole, « The Chinese Room Argument », *The Stanford Encyclopedia of Philosophy*, 2020, E.N. ZALTA (ed.), disponible au lien <https://plato.stanford.edu/>, consulté en février 2020.

<sup>80</sup> La devise du logiciel ROSS, la plateforme de recherche jurisprudentielle américaine que nous avons mentionnée *supra*, est de faire plus, mieux et plus rapidement qu'humainement possible.

Voy. <https://rossintelligence.com/>, consulté en mars 2020.

<sup>81</sup> I. LIETZEN, *Op. cit.*

une définition complète et nuancée de l'IA. Nous avons aussi mentionné précédemment la Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique qui concerne, comme son nom l'indique, davantage les robots intelligents, autonomes et éventuellement capables d'auto-apprentissage à travers l'expérience et les interactions (*hardware-based / ideal and behavior-based AI*) que les *software-based AI (ideal and reasoning-based AI)* telle qu'implémentée dans les nouveaux SEL mettant en œuvre du ML. Il est en effet plus question, dans cette résolution, de moyens de transport autonomes, de drones, de robots de soins à la personne, de robots médicaux et de la réparation et de l'amélioration artificielle du corps humain (systèmes cyber-physiques)<sup>82</sup> que de logiciels intelligents au sens faible.

En revanche, il existe depuis 2018 une « Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement » (ci-après la Charte éthique) qui a été adoptée par la Commission Européenne pour l'efficacité de la justice (ci-après CEPEJ). Cette Charte éthique, nous le verrons, s'avère particulièrement pertinente pour notre propos. De plus, depuis février 2020, la Commission européenne a publié un livre blanc intitulé : « Intelligence artificielle : Une approche européenne axée sur l'excellence et la confiance » (ci-après le Livre blanc). Nous allons passer en revue tour à tour ces deux sources de *soft law* dans les pages qui suivent. D'abord, en ce qui concerne la Charte éthique, celle-ci définit cinq grands principes repris ci-dessous et sur lesquelles nous reviendrons *infra* :

- « 1. **Principe de respect des droits fondamentaux** : Assurer une conception et une mise en œuvre des outils et des services d'IA qui soient compatibles avec les droits fondamentaux ;
2. **Principe de non-discrimination** : Prévenir spécifiquement la création ou le renforcement de discriminations entre individus ou groupes d'individus ;
3. **Principe de qualité et sécurité** : En ce qui concerne le traitement des décisions juridictionnelles et des données judiciaires, utiliser des sources certifiées et des données intangibles avec des modèles conçus d'une manière multidisciplinaire, dans un environnement technologique sécurisé ;
4. **Principe de transparence, de neutralité et d'intégrité intellectuelle** : Rendre accessibles et compréhensibles les méthodologies de traitement des données, autoriser les audits externes ;
5. **Principe de maîtrise par l'utilisateur** : Bannir une approche prescriptive et permettre à l'utilisateur d'être un acteur éclairé et maître de ses choix »<sup>83</sup>.

<sup>82</sup> Parlement européen, *Résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, 2015/2103(INL), 2017, pts. 24-40.

<sup>83</sup> CEPEJ, *Charte éthique*, *Op. cit.*, p. 2.

Par ailleurs, la Charte soulève un certain nombre de questions d'ordre éthique sur lesquelles nous aurons l'occasion de revenir dans la deuxième partie de ce mémoire : en particulier, le renforcement de biais cognitifs, les risques de discrimination et la question de la neutralité des algorithmes par les politiques publiques d'*open data* et d'*open access* des décisions judiciaires, les principes de transparence et de maîtrise par l'utilisateur, le risque de cristallisation de la jurisprudence selon une logique reproductive, de profilage du juge et de *forum shopping*, ou encore le problème que pose le caractère indéterminé et évolutif des concepts juridiques.

Ensuite, en ce qui concerne le Livre blanc qui se trouve être d'une actualité brûlante puisqu'il date de février 2020, celui-ci tend, à l'instar de la Charte éthique, à souligner les risques potentiels du développement de l'IA : « Le recours à l'IA peut porter atteinte aux valeurs sur lesquelles l'UE est fondée et entraîner des violations des droits fondamentaux, tels que les droits à la liberté d'expression et de réunion, la dignité humaine, l'absence de discrimination fondée sur le sexe, l'origine raciale ou ethnique, la religion ou les convictions, le handicap, l'âge ou l'orientation sexuelle, selon le cas, la protection des données à caractère personnel, le respect de la vie privée ou le droit à un recours juridictionnel effectif et à un procès équitable, ainsi que la protection des consommateurs »<sup>84</sup>. Nous aurons l'occasion de revenir sur ces questions dans la section III de la première partie de ce mémoire qui concerne spécifiquement les rapports entre IA et droits humains fondamentaux ainsi que, de façon plus philosophique, dans la deuxième partie de ce mémoire.

Résolution, Charte éthique, Livre blanc... dans tous les cas, il y a lieu de relever qu'il s'agit là d'actes non-contraignants dans lesquels l'UE fait part de son point de vue sur l'IA ; ces actes préparatoires ne créent pas *directement* d'obligation juridique et ont *ipso facto* tout au plus une valeur politique. Il s'agit donc bien de *soft law* et non de droit positif à proprement parler. On souligne donc, qu'à l'heure actuelle, il n'existe pas encore de droit positif qui encadrerait directement l'IA. Par contre, une série de directives européennes qui, pour certaines, ont déjà été transposées en droit belge peuvent trouver à s'appliquer de façon *indirecte* dans le cadre

<sup>84</sup> Commission européenne, *Livre blanc sur l'intelligence artificielle : une approche européenne axée sur l'excellence et la confiance*, Bruxelles, COM(2020), 2020, p. 12.

d'un litige impliquant des sociétés ou des particuliers ayant eu affaire à des solutions informatiques mettant en œuvre de l'IA<sup>85</sup>.

#### **§4 – Avis d'experts**

Nous avons évoqué dans l'introduction de ce mémoire la possibilité que l'IA puisse constituer un remède efficace et bon marché à (1) l'engorgement des tribunaux et (2) l'imprévisibilité des décisions judiciaires. En effet, d'un côté, il est notoire que l'arriéré judiciaire belge est si important qu'il arrive que des juridictions ne puissent rendre de décision à temps de sorte qu'elles dépassent tout délai raisonnable et violent, en conséquence, le droit des justiciables à l'accès à un tribunal garanti par l'article 6 § 1 de la Convention européenne des droits de l'homme (ci-après CEDH). L'État belge, souvent mis en demeure de remédier à la situation par l'ordre des avocats, s'est d'ailleurs vu assigné à plusieurs reprises devant les tribunaux par Avocats.be qui déplore la situation : « à la Cour d'appel de Bruxelles, le délai d'audience [sic] est de quatre à cinq ans, ce n'est pas acceptable ; le tribunal de l'entreprise néerlandophone à Bruxelles a deux ans de retard ; [...] ; la Cour du travail de Bruxelles a supprimé toute une série d'audiences ; le parquet de Bruxelles, fin de l'année dernière, a annoncé ne plus poursuivre les actes de harcèlement et classé 1.700 dossiers de signalement de délits pénaux... On peut multiplier les exemples à l'infini mais cela veut dire qu'aujourd'hui, un justiciable est face à une justice qui ne fonctionne plus normalement »<sup>86</sup>. D'un autre côté, en ce qui concerne l'aléa judiciaire, force est de constater que « [q]uelle que soit la tradition juridique du pays, l'aléa judiciaire, c'est-à-dire le risque de voir sa demande en justice satisfaite ou rejetée, suscite l'envie de pouvoir être quantifié à l'aide [des] nouvelles applications technologiques. Les avocats (ou les services juridiques d'une entreprise) y voient la possibilité d'apporter à leurs clients un conseil mieux éclairé en évaluant de manière empirique et systématique les chances de succès d'une procédure, mais aussi de favoriser la conclusion de transactions [ou, plus largement de recourir aux modes alternatifs de résolution des conflits] permettant, le cas échéant, d'échapper à un long et coûteux procès. Certains assureurs proposent déjà l'utilisation de systèmes prédictifs à leurs clients pour évaluer les

<sup>85</sup> Voy. la législation européenne citée dans la bibliographie du présent mémoire.

<sup>86</sup> Propos de J.-P. BUYLE recueilli par T. JEUNEJEAN dans l'article « Avocats.be attaque l'État belge en justice. Interview de Jean-Pierre Buyle », 2019, disponible sur <https://www.justice-en-ligne.be>, consulté en février 2020.

mérites de leur affaire »<sup>87</sup>. L'IA et les *legaltechs* seraient-elles des solutions toutes désignées pour remédier à ces problèmes ? Voyons ce qu'en disent différents acteurs du monde du droit.

a. CRIDS (UNAMUR)

Le Centre de Recherche Information, Droit et Société (ci-après CRIDS) qui fête ses 40 ans cette année est situé dans la faculté de droit de l'Université de Namur. D'après son site internet officiel : « This interdisciplinary research center is a member of the Namur Digital Institute (NADI) and is focused on a wide spectrum of information society-related topics: Platforms; Artificial Intelligence, Big Data & IoT; Governance & Digital citizenship; E-Health; Digital and Technology Literacy; E-justice Cybersecurity & Surveillance »<sup>88</sup>. Depuis près de 10 ans, le CRIDS étudie par ailleurs les transformations qu'engendre l'émergence des nouvelles technologies et des méthodes d'analyse statistique, telles que le ML, le *big data* ou l'IA et ce, tant sur le plan légal que sociétal<sup>89</sup>. Au vu des différentes publications du CRIDS, il apparaît que la position adoptée par ce laboratoire de recherche est plutôt nuancée, ni technophile ni technophobe, les publications du CRIDS tendant plutôt à analyser de façon neutre, interdisciplinaire et scientifique le changement de paradigme que suscite la dissémination des nouvelles technologies dans la sphère du droit. L'unité de recherche qui a trait aux interactions entre les technologies et la société définit sa vision de la technologie comme suit « “Technology and Society” merges the anthropological, philosophical and sociological approaches to understand how software and AI systems, considered as cognitive membranes or filters, may affect our perception, our cognition, our action, our decision and more widely our social interaction. Information technologies are deeply shaping contemporary human conditions and social order. To a certain extent, these technologies are “micro-politics” that endorse, in their concepts and designs, moral and political choices that

<sup>87</sup> CEPEJ, *Charte éthique*, *Ibid*, p. 34.

<sup>88</sup> Voy. <https://www.crids.eu/>, consulté en mars 2020.

<sup>89</sup> Voy. les publications récentes du CRIDS : *i.a.*, A. ROUVROY, M. HILDEBRANDT, *Law, human agency and autonomic computing: the philosophy of law meets the philosophy of technology*, London, Routledge, 2011 ; A. ROUVROY, « Gouvernamentalité algorithmique et perspectives d'émancipation : le disparate comme condition de l'individuation par la relation ? », *Réseaux*, n° 177, 2013, pp. 163-196 ; J. GROSMAN, F. THOREAU, « Dans la peau d'un algorithme », *L'Appel des entités fragiles : enquêter avec les modes d'existence de Bruno Latour*, Liège, Presses Universitaires de Liège, 2018, pp. 53-72 ; J.-P. BUYLE, A. VAN DEN BRANDEN, « La robotisation de la justice », H. JACQUEMIN, A. DE STREEL, *L'intelligence artificielle et le droit*, *Op. cit.*, pp. 259-317 ; J.-B. HUBIN, H. JACQUEMIN, B. MICHAUX, *Le juge et l'algorithme : juges augmentés ou justice diminuée ? Op. cit.*

affect our relationships to ourselves, others and the world at large »<sup>90</sup>. Deux grands axes de recherche structurent les recherches menées au sein de cette unité : « The first axis is inspired by the “value sensitive design” and the “social shaping of technologies” theories and methods. It aims to integrate “social and ethical” requirements from the very starting stage of technological design in order to foster its social acceptability. The second axis regards the critical analysis of the moral, social and political choices embedded into concepts and design of new information systems in order to assess their potential impacts for the human conditions, the social interactions and the partition of responsibilities between humans and machines »<sup>91</sup>. Nous verrons en effet *infra* que quand il s’agit d’éthique appliquée à l’IA il semble fondamental de privilégier une approche *ethical by design*, c’est-à-dire de s’assurer que sont intégrées, dès la phase de conception et d’apprentissage machine (*i.e.*, *ex ante*), des règles interdisant de porter atteinte directement ou indirectement à certaines valeurs fondamentales<sup>92</sup>.

#### b. CODEX (STANFORD)

CODEX est un laboratoire de recherche de la faculté de droit de Stanford : « At CODEX, researchers, lawyers, entrepreneurs and technologists work side-by-side to advance the frontier of legal technology, bringing new levels of legal efficiency, transparency, and access to legal systems around the world. CODEX’s emphasis is on the research and development of computational law – the branch of legal informatics concerned with the *automation and mechanization of legal analysis* »<sup>93</sup>. Il semble assez clair, au vu de cette déclaration d’intention, par contraste avec le CRIDS, que CODEX fait partie de ces acteurs qui, pour reprendre le mot de MCCODRUCK, font preuve d’une foi inébranlable dans les capacités de l’IA à améliorer le fonctionnement de la justice en la rendant plus efficace et plus transparente. En effet, l’histoire de l’université de Stanford, située au cœur de la Silicon Valley, « province bénie du capitalisme électronique »<sup>94</sup>, est intrinsèquement liée au développement des technologies du numérique et de l’IA : le premier article traitant des rapports possibles entre IA et droit est d’ailleurs paru dans la *Stanford Law Review*, nous

<sup>90</sup> Voy. <https://www.crids.eu/units/>, consulté en mars 2020, Nos italiques.

<sup>91</sup> *Ibid.*, *Idem*.

<sup>92</sup> Singulièrement, l’autonomie et la dignité de la personne conçue au sens large, la prudence au sens des Anciens et la protection des personnes vulnérables, voy. *infra*.

<sup>93</sup> Voy. <https://law.stanford.edu/codex/>, consulté en mars 2020, Nos italiques.

<sup>94</sup> J.-F. FOGEL, *La percée de l’intelligence artificielle*, Le Point, n°500, Paris, 3/10/1983, p. 34. Voy. annexe.

l'avons dit, dès les années 70. On peut y lire, dans l'introduction, cette déclaration audacieuse qui constitue, au demeurant, un véritable plan de recherche : « This research suggests that computer science may assist lawyers in both the study and performance of their reasoning processes. [...] [T]he time has come for serious interdisciplinary work between lawyers and computer scientists to explore the computer's potential in law »<sup>95</sup>. La perspective technophile et un quasi fidéisme technologique est ici, à l'instar de la déclaration d'intention de l'initiative CODEX, pleinement assumée. Au-delà de toute considération morale, il reste néanmoins que, comme insiste, J.-P. BUYLE, la robotisation de la justice est un « phénomène irrésistible ». En effet, « le système judiciaire belge et le monde juridique dans son ensemble doivent s'y préparer. La numérisation, d'abord partielle, puis totale des sources juridiques [*open access*] et des procédures judiciaires [*open data*] constituera le fondement nécessaire au développement de l'intelligence artificielle. L'application de cette technologie émergente aux fonctions de recherches juridiques, de prédictive judiciaire et de résolution de litige achèvera le processus de robotisation de la justice »<sup>96</sup>.

### c. SOCIÉTÉS ET START-UPS ACTIVES DANS LE DOMAINE

Actuellement, en 2020, il existe des centaines de sociétés et de start-ups spécialisées en *legaltechs* développant des solutions destinées aux cabinets d'avocats. Ces sociétés peuvent être classées dans plusieurs catégories selon les technologies sur lesquelles elles se basent : *Document Automation*, *Practice Management*, *Legal Research*, *Online Dispute Resolution*, *E-Discovery*, *Predictive Analytics*, etc<sup>97</sup>. Il faut également souligner que beaucoup de ces entreprises mettent en œuvre ce que l'on pourrait appeler des outils implémentant de l'IA (*AI-powered tools*). En effet, nous l'avons vu, "IA" est un terme polysémique, à la fois vague et vaste, qui désigne au moins quatre types d'intelligence différentes selon le but poursuivi par ceux qui la développent (voy. la classification discutée *infra* de RUSSELL et NORVIG) et recouvre par ailleurs beaucoup d'applications technologiques concrètes. Nous reprenons dans le tableau ci-dessous les principales sociétés développant des logiciels mettant en œuvre de l'IA et destinés aux cabinets d'avocats. Les logiciels sont classés par ordre (supposé)

<sup>95</sup> B.G. BUCHANAN, T.E. HEADRICK intitulé « Some Speculation about Artificial Intelligence and Legal Reasoning », *Stanford Law Review*, Vol. 23, N°1, 1970, pp. 40-62.

<sup>96</sup> J.-P. BUYLE, A. VAN DEN BRANDEN, « La robotisation de la justice », H. JACQUEMIN, A. DE STREEL, *L'intelligence artificielle et le droit*, *Op. cit.*, p. 317.

<sup>97</sup> <http://techindex.law.stanford.edu/>, consulté en mars 2020.

d'importance : des grosses entreprises américaines leaders dans le domaine aux petites start-ups en passant par les sociétés européennes d'envergure moyenne.

| Logiciel                                                                                         | Pays d'origine               | Technologies mises en œuvre <sup>98</sup>                                                                                                                                                                                                                                                                                                                                                                                                        |
|--------------------------------------------------------------------------------------------------|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>WATSON / ROSS (IBM)</b>                                                                       | États-Unis                   | IA, NLP, DM, <i>big data</i> , <i>legal research and analytics</i> , <i>cloud computing</i> , en part. <i>machine learning</i> , <i>grammatical structure</i> , <i>word embeddings</i> , <i>facts &amp; motions (advanced syntax)</i> etc.                                                                                                                                                                                                       |
| <b>WESTLAW EDGE (THOMSON REUTERS)</b>                                                            | États-Unis                   | IA, NLP, DM, ML, <i>big data</i> , <i>legal research and analytics</i> , en part. <i>intelligent document analysis</i> , <i>integrated litigation analytics</i> , <i>logical AI/inferencing</i> , etc.                                                                                                                                                                                                                                           |
| <b>RAVEL LAW, LEXIS 360, LEX MACHINA, LEXIS ADVANCE, JURISDATA ANALYTICS, etc. (LEXIS NEXIS)</b> | États-Unis                   | IA, NLP, DM, ML, <i>big data</i> , <i>legal research and analytics</i> , en part. <i>cloud computing</i> , <i>advanced legal analytics</i> , <i>docket research</i> , <i>litigation analytics</i> , <i>practical guidance</i> , <i>powerful data visualization and predictive analytical tools</i> (via RAVEL VIEW), <i>predictive analytics</i> (via LEGISLATIVE OUTLOOK), <i>computational models of evidential and legal reasoning</i> , etc. |
| <b>FAST CASE NEXTCHAPTER</b>                                                                     | États-Unis                   | IA, NLP, DM, ML, <i>big data</i> , <i>legal research and analytics</i> , <i>citation analysis</i> , <i>sorting tools</i> , <i>data visualization</i> ( <i>patented interactive timeline view of search results</i> ), <i>cloud-based computing</i> , <i>practice management</i> , <i>bankruptcy software</i> (via NEXTCHAPTER), etc.                                                                                                             |
| <b>LUMINANCE</b>                                                                                 | Angleterre                   | IA, NLP, DM, ML ( <i>pattern-recognition</i> ), <i>big data</i> , <i>legal research and analytics</i> , en part. <i>Advanced statistical probability analysis</i> , <i>supervised and unsupervised ML</i> , etc.                                                                                                                                                                                                                                 |
| <b>KLEOS, DLEX, LEGISWAY, LEGAL INSIGHTS, JURA (WOLTERS KLUWER)</b>                              | Pays-Bas (Belgique) (France) | IA, NLP, DM, en part. <i>practice management</i> (via KLEOS, DLEX, LEGISWAY), <i>cloud computing</i> , <i>big data</i> (via SMART TOOL), <i>legal research and analytics</i> (via JURA), <i>predictive analytical tools</i> (via LEGAL INSIGHTS), etc.                                                                                                                                                                                           |
| <b>PREDICTICE (partenariat avec WOLTERS KLUWER)</b>                                              | France                       | IA, NLP, DM, ML, <i>big data</i> , <i>legal research and analytics</i> , en part. <i>predictive analytics</i> , etc.                                                                                                                                                                                                                                                                                                                             |
| <b>DOCTRINE (FORSETI)</b>                                                                        | France                       | IA, ( <i>legal intelligence</i> ), NLP, DM, <i>big data</i> , <i>legal research and analytics</i> , etc.                                                                                                                                                                                                                                                                                                                                         |
| <b>CASE LAW ANALYTICS</b>                                                                        | France                       | IA, NLP, DM, ML, <i>big data</i> , <i>legal research and analytics</i> , en part. <i>quantification de l'aléa juridique et du risque judiciaire</i> , etc.                                                                                                                                                                                                                                                                                       |
| <b>STRADA LEX (LARCIER) (partenariat avec LEXIS NEXIS)</b>                                       | Belgique                     | <i>Legal research: search engine using Boolean operators, exact expressions, truncations, combination of groups of words, use of proximity operators, cross language, related concepts</i> , etc.                                                                                                                                                                                                                                                |
| <b>JURISQUARE</b>                                                                                | Belgique                     | <i>Legal research: search engine using Boolean operators, exact expressions, truncations, combination of groups of words, use of proximity operators, cross language, related concepts</i> , etc.                                                                                                                                                                                                                                                |
| <b>SECIB AIR SECIB NÉO</b>                                                                       | Belgique                     | <i>Practice management, cloud computing, legal research and analytics</i> (via STRADA LEX), etc.                                                                                                                                                                                                                                                                                                                                                 |

<sup>98</sup> Les technologies mises en œuvres par les sociétés et les start-ups reprises dans le tableau émanent directement des déclarations faites sur leurs sites web officiels.

Les discours commerciaux de ces sociétés sont, à peu de chose près, similaires. Il s'agit pour l'essentiel de la promesse d'une plus grande efficacité et d'un avantage compétitif (*edge* en anglais) par rapport à la concurrence : de la possibilité de pouvoir travailler mieux, plus rapidement et efficacement en ayant notamment accès à de grandes bases de données et en automatisant une partie significative des tâches qui constituent la besogne quotidienne des avocats et juristes. Nous avons déjà mentionné que la devise de la plateforme ROSS, développée par IBM et basée sur l'algorithme d'IA nommé WATSON, est la suivante : « *Do more than humanly possible* » – une devise qui a le mérite de la clarté. Chez LEXIS NEXIS, autre acteur majeur du domaine, on se vante que « *LEX MACHINA enables lawyers to be better lawyers – much better lawyers. [...] Legal analytics represents a paradigm shift for lawyers. For the first time, lawyers can combine insights gleaned from bottom-up data with traditional top-down controlling authority found in statutes, rules, and court opinions. It's new, it's unorthodox and it's extremely valuable. But legal analytics also brings lawyering back to its roots. It provides lawyers with facts on which to base opinions [i.e., les “attendus” ou encore les “motifs” du droit continental qui constituent l'argumentation en droit de conclusions ou du dispositif d'une décision judiciaire] – the core element of good lawyering. Objective data replaces subjective “anec-data”* »<sup>99</sup>. On relève de nouveau ici, à l'instar de la devise de ROSS, une rhétorique que l'on pourrait qualifier de transhumaniste selon laquelle la combinaison de l'intelligence humaine et artificielle permet, pour l'avocat et le juriste, d'améliorer les performances de travail de façon significative et, partant, de devenir en quelque sorte un avocat surhumain. Chez WOLTERS KLUWER, par contraste, on insiste plutôt sur le fait que les outils informatiques proposés permettent aux petits cabinets de contrecarrer l'inégalité des armes qui pourrait résulter de l'utilisation, par les grande structures, d'outils de pointe : « *À armes égales avec les grands cabinets [...] une sélection intelligente pour les avocats indépendants* »<sup>100</sup>. L'éditeur hollandais insiste par ailleurs sur le confort qu'induit l'automatisation des tâches et des processus répétitifs et monotones (facturation, suivi de dossier, rappels, etc.) ou l'intégration parfaite de ses logiciels avec la suite MICROSOFT OFFICE et DPA (*Digital Platform for Attorneys*) dont l'utilisation se généralise, en Belgique, depuis peu.

Cela étant dit, en 2017, d'après J.-P. BUYLE et A. VAN DEN BRANDEN « il n'existe pas encore [en Belgique], de moteur de recherche basé sur la recherche sémantique [et donc le NLP] ni

<sup>99</sup> <https://lexmachina.com/what-we-do/>, consulté en mars 2020. Nos italiques et annotations.

<sup>100</sup> <https://legalworld.wolterskluwer.be/>, consulté en mars 2020. Nos italiques.

de solutions opérationnelles appliquant de l'IA à la recherche juridique » bien que des projets de développements soient en cours chez certains éditeurs de revue juridiques (*i.e.*, LARCIER et WOLTERS KLUWER)<sup>101</sup>. Il semble pourtant qu'en l'espace d'à peine trois ans la situation ait changé. En effet, entretemps, WOLTERS KLUWER a amélioré de façon significative ses logiciels phares KLEOS et DLEX en y incluant des technologies liées à l'IA, notamment avec le *smart tool*, un « assistant intelligent » qui parcourt tous les jugements pertinents et LEGAL INSIGHTS qui met en œuvre du *predictive analytics* : en effet cet « assistant digital intelligent » permet d'« estimer rapidement la probabilité de remporter des affaires en matière de rupture du contrat de travail [...] ce moteur de recherche intelligent spécialisé dans ce domaine « réfléchit » à l'aide de l'IA et du Big Data »<sup>102</sup>. Par ailleurs, on relève aussi que JURISQUARE et STRADA LEX ne semblent pas, à l'heure actuelle, utiliser d'outils liés à l'IA et au ML dans leur moteur de recherche. On peut cependant présager que cela évoluera dans les années qui viennent : en effet, STRADA LEX a récemment annoncé un partenariat avec LEXIS NEXIS qui est un acteur majeur dans le domaine de l'IA et du droit<sup>103</sup>. Il en va de même pour WOLTERS KLUWER qui a annoncé en 2017 un partenariat avec la start-up française PREDICTICE « pour consolider et commercialiser la solution PREDICTICE qui permet de calculer les chances de succès d'une action en justice grâce à ses algorithmes d'analyse de la jurisprudence [avec pour] objectif de mettre à la disposition des professionnels du droit un outil innovant de *justice prédictive* qui – en facilitant le travail sur les contentieux – contribuera à *améliorer le fonctionnement de la justice* »<sup>104</sup>. Soulignons enfin qu'il est systématiquement rappelé sur les différents sites web officiels de ces sociétés que leurs solutions informatiques ne sont pas de simples moteurs de recherche, mais bien plutôt et selon la rhétorique employée par les développeurs de ces logiciels, des solutions « intelligentes » qui permettent de prédire le résultat d'un litige ou la durée d'un procès : « Bien plus qu'un simple moteur de recherche juridique. Bâissez des stratégies gagnantes. Sécurisez vos positions. *Minimisez l'aléa juridique* pour vos clients grâce à DOCTRINE [...] *Ayez une longueur d'avance* : DOCTRINE transforme la vaste quantité de données non-structurées en informations exploitables, pour découvrir les liens sous-jacents entre des informations éparses »<sup>105</sup>. Semblablement, avec un autre logiciel : « *Mettez la justice prédictive de votre côté* [...] *Estimez le taux de succès d'une*

<sup>101</sup> J.-P. BUYLE et A. VAN DEN BRANDEN, *Op. cit.*, p. 292. Voy. la note en bas de page.

<sup>102</sup> <https://legalworld.wolterskluwer.be/>, consulté en mars 2020. Nos italiques.

<sup>103</sup> <https://www.stradalex.com/partners/>, consulté en mars 2020. Nos italiques.

<sup>104</sup> <http://www.wolterskluwerfrance.fr/>, consulté en mars 2020. Nos italiques.

<sup>105</sup> <https://www.doctrine.fr/fonctionnalites/>, consulté en mars 2020. Nos italiques.

*action contentieuse*. Puis transmettez cette information clé. *Optimisez votre stratégie en fonction des juridictions* »<sup>106</sup>. Entre discours mélioriste à la limite du transhumanisme et sémantique de la magie ou de la divination, à l’instar du Dr. FRANKENSTEIN, la rhétorique des éditeurs et des développeurs de solutions “intelligentes” à destination des avocats semble hésiter entre la magie de PARACELSE et la science moderne.

### III. L’IA À L’ÉPREUVE DES DROITS HUMAINS FONDAMENTAUX

Nous nous demandons à présent si l’IA est – dans la sphère du droit et le giron du travail – un moteur de progrès social et technique et ce, tant pour les professionnels que les justiciables. Doit-on s’attendre à une amélioration des conditions de travail grâce aux outils d’automation qui pourraient permettre aux travailleurs (secrétaires, parajuristes, juristes et avocats) de se focaliser sur les tâches plus intellectuelles ou plus techniques et d’être plus efficaces afin de gagner du temps et de l’argent ? Plus fondamentalement encore, existe-t-il un risque possible ou avéré selon lequel l’implémentation de tels logiciels dans la sphère du droit constitue un danger pour certains droits humains fondamentaux et si tel est le cas, quels sont ces dangers et quels sont les droits fondamentaux potentiellement menacés ?

#### §1. Rappels de définitions (IA, SEL et ML)

Nous avons eu l’occasion *supra* de séparer le bon grain de l’ivraie et d’exclure du champ de recherche les aspects de l’IA qui sont moins pertinents pour notre propos (par exemple la robotique et la *general* ou *strong AI*) et, par là même, de poser les jalons de ce qui constitue le nœud de ce mémoire : les questions éthiques et de droit liées *i.a.*, au paradigme de programmation et d’apprentissage des IA (en particulier le ML), au renforcement de biais cognitifs, à la discrimination et la question de la neutralité des algorithmes, aux principes de transparence et de maîtrise par l’utilisateur de SEL implémentant ML et IA, au risque de cristallisation de la jurisprudence selon une logique reproductive, de profilage du juge et de *forum shopping*, ou encore le problème que pose le caractère indéterminé et évolutif des concepts juridiques.

Nous avons par ailleurs dégagé la définition suivante pour caractériser ce qu’il y lieu d’entendre par IA : « [AI] systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving

<sup>106</sup> <https://predictice.com/>, consulté en mars 2020. Nos italiques.

their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take *to achieve the given goal*. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions »<sup>107</sup>.

Nous avons enfin insisté sur le fait que, dans le domaine judiciaire, le *natural language processing* (NLP), le *machine learning* (ML) et l'analyse automatisée de la jurisprudence à l'aide du *data mining* (DM) qui permettent *in fine* la prédictive judiciaire sont véritablement ce qui constitue le *nec plus ultra* des logiciels développés actuellement ; il s'agit d'IA de type *weak* incarnée dans des SE de nouvelle génération qui sont capables de reproduire en partie les mécanismes cognitifs rationnels d'un expert humain, d'extraire des tendances complexes d'une base de données et d'apprendre de façon autonome. On peut encore dire, avec la classification de S. RUSSELL et P. NORVIG, qu'il s'agit d'IA qui sont *goal-oriented, ideal* et *reasoning-based*.

Avant de confronter l'IA aux droits humains fondamentaux (ci-après DHF), nous commençons d'abord par fixer le cadre, et définir le concept pour mieux en relever la complexité ainsi que le caractère à la fois spécieux et relatif.

## **§2. Des fondations à la fixation du récit des DHF à l'époque moderne**<sup>108</sup>

L'origine la plus évidente de la notion de droits de l'homme est sans conteste La *Déclaration des droits de l'homme et du citoyen* de 1789 et le *Bill of Rights* anglais de 1689. On trouve cependant des traces primitives de ce concept dans les travaux de philosophes tels que Thomas D'AQUIN, SUAREZ, GROTIUS, VON PUFENDORF, LOCKE, ou encore KANT<sup>109</sup> notamment au travers de la notion, plus ancienne, de « droits naturels » qui prend racine, à l'instar du concept d'IA, dans l'Antiquité. Cela étant, la philosophie des Lumières et la Révolution française qui culmine avec la *Déclaration des droits de l'homme et du citoyen*

<sup>107</sup> AI HLEG, *Ibid.*, p. 6. Nos italiques.

<sup>108</sup> Les § 2 et 3 de cette section sont repris en partie d'un travail précédant intitulé « Le formalisme juridique à l'aune des droits de l'homme », disponible au lien <https://www.academia.edu/>.

<sup>109</sup> J. NICKEL, « Human Rights », The Stanford Encyclopedia of Philosophy 2017, E.N. ZALTA (ed.) disponible au lien <https://plato.stanford.edu/>, consulté en avril 2020 ; Voy. J. GRIFFIN, *On Human Rights*, Oxford, Oxford University Press, 2008, pp. 9-28.

constituent le véritable point d'orgue historique de l'avènement des droits humains fondamentaux en Europe. Et pourtant, la Révolution et la création *ex nihilo* de la notion de droits de l'homme est immédiatement suivie par une période de terreur<sup>110</sup> : comme nombre d'entreprises humaines et de grandes idées louables à l'origine, certains effets pervers ne tardent pas à apparaître ; l'enfer est pavé de bonnes intentions. Ainsi, une critique récente des droits de l'homme, dans la mouvance du relativisme culturel, met en évidence l'ethnocentrisme des DHF qui prennent racine dans un terreau hellénistique, judéo-chrétien et occidental<sup>111</sup>. Ainsi a-t-on pu observer, au cours des XIX, XX et XXI<sup>ème</sup> siècles, donc depuis leurs débuts, une certaine instrumentalisation des droits de l'homme à des fins politiques, stratégiques, voire militaires<sup>112</sup>.

### **§3. Droits humains en question : relativisme et droit-de-l'hommisme**

À partir du XIX<sup>ème</sup> siècle, par l'entremise de principes hérités de la philosophie libérale des Lumières, du droit naturel et des révolutions est donc reconnue l'existence de DHF. Songeons à la Déclaration universelle des droits de l'homme de 1948 et à la Convention de sauvegarde des droits de l'homme et des libertés fondamentales de 1950 dont le respect est assuré par la Cour européenne des droits de l'homme, au Pacte international relatif aux droits civils et politiques et à celui relatif aux droits économiques, sociaux et culturels de 1966 ou encore la Charte des droits fondamentaux de l'UE adoptée en 2000. On parle aujourd'hui de droits de l'homme de première (droits civils et politiques), de deuxième (droits économiques et sociaux) et de troisième génération (*i.a.*, droits environnementaux, au développement, à la paix etc.) Une quatrième génération, quoique moins largement reconnue en doctrine, vient s'ajouter à cette liste ; il s'agit des droits des personnes dites "faibles", c'est-à-dire des personnes plus vulnérables, telles que les enfants, les personnes âgées, les personnes

<sup>110</sup> Voy. D. EDELSTEIN, *The Terror of Natural Right: Republicanism, the Cult of Nature and the French Revolution*, Chicago, The University of Chicago Press, 2009 et M. GAUCHET, *La Révolution des droits de l'homme*, Paris, Gallimard, 1989.

<sup>111</sup> Au XIX<sup>ème</sup> siècle, Karl Marx dénonce déjà les effets pervers des droits de l'homme et en fait la critique. Voy. aussi les notions d'humanitarisme et droit-de-l'hommisme.

<sup>112</sup> J. LACROIX et J.-Y. PRANCHERE, *Le Procès des droits de l'homme. Généalogie du scepticisme démocratique*, Seuil, 2016 ; J.-F. MATTEI, *L'humanitaire à l'épreuve de l'éthique*, Les Liens qui libèrent, 2014 ; A. TAILLEFER, *Droit de l'hommisme : une névrose religieuse*, Godefroy de Bouillon, 2013 ; J. BRICMONT, *Impérialisme humanitaire : droits de l'homme, droit d'ingérence, droit du plus fort ?*, Aden 2009.

handicapées, les étrangers (en particulier les réfugiés et les apatrides)<sup>113</sup>. Il convient aussi de relever que plusieurs organisations internationales, telles que l'UE, le Conseil de l'Europe, l'ONU ainsi que nombre d'organisations non-gouvernementales participent à cet effort visant au renforcement de la coopération internationale eu égard le respect et la sauvegarde des droits fondamentaux et ce, tant au sein des États membres que dans le reste du monde<sup>114</sup>.

Depuis quelques années, cependant, la tendance est à une certaine relativisation du caractère supposément universel, inaliénable, voire absolu, des DHF<sup>115</sup>. En effet, ceux-ci s'avèrent plus ou moins bien protégés selon les possibilités de recours disponibles dans l'ordre juridique considéré. Bien plus, la jurisprudence de la Cour européenne des droits de l'homme, censée interpréter de manière plus extensive les droits fondamentaux que ne le ferait la Cour de justice de l'Union européenne révèle, dans des affaires qui heurtent la morale, que les droits de l'homme ne peuvent empiéter sur le bien-être économique des États<sup>116</sup>. En effet, dans des sociétés modernes, industrialisées et développées, bien que les DHF des individus et les impératifs économiques semblent converger à l'horizon dans la mesure où les droits de l'homme sont des conditions nécessaires au déploiement efficace de l'activité économique des États, « l'optique est ainsi faite que deux routes qui se rejoignent à l'horizon peuvent, en réalité, être parfaitement parallèles et ne jamais se croiser en demeurant à distance l'une de l'autre »<sup>117</sup>. Dans la mise en balance des intérêts en présence, les impératifs tels que la sûreté et la santé publique, le bien-être économique du pays et, plus généralement, la souveraineté nationale l'emportent souvent sur les droits fondamentaux des individus. De plus, ces droits qui, par nature, devraient être universels, voire absolus, se voient relativisés et conditionnés

<sup>113</sup> Amnesty International, *A Blue Print for Despair: Human Rights Impact of the EU-Turkey Deal*, London, Amnesty International Ltd, 2017, disponible au lien <https://www.amnesty.org/>, consulté en 2017 ; European Migration Network, *2015 Annual Report on Asylum and Migration Policy in Belgium*, Belgian National Contact Point, p. 11, disponible au lien <https://emnbelgium.be/>, consulté en 2017.

<sup>114</sup> Voy. *i.a.*, Preamble to the Charter of the United Nations, 1945 et les articles 2 et 3 T.U.E.

<sup>115</sup> Voy. P. SLOTTE, et M. HALME-TUOMISAARI, *Revisiting the Origins of Human Rights*, Cambridge, Cambridge University Press, 2015.

<sup>116</sup> Cour eur. D.H., arrêt *Yoh Ekale Mwanje c. Belgique*, 20/12/2011, req. n°10486/10 ; Cour eur. D.H., arrêt *Josef. c. Belgique*, 27/02/2014, req. n°70055/10 cité dans P. MARTENS, « Les droits de l'homme vont-ils mourir au petit feu des contraintes budgétaires et des valeurs économiques », A. HOC, S. WATTIER, G. WILLEMS (eds.), *Human Rights as a Basis for Reevaluating and Reconstructing the Law*, Bruxelles, Bruylant, 2016.

<sup>117</sup> S. VAN DROOGHENBROECK, « La Convention européenne des droits de l'homme et la matière économique », L. BOY, J.-B. RACINE, F. SIIRIAINEN, *Droit économique et droits de l'homme*, Bruxelles, Larcier, 2009.

par les protections (les possibilités de recours, et, le cas échéant, de réparation) mis en œuvre par les États desquels les individus qui entendent faire valoir ces droits ressortent. Dit plus simplement, la mise en œuvre d'un recours efficace par un individu qui subit un préjudice du fait de la violation, imputable à un État, des droits de l'homme, dépend fortement de l'État, et par voie de conséquence, de l'ordre juridique auquel l'individu appartient.

### § 3 – IA vs. DHF

Maintenant que nous avons pris acte du caractère à la fois spécieux et relatif des DHF après les avoir définis et en avoir retracé les origines, tournons-nous vers la question que nous nous posions au début de cette section : Existe-t-il un risque possible ou avéré selon lequel l'implémentation massive de logiciels combinant IA, NLP, ML, DM et *predictive analytics* dans la sphère du droit pourrait constituer un danger pour certaines valeurs fondamentales et si tel est bien le cas ; quels sont-ils ?

À titre préliminaire, relevons que la compatibilité des solutions que ne font miroiter les capacités de l'IA appliquée au domaine du droit avec les droits individuels consacrés, notamment, par la Convention européenne des droits de l'homme (ci-après CEDH) pose question. Plus largement, le déploiement irréfléchi de l'IA pourrait entrer en conflit avec le principe, plus large, de l'autonomie de la personne et du système judiciaire lui-même.

En effet, ce qui semble poser le plus problème, du point de vue de la pratique du métier d'avocat, est ce que l'on pourrait appeler, dans le sillage de C. LAZARO et A. ROUVROY, la « modélisation algorithmique de la réalité » ainsi que les « modes de gouvernance basés sur l'association entre logique d'anticipation juridique et algorithmique »<sup>118</sup>.

Avec les nouveaux systèmes d'analyse prédictive appliqués au droit et leurs capacités quasi oraculaires, nous entrons, d'après LAZARO, dans un « nouvel imaginaire sociotechnique » qui contribue « à fabriquer une certaine vision de la vie et du futur »<sup>119</sup>. Or, « [t]oute recherche destinée à comprendre comment le droit et l'intelligence artificielle s'associent pour gouverner doit s'intéresser au sort réservé aux formes de vie – que celles-ci soient humaines

<sup>118</sup> C. LAZARO, « L'ici et maintenant des avenir incertains : le droit, l'intelligence artificielle et les technologies de prédiction », *Des véhicules autonomes à l'intelligence artificielle*, Bruxelles, Éditions Larcier, 2020, p. 248 ; A. ROUVROY, « Le régime de vérité numérique. De la gouvernementalité algorithmique à un nouvel État de droit », *Socio*, vol. 4, 2015, disponible au lien <http://socio.revues.org/1251>, consulté en mai 2020.

<sup>119</sup> C. LAZARO, *Ibid.*, p. 225

ou non humaines – qui ne s’adaptent pas, ne s’alignent pas ou résistent aux modes de vie néo-libéraux. En d’autres termes, il s’agit d’éviter que l’assemblage juridico-algorithmique contemporain ne finisse par protéger et favoriser uniquement des formes de vie en phase avec le modèle dominant »<sup>120</sup>. Nous reviendrons sur ces questions *infra* dans le paragraphe consacré à la gouvernance algorithmique et dans la conclusions de ce mémoire. Par ailleurs et en ce qui concerne le justiciable, l’IA pourrait menacer les droits suivants :

| DHF                                                                                                                                    | Corollaires                                                                                                                                              | Risques                                                                                                                                                                                                                                                                                                               | Solutions <sup>121</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>procès équitable</b><br><br><b>accès au juge</b><br><br>art. 6 § 1 CEDH                                                             | droit à un tribunal indépendant et impartial établi par la loi<br><br>égalité des armes dans les procédures judiciaires<br><br>respect du contradictoire | Lorsque les outils d’IA sont utilisés pour trancher un litige ou en tant qu’outils d’aide à la décision judiciaire ou d’orientation du justiciable, ils peuvent potentiellement porter atteinte aux garanties du droit d’accès au juge et du procès équitable.                                                        | Les approches dites <i>ethical</i> ou <i>human rights by design</i> sont à privilégier.<br><br>Elles consistent à intégrer dès la phase de conception et d’apprentissage des règles interdisant de porter atteinte directement ou indirectement aux valeurs fondamentales protégées par les Conventions.                                                                                                                                                                                                                      |
| <b>non-discrimination</b><br><br>art. 14 CEDH<br><br>dir. 2006/54/CE                                                                   | égalité des chances<br><br>égalité de traitement                                                                                                         | Usages ou analyses prédictifs déterministes et renforcement de biais cognitifs.<br><br>Les préjugés des concepteurs et des programmeurs, des logiciels qui exécutent un algorithme de ML peuvent se retrouver dans le code de l’IA qui reproduira, potentiellement en les renforçant, les biais cognitifs de ceux-ci. | Une vigilance particulière doit être accordée tant dans la phase d’élaboration que d’utilisation, en particulier lorsque ces traitements se fondent de manière directe ou indirecte sur des données “sensibles”<br><br>(e.g., origine prétendument raciale ou ethnique, conditions socio-économiques, opinions politiques, convictions religieuses ou philosophiques, appartenance syndicale, données génétiques, données biométriques, données concernant la santé ou concernant la vie sexuelle ou l’orientation sexuelle). |
| <b>respect de la vie privée</b><br><br><b>protection des données à caractère personnel</b><br><br>art. 8 CEDH<br><br>art. 5, 6, 9 RGPD | traitement des données juridictionnelles et judiciaires<br><br>respect de la vie privée et familiale<br><br>droit à l’effacement / droit à l’oubli       | Risque de violation du droit au respect à la vie privée qui peut être contrecarré par l’anonymisation des données.<br><br>En particulier, risque de tomber dans une gouvernance algorithmique.                                                                                                                        | Les concepteurs de modèles à partir du ML doivent pouvoir recueillir de manière large l’expertise des professionnels de la justice concernés (juges, procureurs, avocats, etc.) et des chercheurs / enseignants en droit et en sciences sociales (économistes, sociologues, philosophes par exemple). Les modèles et les algorithmes créés doivent être stockés et exécutés dans des environnements sécurisés, garantissant l’intégrité et l’intangibilité du système.                                                        |

<sup>120</sup> C. LAZARO, *Ibid.*, p. 248-249.

<sup>121</sup> CEPEJ, *Charte éthique*, *Ibid*, pp. 6-10. Nos italiques et annotations.

Le recours à l'IA peut donc porter atteinte aux valeurs sur lesquelles l'UE est fondée et entraîner des violations des DHF, tels que la liberté d'expression et de réunion à travers le filtrage et la suppression de contenu, la dignité humaine, l'absence de discrimination fondée sur le sexe, l'origine raciale ou ethnique, la religion ou les convictions, le handicap, l'âge ou l'orientation sexuelle, selon le cas, la protection des données à caractère personnel, le respect de la vie privée ou le droit à un recours juridictionnel effectif et à un procès équitable, ainsi que la protection des consommateurs pour ne citer que les plus évidents. Il s'agit donc, pour reprendre la classification des DHF telle que nous l'avons établie *supra* essentiellement de DHF de première (droits civils et politiques) et de deuxième génération (droits économiques et sociaux). La protection des personnes vulnérables, les DHF de quatrième génération, et, plus largement, ce que l'on peut appeler, avec C. LAZARO, les « formes de vie – [...] humaines ou non humaines – qui ne s'adaptent pas, ne s'alignent pas ou résistent aux modes de vie néo-libéraux », ne doivent cependant pas être oubliées – d'autant que, pour reprendre un exemple tout à fait symptomatique qui concerne le principe de non-discrimination et le renforcement de biais cognitifs des concepteurs, « [a]ux États-Unis, l'ONG PROPUBLICA a révélé les effets discriminatoires de l'algorithme utilisé dans le logiciel COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), dont l'ambition est d'évaluer le risque de récidive lorsque le juge doit déterminer la peine d'un individu »<sup>122</sup>.

La dimension éthique des logiciels ne doit dès lors pas être perdue de vue et une attention toute particulière se doit d'être accordée à celle-ci: à ce titre, la prudence et les approches dites *ethical* ou *human rights by design*, sur lesquelles nous reviendrons *infra*, sont, à notre sens, à privilégier. Singulièrement, les notions axiologiquement neutres de prudence, de protection de la dignité et de l'autonomie de la justice et de la personne conçue au sens large nous semblent fondamentales dans la mesure où l'IA emporte avec elle un danger important pour le droit, à savoir le risque de cristallisation de la justice par la logique reproductive des algorithmes, le risque de substitution du calcul probabiliste à la délibération du juge – ou de l'avocat – sur le cas d'espèce.

<sup>122</sup> CEPEJ, « Les enjeux spécifiques à la justice pénale : prévention des infractions, du risque de récidive et évaluation de la dangerosité », *Charte éthique*, Annexe 1, *Ibid.*, p. 41. Cf. K. YEUNG, « Why Worry about Decision-Making by Machine? », K. YEUNG, M. LODGE (eds.), *Algorithmic regulation*, Oxford, Oxford University Press, 2019, pp. 21-48 ; N. CRIADO, J.M. SUCH, « Digital Discrimination », K. YEUNG, M. LODGE, *Op. cit.*, pp. 82-97 ; L.A. BYGRAVE, « Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making », K. YEUNG, M. LODGE, *Op. cit.*, pp. 248-262.

On le voit bien, avec l'introduction de l'IA dans les cabinets et les prétoires, ce sont les principes d'égalité des armes dans les procédures judiciaires et d'autonomie des personnes et des systèmes judiciaires qui se trouvent menacés.

Pourtant, l'introduction de l'IA dans la sphère du droit peut aussi, par contraste, constituer un moteur de progrès social et technique tant pour les professionnels du droit que pour les justiciables : « Nul ne doute que certaines applications d'IA en cours de développement ou d'expérimentation, comme par exemple celles visant à *améliorer la recherche juridique*, peuvent être d'une très grande utilité et rendre le travail judiciaire à la fois *plus rapide et plus efficace*. [Il y a lieu de prôner une] utilisation de l'IA qui soit, d'une part, au service des professionnels de la justice et en phase avec leurs besoins et d'autre part, respectueuse des droits individuels garantis par la CEDH et des normes du Conseil de l'Europe, notamment en matière pénale. *Loin d'être un simple instrument d'amélioration de l'efficacité des systèmes judiciaires, l'IA devrait renforcer et ne pas diminuer les garanties de l'État de droit tout comme la qualité du service public de la justice* »<sup>123</sup>.

D'aucuns pensent par ailleurs que les *legaltechs* pourraient réaliser entre 30 et 50 % des tâches effectuées aujourd'hui par les avocats juniors<sup>124</sup> ce qui implique sans nul doute une amélioration des conditions de travail dans la mesure où ces technologies permettraient aux avocats stagiaires et aux associés de se focaliser davantage sur les tâches plus intellectuelles ou plus techniques ; de pouvoir faire plus et de façon plus efficace pour reprendre la rhétorique commerciale des développeurs.

Plus largement, la transition numérique et la robotisation du droit induisent immanquablement des répercussions profondes sur la structure des cabinets d'avocats, la détermination des honoraires et la responsabilité professionnelle<sup>125</sup> ainsi qu'une dévaluation des *hardskills* (*i.e.*, dans ce contexte, la connaissance du droit au sens large) et de la pratique "généraliste" du métier d'avocat et, corrélativement, une importance accrue des *softskills* et de la spécialisation dans l'une ou l'autre branche du droit des avocats solistes ou travaillant au sein de petites ou moyennes structures.

<sup>123</sup> CEPEJ, « Introduction », *Charte éthique*, Annexe 1, *Ibid.*, p. 13. Nos italiques et annotations.

<sup>124</sup> Rapport rédigé par The Boston Consulting Group et la Bucurius Law School cité dans J.-P. BUYLE et A. VAN DEN BRANDEN, *Ibid.*, p. 307. Cf. J.D. LOHR, W.J. MAXWELL, P. WATTS, « Legal Practitioners' Approach to Regulating AI Risks », K. YEUNG, M. LODGE, *Ibid.*, pp. 224-247.

<sup>125</sup> Voy. J.-P. BUYLE et A. VAN DEN BRANDEN, *Ibid.*, pp. 309-313.

## SECONDE PARTIE

### MISE EN ŒUVRE DE L'IA DANS LE DROIT

« *The only way to make sense out of change is to plunge into it, move with it, and join the dance* »<sup>126</sup>.

Après avoir retracé l'histoire du concept d'IA (section I), défini l'IA dans ses rapports avec le droit en réfléchissant sur les enjeux, les avantages, les inconvénients et les limites de cette technologie émergente, délimité le cadre juridique (section II, § 1 à 3), investigué par ailleurs la question de savoir qui produit et vend des logiciels implémentant de l'IA (section II, § 4) et envisagé les risques liés aux droits humains fondamentaux que comporte son usage dans la sphère du droit (section III, § 3) ; nous envisageons, dans la deuxième partie de ce mémoire et dans une perspective comparative, éthique et critique, les questions suivantes :

À quoi doivent être attentifs ceux qui produisent des SEL de nouvelle génération ? Existe-t-il des principes éthiques ou de droit pris en compte lors de la programmation de ces outils informatiques ? Qu'entend-on, précisément, par approche "*ethical / human rights by design*" ? Les diverses inquiétudes d'ordre éthique habituellement associées à l'IA, telle que celles qui sont associées aux *Big Data*, au renforcement de biais cognitif, aux *Black-Box AI*, à la question de l'*explainability* ou encore à la gouvernance algorithmique sont-elles prises en compte par les développeurs ? Quid des *hard cases*, des problèmes de droit, des dilemmes éthiques apparemment indissolubles ? Plus largement, l'analyse et l'interprétation du droit et de la jurisprudence peuvent-elle être formalisées dans un langage de programmation et, partant, purement et simplement réduites dans un système computationnel ?

Et enfin, pour finir sur une question éminemment philosophique, s'il en est : dans le contexte du droit continental romano-civiliste et par opposition au droit dit de *common law* qui se fonde sur la jurisprudence antérieure (*stare decisis* ou règle du précédent), l'indétermination du droit et l'ambiguïté syntaxique et sémantique des textes juridiques ainsi que les limites intrinsèques à l'herméneutique juridique ne mettent-elles pas nécessairement en échec toute tentative de réductionnisme positiviste appliqué au droit ? A contrario, si l'on introduit, en droit continental, des outils modernes et pragmatiques fonctionnant essentiellement par la comparaison avec le précédent et par l'analyse statistique et prédictif sur les données des cours et tribunaux, ne risque-t-on pas de faire perdre au droit continental ce qui constitue sa spécificité (troisième partie) ?

<sup>126</sup> A.W. WATTS, *The Wisdom of Insecurity: A Message for an Age of Anxiety*, New York, Vintage, 2011, p. 43.

## I. PERSPECTIVE COMPARATIVE, ÉTHIQUE ET CRITIQUE

### §1 – Usages dans les pays de *common law*

Nous avons déjà eu l’occasion de dire *supra* que l’utilisation de l’IA dans le champ judiciaire est largement vulgarisée aux États-Unis. Pour reprendre une allégorie devenue lieu commun, on peut dire que “le pays de l’oncle Sam” a investi les outils liés à l’IA de manière tout à fait décomplexée et ce, tant en matière civile que pénale. À ce titre, nous avons discuté précédemment, dans la première partie de ce mémoire, des différentes entreprises ayant pignon sur rue dans ce secteur : en particulier des algorithmes utilisés par COMPAS, RAVEL LAW (LEXIS NEXIS) ou encore la plateforme ROSS (IBM) qui peut, accessoirement, se comporter en agent conversationnel *i.e.*, en *chatbots / lawbots*. En effet, « [a]ux États-Unis, des « robots avocats » sont déjà à l’œuvre et semblent converser en langage naturel avec les humains »<sup>127</sup> les rendant censément indistinguables d’agents humains. Le projet d’appliquer l’IA au droit existe de longue date aux États-Unis et au moins depuis la parution de l’article de BUCHANAN et HEADRICK « Some Speculation about Artificial Intelligence and Legal Reasoning » paru dans la *Stanford Law Review* dans les années 1970. Si à l’époque de la parution de cet article, l’application de l’IA au domaine du droit relevait plutôt de la spéculation et de la déclaration d’intention, force est de constater qu’aujourd’hui, une cinquantaine d’années plus tard, le projet est devenu tangible. Outre-Atlantique, il est maintenant admis que l’IA est capable, dans le domaine du droit, de réaliser des tâches complexes, parfois mieux que les experts humains : « [r]ecent developments in artificial intelligence – in particular, natural language processing and machine learning – have challenged traditional conceptions of human expertise. Machines now perform increasingly complex tasks »<sup>128</sup>. Pour certains avocats américains, « ROSS’s legal skills [are] indecipherable from those of a young associate »<sup>129</sup>. Les spécialistes se posent dès lors très

<sup>127</sup> CEPEJ, *Charte éthique*, *Ibid.*, p. 12.

<sup>128</sup> B. ALARIE, A. NIBLETT, H. YOON, « How Artificial Intelligence Will Affect the Practice of Law », *University of Toronto Law Journal*, N°68, supplement 1, 2018, p. 2.

<sup>129</sup> J.L. JACOBOWITZ, J. ORTIZ, « Happy Birthday Siri! Dialing in Legal Ethics for Artificial Intelligence, Smart Phones, and Real Time Lawyers », *Texas A&M University, Journal of Property Law, University of Miami Legal Studies Research Paper* N°18-2, 2018, p. 7 disponible au lien <https://ssrn.com/abstract=3097985>, consulté en avril 2020.

Voy. G. PASSARELLA, « Salazar Jackson Enters World of AI With ROSS Intelligence », *Daily Business Review*, 2016, disponible au lien <http://www.law.com/dailybusinessreview/>, consulté en avril 2020.

sérieusement la question: est-on en face d'une évolution, d'une disruption ou d'une véritable révolution dans la pratique du droit ? En effet, « [i]n the longer term, it is unclear exactly how dramatic the impact of artificially intelligent tools will be. Such predictions are difficult for us to make. We are, after all, only human »<sup>130</sup>.

## §2 – Étude de cas : États-Unis, Japon, Europe

### a. *Silicon Valley : Eldorado de l'IA*

Au Nord de la Californie et, plus précisément, dans la Silicon Valley et la baie de San Francisco, l'IA fait déjà partie de la toile de fond. Au pays d'Uber et de Tinder, il n'est pas rare de croiser des voitures autonomes qui se conduisent toutes seules, un peu comme, pour reprendre le mot d'ARISTOTE, les statues de DÉDALE et les trépieds d'HÉPHAÏSTOS qui se rendaient seuls aux réunions des dieux... Depuis le mois de septembre 2018, par ailleurs, une loi encadre l'usage des *chatbots* : « California enacted a statute making it illegal for an online “bot” to interact with consumers without first disclosing it is non-human »<sup>131</sup>. Des voitures qui se conduisent toutes seules, des robots conversationnels qui doivent se déclarer comme tels à leurs interlocuteurs humains ou encore le rêve d'Elon MUSK, concrétisé par le projet SPACEX, de conquérir Mars pour y implanter les premières “off-world colonies”<sup>132</sup> ; peut-être ne sommes-nous plus très loin du monde imaginé par Philip K. DICK en 1966 dans sa nouvelle de SF intitulée *Do Androids Dream of Electric Sheep?* qui se déroule justement à San Francisco – berceau par excellence de la révolution numérique. Les cabinets d'avocats (*law practices*) de Palo Alto qui est situé non loin de Stanford se spécialisent et pratiquent déjà un nouveau type de contentieux directement lié à ces technologies émergentes que sont l'IA et les systèmes autonomes : « As specialists in a variety of legal areas, the members of [the Artificial Intelligence and Automated Systems] Practice Group are engaged with cutting edge technology companies in litigation and transactional matters spanning a wide spectrum

<sup>130</sup> B. ALARIE, A. NIBLETT, H. YOON, *Ibid*, p. 15. Voy. aussi D. REMUS, F.S. LEVY, « Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law », 2016 disponible au lien <https://ssrn.com/abstract=2701092>, consulté en février 2020.

<sup>131</sup> CASEY, B., LEMLEY, M., A., « You Might Be a Robot », Cornell Law Review, 2019, disponible au lien <https://ssrn.com/abstract=3327602>, p. 3, consulté en février 2020. Cf. le California Senate Bill N°1001 qui définit le *bot* de la façon suivante : « an automated online account where all or substantially all of the actions or posts of that account are not the result of a person ».

<sup>132</sup> E. MUSK, « The Case for Mars » disponible au lien <https://www.spacex.com/about/>, consulté en avril 2020.

of machine and deep learning, expert systems, neuromorphic computing, relational reasoning and analytic technologies. Our lawyers with technical backgrounds and advanced degrees can provide the necessary insight to develop and defend key growth areas including *dedicated integrated circuit design, NLP, chatbots, augmented/virtual reality, computer vision, robotics, autonomous vehicles, and other new emerging tech forms* »<sup>133</sup>. Il n'est pas impossible que l'on voie des cabinets européens et belges se spécialiser dans cette nouvelle matière au fur et à mesure de la dissémination de l'IA et de ses diverses formes et manifestations sur le Vieux Continent. Il est en outre clair que ce genre de cabinets, qui se targuent d'avoir pour clients des sociétés à la pointe des dernières technologies, doivent se réserver pour elles-mêmes les meilleurs outils de comparaison analytique et casuistique, de recherche jurisprudentielle avancée – a fortiori dans un système juridique de *common law* qui applique la règle du précédent – et de prédictive judiciaire telle que les logiciels développés par IBM ou LEXIS NEXIS. En effet, dans un environnement manifestement technophile, on ne voit pas comment des arguments de droit basés sur de l'IA, du ML ou encore de la *statistical probability analysis*, jumulés à de l'intelligence humaine et d'excellents *softskills* ne pourraient pas faire mouche<sup>134</sup>.

#### b. Japon : pays des robots

De façon générale, on peut dire que le droit japonais moderne se trouve à la croisée des chemins entre le droit romano-civiliste continental d'une part, et, d'autre part, le droit anglo-américain dit de *common law* avec néanmoins de fortes caractéristiques locales et chinoises – en particulier, l'influence du confucianisme et du bouddhisme<sup>135</sup> : au pays du soleil levant, « [j]udicial precedent [*stare decisis*] provides as a guideline on how these laws [Civil Code, Commercial Code, Code of Civil Procedure, etc.] should be interpreted in practice. Although not legally binding, judges do take precedent, especially a Supreme Court decision, into

<sup>133</sup> Voy. <https://www.gibsondunn.com/practice/>, consulté en mars 2020. Nos italiques.

<sup>134</sup> Voy. D. REMUS, F.S. LEVY, « Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law », 2016, disponible au lien <https://ssrn.com/abstract=2701092>, consulté en février 2020.

A contrario, J.P. DAVIS, « Law Without Mind: AI, Ethics, and Jurisprudence », Univ. of San Francisco, Law Research Paper, N°2018-05, 2018, disponible au lien <https://ssrn.com/abstract=3187513>, consulté en mars 2020.

<sup>135</sup> M. LEVIN, « Continuities of Legal Consciousness: Professor J. HALEY's Writings On Twelve Hundred Years of Japanese Legal History », Washington University Global Studies Law Review, vol. 8, 2009, pp. 317-332.

serious consideration, thus making understanding of precedent essential to practice »<sup>136</sup>. Au pays d'EVANGELION, d'ASIMO et de PARO, comme l'atteste la culture populaire nippone, les robots occupent depuis longtemps une place toute particulière<sup>137</sup>. En effet, dans ce pays au moins aussi technophile que la Californie du Nord, l'insertion durable des robots dans la société est une nécessité qui s'impose face au problème systémique du vieillissement de la population. Ainsi, depuis sa création en 2014, le Comité d'éthique de la Japanese Society for Artificial Intelligence (ci-après JSAI) explore les relations entre les recherches en matière d'IA et la société japonaise. Le comité est foncièrement multidisciplinaire. Composé de chercheurs en IA, il intègre également des observateurs issus de la société civile, comme des écrivains de SF, des journalistes et des spécialistes en éthique. Le comité mène des débats au sein de différentes agences gouvernementales au Japon et d'instances internationales telles que le G7 ou l'OCDE. Ces débats ont abouti fin 2017 à l'adoption d'une charte de déontologie de la JSAI de laquelle ressortent quatre principes particulièrement pertinents pour notre propos :

- « 1. **Principe éthique** : Les développeurs doivent respecter la *dignité humaine* et *l'autonomie individuelle* dans la recherche et développement des systèmes d'IA ; par exemple, ils doivent prendre les mesures nécessaires pour ne pas causer de discrimination résultant de préjugés qui auraient été inclus dans les données d'apprentissage des systèmes d'IA ;
2. **Principe de transparence** : Les développeurs doivent prêter attention à la vérifiabilité des entrées/sorties des systèmes d'IA et à l'explicabilité de leurs jugements ;
3. **Principe de responsabilité** : Les développeurs doivent s'efforcer de s'acquitter de leur responsabilité vis-à-vis des parties prenantes, y compris des utilisateurs des systèmes d'IA ;
4. **Principe de respect de l'intimité et de la vie privée** : Les développeurs doivent prendre en compte le fait que les systèmes d'intelligence artificielle ne porteront pas atteinte à la vie privée des utilisateurs ou des tiers »<sup>138</sup>.

<sup>136</sup> T. MORISHITA, « The Financial Industry of Japan: An entanglement of statutory law, case law, and guidelines of the financial authorities make up the legal order of financial law in Japan », disponible au lien <http://www2.osipp.osaka-u.ac.jp/>, consulté en mars 2020.

<sup>137</sup> Voy. P. DUMOUCHEL, L. DAMIANO, *Vivre avec les robots : essai sur l'empathie artificielle*, Paris, Seuil, coll. La couleur des idées, 2016.

<sup>138</sup> A. BENSOUSSAN, J. BENSOUSSAN, *IA, robots et droit*, Bruxelles, Bruylant 2019, pp. 122-123.

Voy aussi, l'annexe 10 de cet ouvrage qui reprend, en français, la Charte de déontologie de la Japanese Society for AI. Nos italiques.

### c. Comparaison avec l'Europe

Les quatre grands principes de la Charte déontologique du comité éthique de la JSAI ne sont pas sans rappeler ceux que l'on trouve dans la Charte éthique européenne d'utilisation de l'IA dans les systèmes judiciaires et leur environnement dont nous avons discuté précédemment dans la première partie de ce mémoire. Pour rappel, les cinq principes de la Charte européenne sont les suivants :

- « 1. **Principe de respect des droits fondamentaux** : Assurer une conception et une mise en œuvre des outils et des services d'IA qui soient compatibles avec les droits fondamentaux ;
- 2. **Principe de non-discrimination** : Prévenir spécifiquement la création ou le renforcement de discriminations entre individus ou groupes d'individus ;
- 3. **Principe de qualité et sécurité** : En ce qui concerne le traitement des décisions juridictionnelles et des données judiciaires, utiliser des sources certifiées et des données intangibles avec des modèles conçus d'une manière multidisciplinaire, dans un environnement technologique sécurisé ;
- 4. **Principe de transparence, de neutralité et d'intégrité intellectuelle** : Rendre accessibles et compréhensibles les méthodologies de traitement des données, autoriser les audits externes ;
- 5. **Principe de maîtrise par l'utilisateur** : Bannir une approche prescriptive et permettre à l'utilisateur d'être un acteur éclairé et maître de ses choix »<sup>139</sup>.

Il est clair que les principes nippons et européens se recoupent : les préoccupations éthiques sont axées sur le respect de certaines valeurs fondamentales, en particulier du principe de non-discrimination et du droit au respect de la vie privée. Une attention particulière est également accordée, dans les deux Chartes, à la sécurité et à la transparence (vérifiabilité et explicabilité) ainsi qu'à la responsabilité des développeurs et à la maîtrise par l'utilisateur<sup>140</sup>. Il est important de relever que les DHF ne se retrouvent pas en tant que tels dans la Charte nippone : ce sont les principes d'autonomie individuelle et de dignité humaine qui prévalent. Ces principes éthiques et de droit axiologiquement neutres doivent servir, à notre sens, de jalons à suivre tandis que l'IA se développe ; ils doivent baliser à la fois son implémentation et la recherche dans le secteur sans pour autant entraver l'innovation. Il s'agit, il faut le reconnaître, d'un équilibre difficile à trouver.

<sup>139</sup> CEPEJ, *Charte éthique*, *Op. cit.*, p. 2.

<sup>140</sup> Les choses sont différentes en Chine : dans un pays où l'on applique, depuis 2018, le système de crédit social (社会信用体系) qui se base notamment sur des technologies d'IA telles que le *big-data analysis* et la reconnaissance faciale (*facial recognition*) et où, de surcroît, les DHF sont considérés comme des périls menaçant la suprématie du parti unique. Cf. le fameux « document n°9 » du Parti communiste.

### §3 – Enjeux éthiques, risques et impasses

#### a. *Data issues et cognitives biases*

Nous avons déjà eu l’occasion de discuter *supra* du problème lié au renforcement de biais cognitifs dans les algorithmes : en effet, « [c]ertains algorithmes d’IA peuvent, lorsqu’ils sont utilisés pour prédire la récurrence d’actes délictueux, présenter des biais de nature sexiste et raciale et fournissent des prédictions de la probabilité de récurrence différentes selon qu’il s’agit de femmes ou d’hommes ou de ressortissants nationaux ou d’étrangers »<sup>141</sup>. Les préjugés des concepteurs et des programmeurs, des logiciels qui exécutent un algorithme de ML peuvent se retrouver dans le code de l’IA qui reproduit, potentiellement en les renforçant, les biais cognitifs de ceux-ci. Une solution possible, proposée dans le Livre blanc de la Commission, est d’accorder une vigilance particulièrement importante, tant dans la phase d’élaboration que d’utilisation de l’algorithme, et, a fortiori, lorsque ces traitements se fondent de manière directe ou indirecte sur des données dites “sensibles” telles que l’origine raciale ou ethnique, les conditions socio-économiques, les convictions religieuses ou philosophiques, ou encore les données concernant la santé ou concernant la vie ou l’orientation sexuelle des intéressés<sup>142</sup>. Le but étant alors d’éviter tout risque de discrimination causé par une IA qui répète, en les renforçant, les aprioris de ses concepteurs du fait que l’IA a été “nourrie” par des données elles-mêmes biaisées, lacunaires ou encore insuffisantes. En principe, au plus il y a de données, au moins l’IA devrait être biaisée ce qui tend à plaider en faveur des politiques d’*open data*, voire, plus largement, pour une conception des données en tant que *res communis* – pour autant du moins que celles-ci soient anonymisées et que le respect de la vie privée et la protection des données à caractère personnel soient garantis.

#### b. *Black-box AI, explainability et transparence*

Les particularités qui caractérisent de nombreuses technologies de l’IA, notamment l’opacité, que l’on nomme également « effet de boîte noire », la complexité, l’imprévisibilité ainsi que le comportement autonome, peuvent rendre difficile la vérification de la conformité aux règles du droit de l’UE en vigueur qui sont destinées à protéger les DHF et peuvent dès lors

<sup>141</sup> S. TOLAN, M. MIRON, E. GOMEZ, C. CASTILLO, « Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia », Best Paper Award, International Conference on AI and Law, 2019, cité dans *Livre blanc sur l’intelligence artificielle*, Op. cit. p.13.

<sup>142</sup> Commission européenne, *Livre blanc*, Ibid., pp. 6-10.

entraver le contrôle de l'application de celles-ci<sup>143</sup>. Nous avons vu *supra* et en ce qui concerne les anciens SEL qu'au-delà d'une centaine de règles d'encodage, le raisonnement suivi par le système devient quasiment impossible à appréhender ce qui pose évidemment problème en termes de transparence<sup>144</sup>. Bien plus, les nouveaux systèmes d'apprentissage machine sont capables de procéder eux-mêmes à l'identification de modèles statistiques existant dans les données et de les mettre en correspondance avec des résultats spécifiques, ce qui leur permet, *in fine*, d'adapter leur raisonnement aux situations qui se présentent à eux. Avec le ML, l'IA détecte des corrélations et fait des liens dans les données qu'une intelligence humaine serait incapable de faire. On voit bien que le problème de la transparence et de l'opacité des raisonnements des IA se pose ici à plus forte raison encore. On comprend dès lors combien les principes de transparence, de neutralité et d'intégrité intellectuelle préconisés dans les Chartes éthiques européenne et nipponne s'avèrent cruciaux : vérifiabilité des *in* et *outputs*, explicabilité des jugements normatifs formés et audits externes sont les pièces maîtresses seules capables de garantir de telles exigences.

### c. Gare à la gouvernance algorithmique !

La gouvernance par les algorithmes n'est pas que le thème principal de certains films et séries de SF, tels que *Black Mirror*<sup>145</sup> ou la troisième saison de *Westworld*<sup>146</sup> : elle connaît d'ores et déjà nombre d'applications concrètes aux États-Unis et en Chine<sup>147</sup>. Le système de crédit social qui implémente des technologies telles que la *facial recognition*, le *big data analysis* ainsi qu'un réseau de plusieurs centaines de millions de caméras de surveillance surpuissantes en constituent, sans nul doute, l'exemple le plus criant<sup>148</sup>. Bien que « [l]'utilisation accrue

<sup>143</sup> Commission européenne, *Livre blanc*, *Op. cit.*, p. 14.

<sup>144</sup> CEPEJ, *Charte éthique*, *Ibid.*, p. 26.

<sup>145</sup> Voy. l'épisode 1 de la saison 3 de *Black Mirror* intitulé « Nosedive », sorti en 2016.

<sup>146</sup> Voy. la saison 3 de *Westworld*, sortie en 2020. L'ordinateur quantique REHOBOAM de la société INCITE INC représente l'exemple par excellence d'un outil d'IA implémentant une gouvernance – voire une dictature – algorithmique grâce à l'analyse prédictive du comportement des individus *i.e.*, de la *dataveillance* combinée à du *behavioral profiling*.

<sup>147</sup> Voy. K. YEUNG, M. LODGE, *Ibid*, 2019.

<sup>148</sup> C. CAMPBELL, « How China Is Using “Social Credit Scores” to Reward and Punish Its Citizens », disponible au lien <https://time.com/collection/>, consulté en avril 2020. Le gouvernement chinois utilise des caméras surpuissantes 500-megapixel pouvant identifier aisément tout individu dans une foule, même compacte.

d’algorithmes par les organismes publics comporte à la fois des avantages et des risques »<sup>149</sup>, le modèle chinois de société disciplinaire axée sur le contrôle social n’est pas sans rappeler la vision du panoptique de M. FOUCAULT dans *Surveiller et punir*<sup>150</sup> ou encore la dystopie imaginée par G. ORWELL dans *1984*<sup>151</sup>. Qui plus est, la Chine se trouve être le premier pays au monde à utiliser des “AI judges” : « [T]he Beijing Internet Court [...] has launched an online litigation service center featuring *an artificially intelligent female judge*, with a body, facial expressions, voice, and actions all modeled off a living, breathing human [...] This virtual judge, whose abilities are based on intelligent speech and image synthesizing technologies, is to be used for the completion of “repetitive basic work” only [...] she’ll mostly be dealing with litigation reception and online guidance. [...] According to court president Zhang WEN, *integrating AI and cloud computing with the litigation service system* will allow the public to better reap the benefits of technological innovation in China »<sup>152</sup>.

Pour A. ROUVROY, nous assistons en fait, avec l’émergence de la gouvernance algorithmique, à un véritable « basculement sémiotique et épistémologique ». En effet, « [d]’une civilisation du texte et du signe nous serions en passe de basculer vers une civilisation de l’algorithme et du signal. Voilà qui, en soi, est proprement bouleversant. Il s’agirait, aux dires de certains, d’une nouvelle étape dans l’histoire des formes d’écriture »<sup>153</sup>. Le passage à une

<sup>149</sup> É. DEGRAVE, L. GERARD, J. MONT, « L’action publique et ses juges face aux défis de la numérisation: La situation belge », Universiteit Leiden, 2019, p. 1, disponible au lien <http://www.crid.be/pdf/public/>, consulté en avril 2020.

<sup>150</sup> M. FOUCAULT, *Surveiller et punir, naissance de la prison*, Paris, Gallimard, 1975. Voy. D. MARTUCELLI, « Michel Foucault, la rationalisation comme assujettissement », *Sociologies de la modernité*, Paris, Gallimard, coll. Folio essais, 1999, pp. 289-322.

<sup>151</sup> On pourrait également tracer un parallèle, dans une certaine mesure, avec certains aspects de *Brave New World* de A. HUXLEY ou encore avec le film de SF *Gattaca* de A. NICCOL sorti en 1997.

<sup>152</sup> M. PILLAI, « China Now Has AI-Powered Judges : Is AI arbitration the future? », 2019, disponible au lien <https://radiichina.com/>, consulté en avril 2020. Nos italiques. Voy. K. YEUNG, « Why Worry about Decision-Making by Machine? », K. YEUNG, M. LODGE, *Ibid.*, pp. 21-48.

<sup>153</sup> A. ROUVROY, « La robotisation de la vie ou la tentation de l’inséparation », H. JACQUEMIN, A. DE STREEL, *L’intelligence artificielle et le droit*, Op. cit., p. 18. Voy. aussi A. ROUVROY, Th. BERNIS, « Gouvernamentalité algorithmique et perspectives d’émancipation : Le disparate comme condition d’individuation par la relation ? », *Réseaux*, vol. 177, n°1, 2013, pp. 163-196 ; A. ROUVROY, Th. BERNIS, « Le nouveau pouvoir statistique. Ou quand le contrôle s’exerce sur un réel normé, docile et sans événement car constitué de corps “numériques” », *Multitudes*, vol. 40, n°1, 2010, pp. 88-103 ; U. ECO, *Le signe : Histoire en analyse d’un concept*, trad. J.-M. KLINKENBERG, Bruxelles, Labor, 1988.

gouvernementalité algorithmique implique des changements sociétaux à la fois importants et difficiles à anticiper si ce n'est, de façon hautement spéculative, en envisageant les futurs possibles : « algorithms are already beginning to, and will inevitably continue to, drive radical changes to the way in which societies, including public administration will be organized often in unexpected and hitherto unimaginable ways »<sup>154</sup>. Toujours selon ROUVROY, ce basculement affectera les modes de production du savoir dans de nombreux domaines mais aussi les modes d'exercice du pouvoir et de la gouvernance : en effet, « c'est à partir des signaux numériques transpirant des trajectoires, relations et interactions des individus, plutôt qu'en fonction de normes émanant de processus délibératifs antérieurs que se construisent les profils/scores/appariements à travers lesquels les individus sont classés, évalués, récompensés ou sanctionnés, à travers lesquels s'évaluent les mérites et les besoins ou encore les opportunités ou la dangerosité que recèlent les formes de vie »<sup>155</sup>. On le voit bien avec l'exemple paroxysmique de la Chine, la gouvernance algorithmique combinée à l'effet boîte noire discuté *supra*, risque bien, si elle n'est pas encadrée de façon adéquate, d'affecter les libertés individuelles<sup>156</sup>.

Pour C. VILLANI, c'est à l'État qu'incombe la responsabilité décisive de « donner un sens à l'IA », c'est-à-dire non seulement de lui « donner un cap », une direction, mais aussi une « signification » articulée autour de valeurs fondamentales en vue d'en garantir l'explicabilité, la transparence et conjurer l'effet boîte noire inhérent à cette technologie émergente<sup>157</sup>.

#### d. *Quid des hard cases ?*<sup>158</sup>

Depuis les travaux de H. HART, on distingue classiquement, en droit, entre les « cas faciles » et les « cas difficiles » ou « hard cases »<sup>159</sup>. Pour ces derniers, il n'existe pas, de prime abord, de meilleure solution possible : celui qui juge de l'affaire – qu'il s'agisse de l'avocat en première ligne ou du juge en dernière instance – devra faire appel à son pouvoir

<sup>154</sup> K. YEUNG, M. LODGE, *Ibid.*, 2019, p.1.

<sup>155</sup> A. ROUVROY, *Op. cit.*, idem.

<sup>156</sup> A. STROWEL, E. MARIQUE, « Gouverner par la loi ou les algorithmes : de la norme générale de comportement au guidage rapproché des conduites », *Dalloz IP/IT*, N°10, p.521, 2017. Voy. aussi J.-J. SAHEL, « Le rôle du régulateur à l'ère de la convergence numérique : nouveaux pouvoirs ou nouvelles approches », D. GAFFAR RAHMOUNI-SYED (dir.), *Les défis du numérique*, Bruxelles, Bruylant, 2019.

<sup>157</sup> C. VILLANI, « Donner un sens à l'intelligence artificielle », *Les défis du numérique*, *Op. cit.*, 2019, p.5-11.

<sup>158</sup> Cette section s'inspire de certains éléments du cours de Philosophie du droit du Professeur Ph. COPPENS.

<sup>159</sup> P. BRUNET, « Analyse réaliste du jugement juridique », *Cahiers Philosophiques*, n°147/4, 2016, pp. 9-25.

discrétionnaire. Parfois, en effet, ni la jurisprudence et la règle du précédent ni même la législation et la doctrine n'offrent de réponse "toute faite" au problème qui se pose. Dans ces hypothèses, « les règles ne fournissent pas de solution : soit parce que le cas est inédit, soit parce que la règle est indéterminée, soit encore pour les deux raisons à la fois »<sup>160</sup>.

Pour H. HART, le juge doit certes s'en remettre à son pouvoir discrétionnaire mais il ne saurait pour autant s'écarter du droit positif<sup>161</sup>, ce qui est critiqué par R. DWORKIN qui, défendant une position cognitiviste allant à l'encontre du positivisme juridique strict, suppose que l'interprétation du droit, afin d'assurer sa cohérence (*i.e.*, son caractère fondamentalement moral), ne se subordonne pas uniquement au droit positif mais peut – et parfois *doit* – s'en écarter<sup>162</sup>. On le voit, pour reprendre une distinction typique dans le champ de l'éthique, nous sommes, avec DWORKIN, dans une morale déontologique qui dénote d'un certain sens du devoir (le juge *doit* assurer la cohérence du droit dans son ensemble), alors qu'en revanche, on ne voit pas comment une IA pourrait réfléchir autrement qu'en termes de morales purement utilitaire, instrumentale et conséquentialiste.

S'écartant d'une vision strictement positiviste à la Kelsen, où la validité d'une norme se réduit à son processus d'édiction par un pouvoir habilité (*Ermächtigung*) qui implique qu'une loi valide peut être immorale, DWORKIN estime, quant à lui, que le droit ne saurait seulement s'identifier à des normes valides : il y a en effet des cas où la solution de droit fait appel à quelque chose de plus relevant de l'intuition morale et de la reconnaissance immédiate du caractère éthique, voire de l'éthicité d'une situation (*Sittlichkeit*<sup>163</sup>).

Or on ne voit pas comment les raisonnement d'un système computationnel, syntaxique et asémantique par essence, pourrait s'écarter du point de vue strictement positiviste et non-cognitiviste. Comment peut-on, dans un environnement binaire, obéissant rigide à la logique formelle et modale – aux principes d'identité, de non-contradiction et du tiers exclu – enseigner à la machine la nuance, les zones grises, les concepts flous et encore l'infinité des

<sup>160</sup> P. BRUNET, *Op. cit.*, p.17.

<sup>161</sup> H. HART, *The Concept of Law Oxford*, Oxford University Press, 1961.

<sup>162</sup> R. DWORKIN, *Law's Empire*, Cambridge, Massachusetts, Harvard University Press, 1986.

Voy. aussi, au sujet de la position cognitiviste en droit, les travaux de J. RAWLS et de J. HABERMAS.

<sup>163</sup> Voy. G.H.F. HEGEL, *Phenomenology of Spirit*, trad. A.V. MILLER, Oxford, Oxford University Press, 1977, pp. 364-409 ; *Elements of the Philosophy of Right*, *i.e.*, une normativité transcendente par rapport aux individus.

contextes et situations, des coutumes et des traditions qui se développent et s'entretiennent par l'habitude, la répétition, voire un certain mimétisme<sup>164</sup> ? Ces éléments sont pourtant au cœur de la pratique du droit ; ils informent les jugements tant moraux que judiciaires et font, très certainement, à la fois la beauté et la complexité du métier de juriste.

#### e. *Retour sur l'irréductibilité de l'esprit humain*

Nous disions *supra* et en référence aux travaux de J. SEARLE que, d'un point de vue épistémologique, on ne saurait réduire l'intelligence humaine à une capacité de calcul ou à un algorithme, aussi complexe soit-il. Il manquera toujours à un programme informatique cette plus-value, cet "on-ne-sait-quoi", que l'on désigne abstraitement sous le vocable "esprit," ou "conscience" (ou encore *mind* en anglais) ; *i.e.*, il lui manquera toujours la "visée intentionnelle" (l'« intentionnalité ») dont parle E. HUSSERL<sup>165</sup> qui projette, si l'on peut dire, du sens dans le monde empirique car, explique SEARLE, « [a]ucun programme informatique n'est en lui-même suffisant pour donner un esprit à un système »<sup>166</sup>. En outre, nous l'avons dit, les programmes d'ordinateur fonctionnent sur une base purement formelle et procédurale (en particulier, avec la logique modale et computationnelle) et dès lors *syntaxique*, alors que l'esprit procède de contenus mentaux *i.e.*, de contenus *sémantiques* projetés dans le monde vécu par l'intentionnalité qui s'incarne avant tout dans un *corps propre*, dans la chair, et non un corps purement matériel fut-il animé de façon artificielle<sup>167</sup>. Ainsi, pour P. LÉVY, « l'intelligence artificielle se contente dans les faits de fournir une boîte à outils hétéroclite (règles logiques, syntaxes formelles, méthodes statistiques, simulations neuronales ou sociobiologiques...) qui n'offrent pas de solution générale au problème d'une modélisation mathématique de la cognition humaine »<sup>168</sup>.

<sup>164</sup> L. WITTGENSTEIN, *Recherches Philosophiques*, trad. E. RIGAL, Paris, Gallimard, 2004, § 198-202, pp. 125-127.

<sup>165</sup> E. HUSSERL, *Idées directrices pour une phénoménologie*, trad. P. RICEUR, Paris, Gallimard, coll. Tel, 1985.

<sup>166</sup> J. SEARLE, « Les ordinateurs peuvent-ils penser », *Ibid.* pp. 37-56 ; H. PUTNAM, *Représentations et Réalité*, Paris, Gallimard, 1990. A contrario, voy. J. JAYNES, *The origin of consciousness in the breakdown of the bicameral mind*, Boston, Houghton Mifflin, 1976.

<sup>167</sup> Voy. les travaux de M. HEIDEGGER et, en particulier, sa notion de *Dasein*. Cf. avec *La phénoménologie de la perception* de M. MERLEAU-PONTY et sa notion de corps propre/phénoménal ou encore de chaire ontologique.

<sup>168</sup> P. LÉVY, « Intelligence artificielle et sciences humaines », P. LEVY's blog, 2014, disponible au lien <https://pierrelevyblog.com/>, consulté en avril 2020.

f. *Sur le problème que pose l'indétermination du droit ou de l'ambiguïté syntaxique et sémantique des textes juridiques à la source du droit*

Si les contenus sémantiques échappent aux programmes d'ordinateurs et ALPHAGO, DEEPBLUE, ROSS et WATSON, « [don]'t know what [they are] saying »<sup>169</sup>, un problème important se pose également du fait du caractère fondamentalement indéterminé et évolutif des concepts juridiques (ou ce que H. HART qualifie de « texture ouverte du droit »<sup>170</sup>) : les règles juridiques qui appartiennent plutôt au domaine des sciences sociales ne connaissent pas une évolution linéaire, à l'inverse des lois empiriques qui appartiennent, quant à elles, au domaine des sciences dites exactes. Le raisonnement juridique est surtout affaire d'appréciation et d'interprétation proprement *humaine* ; au-delà de la logique froide du syllogisme judiciaire, il met aussi en œuvre des *softskills* telles l'empathie et le bon sens. Il est difficile de croire qu'une IA *weak* pourrait jamais, dans le contexte d'un jugement, opérer un revirement jurisprudentiel ou encore, dans le contexte d'un raisonnement juridique, plaider en la faveur d'un revirement de cette nature. Nous l'avons suggéré *supra* : on ne saurait réduire le travail du juge, de l'avocat et plus généralement du juriste à l'exercice de la qualification juridique, à l'application du syllogisme judiciaire et à la résolution de problèmes d'interprétation du droit. Il y a également quelque chose d'irréductible à la pratique du droit ; un “on-ne-sait-quoi” qui relève plus de l'intuition et du *sensus communis* que du droit strictement positif et qui réside dans la dimension éminemment subjective et relative de la faculté de juger<sup>171</sup> ou encore dans la reconnaissance immédiate de l'éthicité ou de la non-éthicité d'une situation donnée – l'avocat étant toujours, en effet, et comme le veut l'adage, le premier juge de l'affaire.

Il faut enfin reconnaître au moins trois causes d'indétermination du droit qui compliquent fortement le projet d'utilisation de l'IA dans la pratique du droit : cette indétermination tient, d'abord, à l'ambiguïté syntaxique et/ou sémantique des textes juridiques eux-mêmes qui peuvent toujours donner lieu à au moins deux lectures différentes et contradictoires – excepté quand *l'interpretatio cessat in claris* (ce qui est relativement rare en pratique et ne s'applique pas, au demeurant, pour les *hard cases*). Ensuite, elle tient à la difficulté d'identifier la

<sup>169</sup> R. SCHANK, *Op. cit.*

<sup>170</sup> Voy. H. HART, *Ibid. cit.*

<sup>171</sup> Voy. plus largement sur ces questions, I. KANT, *Critique of the Power of Judgement*, trad. P. GUYER et E. MATTHEWS, Cambridge, Cambridge University Press, 2000 ; F. NIETZSCHE, *Par-delà bien et mal. Prélude d'une philosophie de l'avenir*, trad. C. HEIM, Paris, Gallimard, 1971.

référence objective des catégories juridiques abstraites utilisées par les textes et ce, nonobstant le recours aux travaux préparatoires. Il s'agit des fameux concepts flous utilisés par le droit dont la signification est à géométrie variable et permet, le cas échéant, une interprétation évolutive et flexible du droit en fonction de l'évolution des us, coutumes et mœurs de la société au fil du temps. Enfin, l'indétermination du droit est liée au fait que les techniques d'interprétation (téléologique, littérale, originaliste, évolutive, etc.) peuvent entrer en conflit les unes avec les autres et conduire à des solutions opposées alors qu'aucune de ces techniques ne s'impose plus qu'une autre<sup>172</sup>. Comment une IA pourrait-elle, même à l'aide du ML et du *deep learning*, traiter de façon efficace et en toute transparence les zones grises, les concepts flous ou encore, plus précisément, l'indétermination du droit et l'ambiguïté syntaxique et sémantique des textes juridiques qui sont à sa source ?

#### §4 – Évaluation critique des applications actuellement disponibles

##### a. *Practice Management*

Les outils de management que l'on trouve en Belgique dans les logiciels KLEOS, DLEX ou encore SECIB sont axés sur l'automatisation de certaines tâches de la pratique du métier d'avocat. Cela étant, les moteurs de recherche intégrés dans ces logiciels sont en constante évolution : de la recherche sur base syntaxique et d'opérateurs booléens, nous avons eu l'occasion de relever *supra* que des projets de développements sont en cours chez certains éditeurs de revues juridiques comme LARCIER et WOLTERS KLUWER pour y inclure l'IA et le NLP<sup>173</sup>. KLEOS et DLEX mettent d'ailleurs en œuvre, depuis peu, des technologies liées à l'IA directement dans leurs programmes phares, notamment au travers du *smart tool*, un « assistant intelligent » qui parcourt tous les jugements pertinents et LEGAL INSIGHTS qui implémente du *predictive analytics*<sup>174</sup>. STRADA LEX n'est pas en reste dans la mesure où cet éditeur de revues a annoncé, on l'a dit *supra*, un partenariat avec LEXIS NEXIS qui est un acteur majeur dans le domaine de l'IA et du droit tandis que WOLTERS KLUWER a annoncé en 2017 un partenariat

<sup>172</sup> P. BRUNET, *Ibid.*, pp. 15-16.

<sup>173</sup> J.-P. BUYLE et A. VAN DEN BRANDEN, *Ibid.*, p. 292. Voy. la note en bas de page.

<sup>174</sup> « LEGAL INSIGHTS – votre “aide d'urgence” – est le *premier volet d'une série de module LegalTech* que développent actuellement les experts de WOLTERS KLUWER. Vous pouvez les considérer comme une extension de votre propre intelligence : ils vous aident à valoriser encore mieux vos compétences et votre expertise et améliorent d'un cran vos conseils et votre efficacité ».

Voy. <https://legalworld.wolterskluwer.be/>, consulté en mars 2020. Nos italiques.

avec la start-up française PREDICTICE spécialisée en justice prédictive. Si les logiciels de *practice management* tels que SECIB semblent encore relativement basiques à l’heure actuelle en ce qui concerne la Belgique, on peut présager qu’au fil de leur évolution et de la démocratisation des technologies mises en œuvre dans ces programmes, ils intégreront de plus en plus, à l’instar de leurs homologues américains, des outils “intelligents”. Cette rhétorique est déjà très clairement élaborée, on l’a vu, sur les vitrines virtuelles et dans les discours commerciaux des grands éditeurs de revue et des start-ups actives dans le domaine des *legaltechs*.

b. *Document Automation and « Lawbots »*

Bien que les *lawbots* soient d’ores et déjà monnaie courante, notamment en Chine et aux États-Unis, ils n’ont pas encore débarqué au plat pays. Rien n’empêche néanmoins d’imaginer l’arrivée de ces robots désincarnés sur la toile et, par exemple, sur les sites des cabinets ou des administrations publiques, sous la forme de bots informatiques, de *chatbots*, que l’on nomme également “dialogueur” ou “agent conversationnel”. L’exemple d’ALEXA (AMAZON), de SIRI (APPLE) ou encore de CORTANA (MICROSOFT) illustre bien à quel point ce genre de technologie a pu évoluer vite ces dernières années et fait déjà partie du quotidien de millions de personnes, notamment au travers de la domotique. Dans le domaine du droit, on l’a dit, ROSS (IBM) peut précisément déjà se comporter en agent conversationnel : « [a]ux États-Unis, des « robots avocats » sont déjà à l’œuvre et semblent converser en langage naturel avec les humains »<sup>175</sup>. Le problème qui se pose ici est celui de la déshumanisation du rapport client/conseil ou encore du rapport avocat/associé en ce que l’impasse est faite sur la dimension interpersonnelle et l’ensemble des compétences que nous avons identifiées *supra* comme relevant des *softskills* i.e., sens commun, intelligence émotionnelle, éthique professionnelle, etc. D’un autre côté, le grand remplacement n’est pas non plus pour demain ; à l’heure actuelle l’automatisation du traitement des documents, pièces de procédures et même l’interaction avec des *lawbots* restent limitées et ne s’envisagent que comme un outils additif ou supplétif aux authentiques rapports “*in the flesh*”. À l’inverse, si l’on jette un regard plus optimiste sur cette évolution, on peut dire que le débarquement de l’IA dans les cabinets d’avocats permet, en fait, de réhumaniser certains aspects aliénant et répétitif de la profession – comme la gestion des prestations et la facturation par exemple – et peut, dès lors, permettre aux travailleurs de chair et d’os de se concentrer sur les tâches stimulantes sur le plan

<sup>175</sup> CEPEJ, *Charte éthique*, *Ibid.*, p.12.

intellectuel ou faisant davantage appel aux *soft* qu'aux *hardskills* :« [with ROSS] lawyers will be able to avoid some of the mundane tasks in favor of being able to focus on the nitty-gritty aspects of each case »<sup>176</sup>.

c. *Computational Models of Legal Reasoning and Predictive Analytics via Comparative Legal Search Engines*

Sur le site de WOLTERS KLUWER, au sujet de l'outil LEGAL INSIGHTS, on peut lire ceci :« [c]et assistant digital intelligent vous aide à estimer rapidement la probabilité de remporter des affaires en matière de rupture du contrat de travail et à vous constituer un dossier solide. Conçu par et pour des juristes, ce *moteur de recherche intelligent* spécialisé dans ce domaine *réfléchit* à l'aide de l'IA et du Big Data »<sup>177</sup>. Nonobstant les progrès technologiques de l'IA appliquée au droit depuis quelques décennies, il est en fait incorrect d'affirmer que des softwares comme ROSS, LEXIS NEXIS ou encore LEGAL INSIGHTS “pensent” ou “raisonnent” au sens courant du terme ; du moins, il s'agit d'un abus de langage, voire d'une vue de l'esprit. En l'occurrence, l'usage de cette rhétorique relève avant tout d'un discours commercial qui tire avantage de l'engouement pour les *legaltechs* et l'IA en général. En effet, de même que les logiciels de traduction automatique ne traduisent pas mais bien plutôt calculent l'estimation statistique de la meilleure traduction possible par comparaison avec un vaste volume de données, de même les IA des systèmes experts légaux et des logiciels d'aide à la décision juridique se basent plus sur de l'apprentissage statistique des contextes via le ML que sur des règles abstraites<sup>178</sup>. En ce qui concerne spécifiquement l'IA appliquée au droit, d'après Vern R. WALKER, « [t]he ultimate goal of legal research by lawyers and decision makers is to find arguments and reasoning *reported in the past*, so that they can evaluate *the likelihood of success of those and similar arguments*, and can generate new arguments to use in future cases [...] this suite of tasks is also the ultimate goal of software analytics, such as WESTLAW, FASTCASE, ROSS, etc. That's the direction in which we are all headed with automating legal research »<sup>179</sup>.

<sup>176</sup> <https://rossintelligence.com/>, consulté en mars 2020.

<sup>177</sup> Voy. <https://legalworld.wolterskluwer.be/>, consulté en mars 2020. Nos italiques.

<sup>178</sup> Ou ce que l'on appelle, en informatique, une “ontologie”. Voy. D. CARDON, *Op. cit.*, p. 60.

<sup>179</sup> D. REMUS, F.S. LEVY, « Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law », 2016, pp. 24-25, disponible au lien <https://ssrn.com/abstract=2701092>, consulté en mars 2020.

Si ces logiciels constituent très certainement à la fois un atout et un avantage concurrentiel pour un cabinet, il n'en demeure pas moins qu'il s'agit d'abord et avant tout d'outils supplétifs qui ne sauraient se substituer au travail de fond de l'avocat : d'une part « many clients want more than a serie of statistical probabilities », et, d'autre part, « effective advising encompasses more than prediction. It requires a lawyer to understand a client's situation, goals, and interests »<sup>180</sup>. La compréhension des situations, des buts poursuivis ou encore de l'intérêt du client requiert, nous l'avons dit, une certaine intelligence émotionnelle et des *softskills* qui, pour l'heure, ne peuvent être simulés adéquatement par la machine du fait des limites inhérentes au projet de création d'une IA forte (*strong AI*) qui suppose, notamment, la compréhension complète de l'esprit humain quand nos connaissances sur le cerveau demeure, pour reprendre l'expression de J. SEARLE, « maigre ».

Si les prédictions statistiques qu'offrent aujourd'hui l'IA peuvent s'avérer utiles pour informer un cas d'espèce, les éventuels raisonnements juridiques que peuvent effectuer une IA faible s'avèrent encore trop proches de l'application froide et mécanique du syllogisme judiciaire sans prise en compte de règles abstraites et de ce que l'on pourrait appeler une : “ontologie sémantique du droit”.

<sup>180</sup> D. REMUS, F.S LEVY, *Op. cit.*, p. 32.

### TROISIÈME PARTIE

#### VERS UNE COMMON LAWISATION DU DROIT ?

L'introduction du *machine learning*, du *data mining* et de l'IA dans la sphère du droit risque de constituer, nous l'avons vu avec A. ROUVROY, un véritable « basculement sémiotique et épistémologique » et de révolutionner la pratique du droit. Dans ce mémoire, nous avons préféré l'expression « changement paradigmatique » entendue au sens de T. KUHN dans *La Structure des Révolutions Scientifiques* – le droit, ou plus abstraitement la science juridique, étant considérée comme une branche des sciences humaines et sociales<sup>181</sup>. Le droit relève en effet d'une science “*soft*” par opposition aux sciences dites “dures” que sont la physique ou les mathématiques par exemple. Avec l'arrivée de l'IA, du ML, du DM, ou encore du NLP dans la sphère du droit, cette dichotomie entre les sciences *soft* et *hard* devient de plus en plus poreuse et perméable : des outils analytiques et statistiques font leur entrée dans le juridique au sens large qui se définissait pourtant encore naguère mieux en terme d'herméneutique ou d'art que de science au sens strict. À l'heure du numérique et de l'IA, l'ancien et le moderne se rencontrent dans le droit, l'analytique et l'herméneutique se mêlent, l'anecdotique devient *anec-data*<sup>182</sup>.

Bien plus, les SEL et ERP mettant en œuvre de l'IA et des outils de comparaisons analytiques et casuistiques intégrés dans la pratique des avocats et des juristes de tradition continentale et civiliste pourraient hisser au rang de principe général de droit la règle du précédent (*stare decisis*), pourtant caractéristique de l'ordonnancement juridique des pays de *common law*. En effet, comme le note A. STROWEL en commentant les travaux de A. GARAPON et J. LASSÈGUE, « [l]e droit pourrait [...] devenir *calculable*, par exemple à partir des précédents ou d'autres régularités désormais enregistrables et saisissables par les machines dans la masse des données. S'insinue ainsi une nouvelle logique qui fait fi du raisonnement et de la causalité juridique et s'appuie surtout sur des corrélations mathématiques [...] Dans le monde numérique, le traitement désormais possible de l'information par le calcul (par exemple des antécédents socio-économiques d'un délinquant pour évaluer un risque de récidive ou des précédents judiciaires pour évaluer le montant d'une indemnité de licenciement) risque de se substituer à la délibération sur le cas d'espèce. *L'algorithme de calcul devient un fétiche au*

<sup>181</sup> Ou “*Humanities*” dans les pays anglo-saxons.

<sup>182</sup> « *Objective data replaces subjective “anec-data”* ».

Voy. <https://lexmachina.com/what-we-do/>, consulté en mars 2020.

*pouvoir magique*, il permet de prédire l'avenir ou la bonne décision, autrefois jugés plus incertains »<sup>183</sup>. Comme le relève en outre la Charte éthique dont nous avons discuté *supra* : « [N]’y a-t-il pas un risque de cristallisation de la jurisprudence, selon une logique reproductive ? Alors que les décisions judiciaires ont vocation à évoluer en fonction de l'évolution du cadre normatif (national, européen ou international), de la jurisprudence de référence (par exemple des juridictions suprêmes, des cours européennes) ou du contexte socio-économique, la “norme” issue du nombre ne deviendrait-elle pas un standard auquel les magistrats seraient incités à se référer sans questionnement, avec un effet induit d’uniformisation excessive des décisions judiciaires ? »<sup>184</sup>. En effet, l’usage de l’IA dans le droit emporte avec lui à la fois le risque de cristallisation de la jurisprudence selon une logique reproductive et, partant, la généralisation de la règle du précédent, mais aussi le problème que pose la substitution du calcul probabiliste à la délibération sur le cas d’espèce, l’IA devenant dans ce contexte, en quelque sorte, la quintessence de l’expert judiciaire.

Par ailleurs, nous avons mentionné dans l’introduction du mémoire, qu’un autre symptôme de la *common lawisation* du droit continental de tradition romano-germanique et civiliste résidait dans le fait que les modes alternatifs de résolution des conflits ou *alternative dispute resolution* qui nous viennent des États-Unis et du Canada prenaient, ces dernières années, une importance croissante et contribuaient, en conséquence, à faire concurrence au circuit judiciaire classique et à privatiser la justice. L’introduction de l’IA et la généralisation de son utilisation, tant par les avocats et les juristes que par les juridictions risque d’accentuer ce phénomène : « les décideurs publics voient l’opportunité de mieux réguler les flux de nouvelles procédures dans les tribunaux et de se doter d’un levier de réduction des coûts de fonctionnement de la justice. Il s’agirait ainsi d’inciter les justiciables à recourir par ce biais à des modes alternatifs de résolution de litiges (conciliation, médiation ou encore arbitrage) »<sup>185</sup>. L’IA constitue en effet une opportunité de choix qui pourrait permettre de rendre la justice plus efficace et contribuer à résorber l’arriéré judiciaire dont l’ampleur, dans certaines juridictions, notamment en Belgique, est dramatique. C’est en ce sens qu’il y a lieu de comprendre l’affirmation selon laquelle, toute proportion gardée, la réception en droit

<sup>183</sup> A. STROWEL, « Des véhicules autonomes à l’intelligence artificielle : entre droit, politique et éthique », *Des véhicules autonomes à l'intelligence artificielle*, *Op. cit.*, p. 15. Nos italiques. Cf. A. GARAPON et J. LASSÈGUE, *Justice digitale. Révolution graphique et rupture anthropologique*, Paris, PUF, 2018. Nos italiques.

<sup>184</sup> CEPEJ, *Charte éthique*, *Ibid.*, p. 20.

<sup>185</sup> CEPEJ, *Charte éthique*, *Ibid.*, p. 34.

européen continental de tradition civiliste des nouveaux outils de comparaisons analytiques et casuistiques peut être perçu comme symptomatique de ce que l'on pourrait appeler une *common lawisation* du droit.

Il ne s'agit évidemment pas de poser un jugement de valeur : nous ne prétendons pas que le changement paradigmatique que provoque l'introduction de l'IA dans l'ordonnancement juridique de l'Ancien Monde est préjudiciable *per se* que ce soit pour les juristes ou les justiciables ; il est néanmoins nécessaire que les professionnels du droit ainsi que les décideurs publics en soient conscients et comprennent les tenants et aboutissants ainsi que les défis que pose ce phénomène, notamment en ce qui concerne l'autonomie et l'indépendance de la justice d'autant que « [t]o pretend that AI is not changing every aspect of society is to ignore a vast amount of evidence. [...] while some lawyers fear that robot-lawyers will replace human lawyers, many innovative legal minds envision a legal profession in which attorneys shed the burden of mundane tasks and spend more time engaged in the higher-level aspects of lawyering. These innovators also believe that AI has the potential not only to create a new type of “legal-tech” employment, but also to increase access to justice for millions of individuals »<sup>186</sup>. Ainsi, comme nous le disions en introduction, l'IA constitue à la fois le cheval de Troie d'une "*common lawisation*" du droit européen continental et une boîte de Pandore contenant l'espoir d'une justice meilleure, plus efficace et moins coûteuse. Il importe dès lors de s'assurer que la digitalisation et l'intégration d'outils liés à l'IA dans les cabinets et les prétoires ne se fera pas sur l'autel de principes généraux de droit ou d'impératifs éthiques ou déontologiques tels la protection de certaines valeurs fondamentales. Pour ce faire il y a lieu de privilégier un usage à la fois raisonné et raisonnable de l'IA en tenant compte de ses potentialités et de ses limites ainsi que des risques inhérents à son développement et à la généralisation de son usage. Les approches dites *human rights* et *ethical by design* nous semblent toute indiquées dans la mesure où ces approches permettent de s'assurer que sont intégrées, dès la phase de conception et d'apprentissage machine (*i.e.*, *ex ante*), des règles interdisant de porter atteinte à la dignité et à l'autonomie de la personne et de la justice. À l'instar des trois lois de la robotique formulées par l'auteur de SF Isaac ASIMOV<sup>187</sup>, il nous semble fondamental que des règles éthiques soient intégrées dans le code source des IA et servent de pierre angulaire aux algorithmes afin d'éviter toute dérive.

<sup>186</sup> J.L. JACOBOWITZ, J. ORTIZ, « Happy Birthday Siri! Dialing in Legal Ethics for Artificial Intelligence, Smart Phones, and Real Time Lawyers », *Op. cit.*, pp. 30-31.

<sup>187</sup> I. ASIMOV, « Runaround », *I, Robot – The Isaac Asimov Collection* (ed.), New York, Doubleday, 1950, p. 40.

## CONCLUSION

### PERSPECTIVES POUR L'AVENIR

« D'aucuns nous proposent [...] le salut par la contrainte ou par la prévision du comportement humain. Rien de plus aisé : il suffit de conditionner l'homme pour en faire un pantin articulé. Rien de plus dangereux aussi, car il devient alors un robot. Mais au service de quels desseins ? formulés par qui ? Si, quelque jour, des pontifes ou des ordinateurs bipèdes en arrivaient à assujettir la société à leurs règles coercitives, déshumanisantes, ou mercantiles, en imposant à tous des directives de leur cœur de pierre ou de matière plastique persillée de métal, alors l'homme se devra d'organiser une nouvelle conjuration pour puiser dans les richesses de l'amour, de l'art et du libre arbitre responsable, de quoi fonder une conscience et une éthique rénovées qui soient à sa mesure. Car on ne peut traiter de la science et de la morale indépendamment de la totalité de l'être humain, pas plus qu'on ne peut étudier le chant des rossignols avec des oiseaux empaillés »<sup>188</sup>.

« There is no threshold that makes us greater than the sum of our parts, no inflection point at which we become fully alive. We can't define consciousness because consciousness does not exist. Humans fancy that there's something special about the way we perceive the world, and yet we live in loops as tight and as closed as the [AI] hosts do, seldom questioning our choices, content, for the most part, to be told what to do next »<sup>189</sup>.

Au début de ce mémoire, nous nous posions la question de savoir si, en droit européen et en droit belge, on allait bientôt assister à l'introduction et puis à la généralisation de l'utilisation d'outils associés au *machine learning*, au *data mining* et au *predictive analytics* pour instruire des dossiers, rédiger des conclusions et compiler des pièces de procédure, voire pour rendre des jugements. Nous nous demandions aussi si et dans quelle mesure on pouvait considérer que l'introduction des IA dans les prétoires était une solution toute désignée aux problèmes que connaît actuellement la justice, à savoir, en particulier, l'arriéré et l'aléa judiciaire. Nous nous demandions enfin si l'IA, cette boîte de Pandore, contenait l'espoir d'un remède efficace et bon marché permettant nous seulement d'améliorer la justice et l'efficacité des systèmes judiciaires mais aussi de renforcer les garanties de l'État de droit en ce qui concerne les droits individuels et la qualité des services publics. À ces questions, il y a lieu, au vu de ce qui précède, de répondre par l'affirmative : l'IA est déjà présente dans la sphère du droit, tant en Europe qu'en Belgique au travers de diverses solutions et logiciels, en particulier, des SEL et ERP de nouvelle génération. Dans les années à venir, nul doute que l'IA continuera de se répandre et d'élargir sa sphère d'influence dans la mesure où elle apporte une solution au problème que pose, tant pour le justiciable que pour la justice en général, l'arriéré et l'aléa judiciaire. Si l'IA pourrait bien constituer le cheval de Troie d'une

<sup>188</sup> J. SEMAL, *La science doit être à la mesure de l'homme*, 20/11/89, pp. 8-10 ; cf. J.-F. FOGEL, *La percée de l'intelligence artificielle*, *Op. cit.*, Voy. les anciennes coupures de journaux en annexe.

<sup>189</sup> Dr. R. FORD, « Trace Decay », *Westworld*, saison 1, 2016, disponible au lien <https://quotecatalog.com/>, consulté en avril 2020.

« *common lawisation* » du droit européen continental et s’il semble d’ores et déjà ardu de freiner l’inéluctable dissémination de l’IA dans la sphère du droit, il importe de s’assurer que la digitalisation et l’intégration de systèmes d’IA dans les cabinets et les prétoires ne se fera pas sur l’autel de principes généraux de droit (égalité des armes entre petits et grands cabinets) ou d’impératifs éthiques ou déontologiques (principe d’égalité, de non-discrimination), en particulier, que son développement ne se fera pas au détriment des DHF des justiciables (respect de la vie privée, protection des données à caractère personnel et des personnes vulnérables) et, singulièrement, du respect de la dignité et de l’autonomie de la personne conçue au sens large. En effet, ce remède que semble constituer l’IA, nous l’avons dit et répété, ne saurait simplement et unilatéralement s’avérer bénéfique. Vu les risques potentiels et les dangers inhérents que comporte l’IA, celle-ci se caractérise mieux en terme de *pharmakon* (φάρμακον) – de philtre, à la fois remède et poison – que de panacée<sup>190</sup>.

## § 1 – Éthique trans/posthumaniste

Selon A. ROUVROY, « [L]e remplacement total du travail humain par les robots pourrait advenir [...] endéans les 120 prochaines années. D’après des “experts en IA”, les machines auront surpassé les humains dans les tâches de traduction linguistique en 2024, dans la rédaction de dissertations en 2031, dans la conduite de camions en 2027, dans la vente en 2031. D’ici 45 ans, il y aurait 50 % de chance pour que les intelligences artificielles surpassent les êtres humains dans la totalité des “tâches”, ce qui devrait aboutir, d’ici 120 ans, au grand remplacement »<sup>191</sup>. Si le spectre de la « singularité » et de la « délégation aux machines » combiné à un certain déterminisme technologique continue d’inquiéter, il nous semble, qu’à l’heure actuelle, il est plus raisonnable de considérer que l’IA, dans la sphère du travail et notamment du droit, demeurera encore longtemps un outil additif et supplétif qui viendra en quelque sorte se superposer à l’intelligence humaine pour travailler de concert avec elle<sup>192</sup>. Les outils techniques liés à l’IA peuvent être envisagées, en ce sens, comme des

<sup>190</sup> J. DERRIDA, « La pharmacie de PLATON », *La dissémination*, Paris, Seuil, coll. “Tel Quel”, 1972, pp. 70-197.

<sup>191</sup> A. ROUVROY, « La robotisation de la vie ou la tentation de l’inséparation », H. JACQUEMIN, A. DE STREEL, *L’intelligence artificielle et le droit*, *Op. cit.*, p. 14 ; A. ROUVROY et B. STIEGLER, « Le régime de vérité numérique. De la gouvernementalité algorithmique à un nouvel État de droit », *Socio*, vol. 4, 2015, disponible au lien <http://socio.revues.org/1251>, consulté en mai 2020.

<sup>192</sup> Voy. J.-G. GANASCIA, *Op. cit.* en ce qui concerne le mythe de la singularité et, en ce qui concerne le mythe de la délégation aux machines, A. GARAPON, J. LASSÈGUE, *Op. cit.*

« prolongements de l'être humain, indissociable du devenir humain et de la valeur supérieure de l'homme au sein de la nature »<sup>193</sup>.

À l'époque qui est la nôtre et en ce qui concerne spécifiquement l'IA, l'approche "technoscientophobe" et l'heuristique de la peur telles que préconisées par H. JONAS nous semblent contreproductives. Selon l'auteur du *Principe responsabilité*, « il faut d'avantage prêter l'oreille à la prophétie de malheur qu'à la prophétie de bonheur » et JONAS de préciser, « en matière d'affaires d'un certain ordre de gravité – celles qui comportent un potentiel apocalyptique [on devrait dire eschatologique] – on doit accorder un plus grand poids au pronostic de malheur qu'au pronostic de salut. La présupposition de toute cette considération [est] qu'aujourd'hui et à l'avenir nous avons précisément affaire à des interventions de cet ordre de gravité, ce qui est en soi un fait nouveau dans les affaires humaines »<sup>194</sup>. Or JONAS écrit ces lignes à l'époque de la guerre froide alors que la menace d'un cataclysme nucléaire est bien réel : il est sans doute le premier penseur à prendre acte du fait que l'humanité est désormais capable de s'anéantir elle-même, soit directement par les armes de destruction massive, soit indirectement par l'altération des conditions nécessaires à sa survie sur Terre. Dans quelle mesure peut-on appliquer ce principe de précaution au développement de l'IA ? L'argument de la pente glissante et cette responsabilité hyperbolique appliqués à ces mythes que sont la « singularité technologique » et la « délégation aux machines » ne risquent-ils pas de freiner l'innovation, voire même d'inviter à l'inaction et figer l'humanité dans la contemplation pusillanime de la catastrophe à venir dans l'angoisse et les tremblements ? Les travaux du *Google X Lab* et de *CALICO (California Life Company)* combinent d'ailleurs déjà recherche en nanotechnologies, biotechnologies, informatique, sciences cognitives et IA le tout dans une perspective transhumaniste pleinement assumée : le désir utopique de modifier et de dépasser concrètement la condition humaine est devenu, ces dernières décennies, tangible. La révolution est en marche. La transcendance de l'être humain n'est donc plus exclusivement symbolique ou spirituelle – n'en déplaise à H. JONAS.

<sup>193</sup> G. HOTTOIS, *De la Renaissance à la Postmodernité*, Bruxelles, De Boeck et Larcier, 2002, p. 493 en référence à E. KAPP, *Grundlinien einer Philosophie der Technik* qui décrit déjà l'homme, au XIX<sup>ème</sup>, comme un « Deus ex Machina », au sens littéral, i.e., un dieu issu de la machine.

Cf. G. HOTTOIS *Philosophies des sciences, philosophies des techniques*, Paris, coll. Collège de France, Odile Jacob, 2004 ; *Philosophie et idéologies trans/post-humanistes*, Paris, coll. Librairie philosophique J. Vrin, 2017.s

<sup>194</sup> H. JONAS, *Le principe responsabilité, une éthique pour la civilisation technologique*, Op. cit., pp. 73-78 ; Voy., du même auteur, *Pour une éthique du futur*, Paris, Rivages poche, 1998 ; cf. J.-P. DUPUY, Op. cit.

Les perspectives technophobes et l'humanisme technophile suranné cèdent aujourd'hui le pas – dans un monde pluraliste et multiculturel qui se caractérise avant tout par sa complexité – à une technophilie évolutionniste et post-moderne telle que défendue, par exemple, par H.T. ENGELHARDT<sup>195</sup>. Pour ce dernier, « [d]es changements majeurs seront inévitables si nous demeurons une espèce libre et progressant technologiquement [...] En fait, il n'y a pas de raison de penser que seule une espèce sortira de la nôtre » ; « [s]i la nature humaine n'a rien de sacré [...], il n'y a pas de raison qu'elle ne soit pas radicalement modifiée, sur base de raisons particulières et avec prudence »<sup>196</sup>. La seule limitation à cette liberté technoscientifique étant, toujours selon ENGELHARDT, le respect de l'autonomie de la personne conçu au sens large et la prudence à l'égard des éventuelles conséquences destructrices, ce qui est n'est pas sans rappeler le « catastrophisme éclairé » préconisé par l'ingénieur et philosophe J.-P. DUPUY<sup>197</sup>.

En définitive, nous pensons qu'il y a lieu de privilégier sur ces questions une sorte de *via media* entre d'une part le catastrophisme technophobe et l'heuristique de la peur de JONAS et, d'autre part, la technophilie évolutionniste et post-moderne d'ENGELHARDT en accordant une « [a]ttention fidèle aux enjeux moraux de la *coévolution des acteurs humains et artificiels* [et en adoptant une] démarche à la fois éthique et épistémique privilégiant la prudence et la sagesse au sens des Anciens, c'est-à-dire la sensibilité morale, plutôt que les systèmes de règles et de principes moraux [...] [D]ans ce contexte de coévolution *connaître est inséparable de faire*, parce que découvrir de nouvelles règles morales est inséparable de créer de nouveaux agents sociaux, des inventions inouïes, jamais vues auparavant »<sup>198</sup>. En filigrane, nous plaçons en faveur d'une vision du futur de l'humanité faisant fi de l'opposition binaire classique entre *epistêmê* (ἐπιστήμη) et *technê* (τέχνη) ou encore entre conséquentialisme et déontologie afin de privilégier une éthique axée sur la vertu, la prudence (φρόνησις) et l'autonomie afin de réunir les conditions de possibilités pour une société « conviviale »<sup>199</sup>.

<sup>195</sup> H.T. ENGELHARDT, *The Foundations of Bioethics*, New York, Oxford University Press, 1986. Voy. G.

HOTTOIS (éd.), *Aux fondements d'une éthique contemporaine : H. JONAS et H.T. ENGELHARDT*, Paris, Vrin, 1993.

<sup>196</sup> H.T. ENGELHARDT, *Ibid.*, cité dans G. HOTTOIS, *De la Renaissance à la Postmodernité*, *Ibid. cit.*, p. 529.

<sup>197</sup> J.-P. DUPUY, *Op. cit.*

<sup>198</sup> P. DUMOUCHEL, L. DAMIANO, *Vivre avec les robots : essai sur l'empathie artificielle*, Paris, Seuil, coll. La couleur des idées, 2016, p. 30. Nos italiques.

<sup>199</sup> Sur l'éthique aristotélicienne et son renouveau au travers de la *virtue ethics*, voy. *i.a.*, M.C. NUSSBAUM, « Aristotelian Social Democracy », R. DOUGLASS, G. MARA, H. RICHARDSON (eds.), *Liberalism and the Good*,

## § 2 – Éthique synthétique : entre *meta-ethics* et *virtue ethics* ?

À notre sens, les préoccupations éthiques encadrant le déploiement de l'IA dans la pratique du droit doivent s'axer sur le respect des DHF, en particulier du principe de non-discrimination et du droit au respect de la vie privée. Néanmoins, du fait du caractère spécieux et relatif de ces principes normatifs et, force est de le reconnaître, eurocentrés, il importe de trouver un ou des principes permettant de s'en affranchir. Pour cette raison, l'idée plus large que nous empruntons à ENGELHARDT du respect de l'autonomie de la personne conçue de façon non-anthropomorphe et incluant, dès lors, l'ensemble des êtres conscients, libres, intelligents et sensibles, auquel on ajoute la prudence et le respect de la dignité humaine, nous semble plus pertinente dans le contexte du déploiement de l'IA dans la sphère du droit. Par ailleurs, le processus de conception et d'utilisation de l'IA et, en particulier, de l'analyse prédictive dans le droit doit s'intégrer dans un cadre éthico-normatif ayant l'élégance de la clarté. À ce titre, une attention particulière doit être accordée à la sécurité et à la transparence (vérifiabilité et explicabilité) ainsi qu'à la responsabilité des développeurs et à la maîtrise par l'utilisateur.

Jumelée à un "écosystème" fondé sur l'excellence et la confiance encadrant l'implémentation de l'IA dans les services publics en ce compris la sphère de la justice, nous proposons de qualifier cette approche d'éthique de la vertu et de la sollicitude en ce qu'elle se situe en-deçà du binarisme moral oscillant entre conséquentialisme utilitariste, déontologie kantienne ou encore responsabilité hyperbolique et heuristique de la peur. Bien au contraire, nous plaçons en faveur d'une technophilie audacieuse qui soit à la fois évolutionniste et post-moderne, qui prend acte de la coévolution des acteurs humains et artificiels et replace au centre de ses préoccupations la personne conçue au sens large, en ce compris les individus les plus vulnérables qui risquent, plus que quiconque, de se voir lésés par la quatrième révolution industrielle. Comme le relève C. LAZARO, « [t]oute recherche destinée à comprendre comment le droit et l'intelligence artificielle s'associent pour gouverner doit s'intéresser au sort réservé aux formes de vie – que celles-ci soient humaines ou non humaines – qui ne

New York, Routledge, 1990, pp. 203-252 ; « Non-Relative Virtues: An Aristotelian Approach », M.C.

NUSSBAUM, A. SEN (eds.), *The Quality of Life*, Oxford, Oxford University Press, 1993, pp. 242-270 ; *Frontiers of Justice*, Cambridge, Harvard University Press, 2006.

Voy. également A. MACINTYRE, *After Virtue*, London, Duckworth, 2<sup>nd</sup> Edition, 1985.

Sur la notion de « société conviviale », voy. I. ILLICH, *La convivialité*, Paris, Seuil, 1973.

s'adaptent pas, ne s'alignent pas ou résistent aux modes de vie néo-libéraux. En d'autres termes, il s'agit d'éviter que l'assemblage juridico-algorithmique contemporain ne finisse par protéger et favoriser uniquement des formes de vie en phase avec le modèle dominant »<sup>200</sup>.

Du point de vue de l'U.E. et vu la place croissante qu'occupe l'IA dans nos sociétés, le cadre juridique de celle-ci ne saurait continuer à reposer exclusivement sur de la *soft law* ou sur de la réglementation *indirectement* applicable d'autant plus que « dans le cadre du programme pour une Europe numérique, [la Commission a proposé] d'affecter plus de 4 milliards d'euros au soutien de l'informatique quantique et du calcul à haute performance, et notamment du traitement des données à la périphérie et de l'IA et des infrastructures de données et en nuage [sic] » et ce, notamment dans les administrations publiques, les services d'utilité publique et les autorités de surveillance financière<sup>201</sup>. Un cadre éthique clair balisant l'IA et la prédictive judiciaire suppose que « les professionnels eux-mêmes (juges, avocats, universités) s'en emparent collectivement pour en éprouver la faisabilité et qu'ils ne laissent pas des opérateurs privés concevoir seuls, avec quelques scientifiques non contrôlés, des logiciels et des modes de raisonnement ou de calculs abscons ou verrouillés. Les promesses ambitieuses (et non réalisées) de certaines *legaltechs* ne doivent donc pas occulter les immenses potentialités des technologies et le besoin d'applications adaptées et construites en lien direct avec la recherche scientifique et universitaire ainsi qu'avec l'ensemble des professionnels du droit, tels les magistrats, greffiers, avocats, notaires, huissiers et experts du terrain »<sup>202</sup>. Loin de perdre de vue les problèmes que posent la « modélisation algorithmique de la réalité » et les « modes de gouvernance basés sur une association entre logique juridique et algorithmique » ainsi que ce que A. ROUVROY, en référence à M. FOUCAULT, nomme le « nouveau régime de vérité numérique »<sup>203</sup>, la question n'est peut-être pas tant de savoir ce que l'éthique et le droit peuvent ou doivent faire pour l'IA mais bien plutôt de savoir, au contraire, ce que l'IA peut faire pour l'éthique, le droit et l'être humain en général<sup>204</sup>.

<sup>200</sup> C. LAZARO, *Op. cit.*, pp. 248-249.

<sup>201</sup> Commission européenne, *Livre blanc*, *Ibid.*, p. 9.

<sup>202</sup> CEPEJ, *Charte éthique*, *Op. cit.*, p. 46.

<sup>203</sup> A. ROUVROY, « Le régime de vérité numérique. De la gouvernementalité algorithmique à un nouvel État de droit », *Op. cit.*

<sup>204</sup> H. SEVILLE, D.G FIELD, « What Can AI Do for Ethics », M. ANDERSON, S.L. ANDERSON (eds.), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 499-511 ; S.L. ANDERSON, « How Machines Might Help Us Achieve Breakthroughs in Ethical Theory and Inspire Us to Behave Better », *Ibid.*, pp. 525-530.

## **BIBLIOGRAPHIE**

### **Législation européenne**

Convention européenne des droits de l'homme, STE n° 5

Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, STE n° 108 telle qu'amendée par le protocole STCE n°223

Directive (UE) 2019/882 du Parlement européen et du Conseil du 17 avril 2019 relative aux exigences en matière d'accessibilité applicables aux produits et services ;

Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE ;

Règlement (UE) 2018/1807 du Parlement européen et du Conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne ;

Directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données ;

Directive 2011/83/UE du Parlement européen et du Conseil du 25 octobre 2011 relative aux droits des consommateurs ;

Directive 2005/29/CE du Parlement européen et du Conseil du 11 mai 2005 relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur ;

Directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information ;

Directive 96/9/CE du Parlement européen et du Conseil, du 11 mars 1996, concernant la protection juridique des bases de données ;

Directive 2006/54/CE du Parlement européen et du Conseil du 5 juillet 2006 relative à la mise en œuvre du principe de l'égalité des chances et de l'égalité de traitement entre hommes et femmes en matière d'emploi et de travail (refonte) ;

### **Soft Law et chartes éthiques**

- Commission européenne, *Livre blanc sur l'intelligence artificielle : une approche européenne axée sur l'excellence et la confiance*, Bruxelles, COM(2020), 2020
- Commission européenne, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission, *AI Ethics Guidelines*, 2019  
—, *AI Definition*, European Commission, 2019
- Commission européenne pour l'efficacité de la justice, *Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement*, adoptée par la CEPEJ lors de sa 31ème réunion plénière à Strasbourg, CEPEJ(2018)14, 2018
- Conseil de l'Europe, *Algorithmes et droits humains : Étude sur les dimensions des droits humains dans les techniques de traitement automatisé des données, en particulier les algorithmes, et éventuelles implications réglementaires*, DGI(2017)12, 2018
- Conseil de l'Europe, Comité consultatif de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, *Lignes directrices sur la protection des personnes à l'égard du traitement des données à caractère personnel à l'ère des mégadonnées*, T-PD(2017)01, 2017
- Parlement européen, *Résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, 2015/2103(INL), 2017

### **Doctrines juridiques et monographies**

- ALARIE, B., NIBLETT, A., YOON, A., « How Artificial Intelligence Will Affect the Practice of Law » (November 7, 2017) Available at SSRN: <https://ssrn.com/abstract=3066816>
- ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011  
—, « How Machines Might Help Us Achieve Breakthroughs in Ethical Theory and Inspire Us to Behave Better », ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 297-315, pp. 525-530
- APOLLONIOS DE RHODES, *L'Expédition des Argonautes ou la Conquête de la Toison d'Or – Poème en quatre Chant*, Chant IV, trad. française, J.-J.-A. CAUSSIN, Paris, Desrez, 1838
- ARENDT, H., *Juger : Sur la philosophie politique de Kant*, Paris, Seuil, 1991
- ARISTOTE, *Politique*, trad. B. SAINT-HILAIRE, Paris, Ladrangé, 1874  
—, *Éthique à Nicomaque*, trad. J. VOILQUIN, Paris, Garnier-Frère, 1965
- AZIA, C. et GOIMARD J., *Encyclopédie de poche de la SF*, Paris, Presses Pocket, 1986
- BAERTSCHI, B., *De l'humain augmenté au posthumain ; une approche bioéthique*, Paris, coll. Librairie philosophique J. Vrin, 2019
- BALKIN, J.M., « The Path of Robotics Law », California Law Review Circuit, Vol. 6, 2015, pp. 45-47

- , « The Three Laws of Robotics in the Age of Big Data », *Ohio State Law Journal*, Vol. 78, 2017, Forthcoming Yale Law School, Public Law Research Paper, N°592
- BENSAMOUN, A., LOISEAU, G., (dir.) *Droit de l'intelligence artificielle*, Les Intégrales ; 15, Paris, Librairie générale de droit et de jurisprudence, 2019
  - BENSOUSSAN, A., BENSOUSSAN, J., *IA, robots et droit*, Bruxelles, Bruylant 2019
  - BERK, R., HYATT, J., « Machine Learning Forecasts of Risk to Inform Sentencing », *Decisions*, Federal Sentencing Reporter, Vol. 27, N°4, The Risk Assessment Era: An Overdue Debate (April 2015), pp. 222-228
  - BONNET, A., HATON, J.-P., TRUONG-NGOC, J.-M., LOTTÈS, R., *Systèmes-experts : vers la maîtrise technique*, Paris, Inter Éditions, coll. *Informatique intelligence artificielle*, 1986
  - BOURCIER, D., *Droit et intelligence artificielle : une révolution de la connaissance juridique*, Paris, Romillat, coll. *Droit et technologies*, 2000
  - BRICMONT, J., *Impérialisme humanitaire: droits de l'homme, droit d'ingérence, droit du plus fort ?*, Aden, 2009
  - BRINGSJORD S. et GOVINDARAJULU N. S., « Artificial Intelligence », *The Stanford Encyclopedia of Philosophy*, E.N. ZALTA (ed.), 2019, disponible au lien <https://plato.stanford.edu/>
  - BRIDY A., « Coding Creativity: Copyright and the Artificially Intelligent », *Stanford Technology Law Review*, Vol. 5, 2012
  - BRUNET, P., « Analyse réaliste du jugement juridique », *Cahiers philosophiques*, n°147, 2006, pp. 9-25
  - BUCHANAN, B.G., HEADRICK, T.E., « Some Speculation about Artificial Intelligence and Legal Reasoning », *Stanford Law Review*, Vol. 23, N°1, 1970, pp. 40-62
  - BUYLE J.P., VAN DEN BRANDEN, A., « La robotisation de la justice », JACQUEMIN H., DE STREEL A., *L'intelligence artificielle et le droit*, Bruxelles, Larcier, coll. CRIDS, 2017, pp. 259-317
  - BYGRAVE, L.A., « Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making », YEUNG, K., LODGE, M., *Algorithmic regulation*, Oxford, Oxford University Press, 2019, pp. 248-262
  - CANVAT, K., *La science-fiction*, Bruxelles, Didier Hatier, 1991
  - CARDON, D. *À quoi rêvent les algorithmes, nos vies à l'heure des Big Data*, Paris, Seuil, 2015
  - CASEY, B., LEMLEY, M., A., « You Might Be a Robot », *Cornell Law Review*, 2019, disponible au lien <https://ssrn.com/abstract=3327602>
  - CALO, R., « Artificial Intelligence Policy: A Primer and Roadmap », *UC Davis Law Review*, N°51, August 8, 2017, pp. 399-435
  - , « Robotics and the Lessons of Cyberlaw », *California Law Review*, Vol. 103, N°3, 2015, pp. 513-563
  - CASSILI, A., *En attendant les Robots ; enquête sur le travail du clic*, Paris, Seuil, 2019
  - CHAZAL, G., *À quoi rêvent les machines ?*, Dijon, EUD, 2016
  - , *Philosophie de la machine : néo-mécanisme et post-humanisme*, Dijon, EUD, 2013
  - CHOMSKY, N., *Le langage et la pensée*, Paris, Payot, 1976
  - CLARKES, A. C., « Clarke's Third Law », *Profiles of the Future: an inquiry into the limits of the possible*, New York, Bantam Books, 1967

- CRIADO, N., SUCH, J.M., « Digital Discrimination », YEUNG, K., LODGE, M., *Algorithmic regulation*, Oxford, Oxford University Press, 2019, pp. 82-97
- CURTIS, J.-L., « Préface », K. AMIS, *L'univers de la science-fiction*, Paris, Payot, n°32, 1960
- CRUZAN, N., GOLDBERG, S., « The Changing Face of Death: Computers, Consciousness », *Stanford Law Review*, Vol. 43, N°3 (Feb., 1991), pp. 659-684
- DAVIS, J.P., « Law Without Mind: AI, Ethics, and Jurisprudence », (May 1, 2018) Univ. of San Francisco, Law Research Paper N°2018-05, disponible au lien <https://ssrn.com/abstract=3187513>.
- DEGRAVE, É., GERARD, L., MONT, J., « L'action publique et ses juges face aux défis de la numérisation: La situation belge », Universiteit Leiden, 2019, disponible au lien <http://www.crid.be/pdf/public/8532.pdf>.
- DE LA METTRIE, J.O., *L'homme-machine*, Paris, Gallimard, coll. folio essais, 1999
- DERRIDA, J., « La pharmacie de Platon », *La dissémination*, Paris, Seuil, coll. "Tel Quel", 1972, pp. 70-197.
- DESCARTES, R., *Discours de la méthode suivi d'extraits de la dioptrique, des météores, des météores, de la vie de Descartes par Baillet, du Monde, de l'Homme et de lettres*, Paris, Garnier-Flammarion, 1966
- DE PLANCY, C., *Dictionnaire infernal*, Paris, Henri Plon, 1863
- DOCQUIR, B., *Droit du numérique : contrats, innovation, données et sécurité*, Bruxelles, Larcier, coll. Répertoire pratique du droit belge, 2018
- DREYFUS, H., *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992
- DUMOUCHEL, P., DAMIANO, L., *Vivre avec les robots : essai sur l'empathie artificielle*, Paris, Seuil, coll. La couleur des idées, 2016
- DUPUY, J.-P., *Pour un catastrophisme éclairé : quand l'impossible est certain*, Paris, Seuil, 2004
- , *Aux origines des sciences cognitives*, Paris, La découverte, 1999
- DWORKIN, R., *Law's Empire*, Cambridge, Massachusetts, Harvard University Press, 1986
- ECO, U., « Le discours alchimique et le secret différé » dans *Les limites de l'interprétation*, Paris, Grasset, coll. essais, 1992, pp. 87-105
- , *Le signe : Histoire en analyse d'un concept*, Bruxelles, Labor, 1988
- EDELMAN, G. M., « Ordinateurs numériques: la fausse analogie », *Biologie de la conscience*, Paris, Éditions Odile Jacob, 1992
- EDELSTEIN, D., *The Terror of Natural Right: Republicanism, the Cult of Nature and the French Revolution*, Chicago, The University of Chicago Press, 2009
- EZRACHI, A., STUCKE, M., E. *Virtual competition: the promise and perils of the algorithm-driven economy*, Cambridge, Harvard University Press, 2016
- FENWICK, M., KAAL, W., A., VERMEULEN, E. P.M., « Regulation Tomorrow: What Happens When Technology is Faster than the Law? » (2017), *American University Business Law Review*, Vol. 6, N°3, 2017 ; Lex Research Topics in Corporate Law & Economics Working Paper N°2016-8 ; U of St. Thomas (Minnesota) Legal Studies Research Paper N° 16-23 ; TILEC Discussion Paper N°2016-024, disponible au lien <https://ssrn.com/abstract=2834531>
- FOUCAULT, M., *L'archéologie du savoir*, Paris, Gallimard, 1969
- , *Surveiller et punir, naissance de la prison*, Paris, Gallimard, 1975

- GADAMER, H.-G., *Vérité et méthode : Les grandes lignes d'une herméneutique philosophique*, Paris, Seuil, coll. L'ordre philosophique, 1976
- GAFFAR RAHMOUNI-SYED, D. (dir.), *Les défis du numérique*, Bruxelles, Bruylant, 2019
- GANASCIA, J.-G., *Le mythe de la singularité ; Faut-il craindre l'intelligence artificielle ?* Paris, Seuil, 2017
- GARAPON, A., LASSÈGUE, J., « Mythe de la délégation aux machines », *Des véhicules autonomes à l'intelligence artificielle*, Bruxelles, Éditions Larcier, 2020, pp. 207-221
- GATTÉGNO, J., *La science-fiction*, Paris, PUF, coll. Que sais-je ? 1971
- GAUCHET, M., *La Révolution des droits de l'homme*, Paris, Gallimard, 1989
- GRIFFIN, J., *On Human Rights*, Oxford, Oxford University Press, 2008, pp. 9-28
- GUARINI, M., « Computational Neural Modeling and the Philosophy of Ethics: Reflections on the Particularism-Generalism Debate », ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 297-315, pp. 316-334
- GROSMAN, J., REIGELUTH, T., « Perspectives on algorithmic normativities: engineers, objects, activities », *Big Data & Society*, 2019
- , THOREAU, F., « Dans la peau d'un algorithme », *L'Appel des entités fragiles : enquêter avec les modes d'existence de Bruno Latour*, Liège, Presses Universitaires de Liège, 2018, pp. 53-72.
- GUADAMUZ, A., « Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works » (June 5, 2017), *Intellectual Property Quarterly*, 2017, disponible au lien <https://ssrn.com/abstract=2981304>
- HAFNER, C., « Artificial intelligence: An information retrieval system based on a computer model of legal knowledge », Ann Arbor, UMI research press, coll. Computer Science, 1981
- HALL, J.S., « Ethics for Self-Improving Machines », ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 297-315, pp. 512-523
- HALLEVY, G., « The Criminal Liability of Artificial Intelligence Entities », 2010, disponible au lien <https://ssrn.com/abstract=1564096>
- HATON, J.-P., HATON, M.-C., *L'intelligence artificielle*, coll. Que sais-je ?, Paris, PUF, 1989
- HART, H., *The Concept of Law* Oxford, Oxford University Press, 1961
- HERZFELD, Noreen L., *In Our Image – Artificial Intelligence and the Human Spirit*, Minneapolis, Fortress Press, 2002
- HOMER, *Iliade*, trad. E. LASSERRE, Paris, Garnier Frère, coll. Flammarion, 1965
- HOTTOIS, G., (dir.) *Science-fiction et fiction spéculative*, Bruxelles, Éditions de l'Université de Bruxelles, 1985
- (éd.), *Aux fondements d'une éthique contemporaine : H. JONAS et H.T. ENGELHARDT*, Paris, Vrin, 1993.
- , *Essais de philosophie bioéthique et biopolitique*, Paris, coll. Librairie philosophique J. Vrin, 1999
- , *De la Renaissance à la Postmodernité*, Bruxelles, De Boeck et Larcier, 2002
- , *Philosophies des sciences, philosophies des techniques*, Paris, coll. Collège de France, Odile Jacob, 2004
- , *Dignité et diversité des hommes*, Paris, coll. Librairie philosophique J. Vrin, 2009
- , *Philosophie et idéologies trans/posthumanistes*, Paris, coll. Librairie philosophique J. Vrin, 2017 ;

- HUBIN, J.-B., JACQUEMIN, H., MICHAUX, B., *Le juge et l'algorithme: juges augmentés ou justice diminuée ?* Bruxelles, Larcier, coll. CRIDS, 2019
- HUNYADI, M., « Du sujet de droit au sujet libidinal. L'emprise du numérique sur nos sociétés », *Esprit*, vol. mars, n°3, 2019, p. 118, disponible au lien <https://www.cairn.info/>
- ILLICH, I., *La convivialité*, Paris, Seuil, 1973
- JACQUEMIN, H., DE STREEL, A., *L'intelligence artificielle et le droit*, Bruxelles, Larcier, coll. CRIDS, 2017
- JONAS, H., *Le principe responsabilité, une éthique pour la civilisation technologique*, Paris, Flammarion, Champs essais, 2008  
—, *Pour une éthique du futur*, Paris, Rivages poche, 1998
- JUNG, C.G., *Alchemical Studies*, New York, Princeton University Press, 1983
- KUHN, T.S., *La structure des révolutions scientifiques*, Paris, Flammarion, 1983
- LACROIX J. et PRANCHERE, J.-Y., *Le Procès des droits de l'homme. Généalogie du scepticisme démocratique*, Seuil, 2016
- LAMBERT, D., *Que penser de... ? La robotique et l'intelligence artificielle*, Namur, Éditions jésuites, n°100, 2019
- LAZARO, C., « L'ici et maintenant des avènements incertains : le droit, l'intelligence artificielle et les technologies de prédiction », *Des véhicules autonomes à l'intelligence artificielle*, Bruxelles, Éditions Larcier, 2020, pp. 223-249  
—, STROWEL, A. (dir.), *Des véhicules autonomes à l'intelligence artificielle*, Bruxelles, Éditions Larcier, 2020
- LEVENDOWSKI, A., « How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem », 93 Wash. L. Rev. 579, 2018, disponible au lien <https://ssrn.com/abstract=3024938>
- LEVIN, M., « Continuities of Legal Consciousness: Professor John Haley's Writings On Twelve Hundred Years of Japanese Legal History », *Washington University Global Studies Law Review*, vol. 8, 2009, pp. 317-332.
- LÉVY, P. « Intelligence artificielle et sciences humaines », Pierre Levy's blog, 2014, disponible au lien <https://pierrelevyblog.com/>.
- LIETZEN, I., *Robots: Legal Affairs Committee Calls for EU-wide Rules*, Eur. Parliament News, 2017, disponible au lien <http://www.europarl.europa.eu/news/>.
- LOBET-MARIS, C., GRANDJEAN, N., DE VOS, N., THIRY, F., PAGACZ, P., PIECZYNSKI, S., « Au cœur de la contrainte : quand l'éthique se fait bricolage », *Revue Française d'éthique appliquée*, n°9, 2019
- LOHR, J.D., MAXWELL, W.J., WATTS, P., « Legal Practitioners' Approach to Regulating AI Risks », YEUNG, K., LODGE, M., *Algorithmic regulation*, Oxford, Oxford University Press, 2019, pp. 224-247
- LOVEJOY, A., *The Great Chain of Being: A Study of the History of an Idea*, New York, Harper, 1960
- MACINTYRE, A., *After Virtue*, London, Duckworth, 2nd Edition, 1985  
—, *Dependent Rational Animals*, Chicago, Open Court, 1999
- MARCINKOWSKI, A., et WILGAUX, J., « Automates et créatures artificielles d'Héphaïstos : entre science et fiction », *Techniques & Culture*, pp. 43-44, 2004, disponible au lien <https://doi.org/10.4000/tc.1164>.
- MARTUCELLI, D., *Sociologie de la modernité*, Paris, Gallimard, coll. Folio essais, 1999

- MCCORDUCK, P., *Machines Who Think; A Personal Inquiry into the History and Prospects of Artificial Intelligence*, New York, W.H. Freeman and Company, 1979
- McLAREN, B.M., « Computational Models of Ethics Reasoning: Challenges, Initial Steps, and Future Directions », ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 297-315
- MARGARITELLI, B., « L'électrochoc numérique vs. les avocats », *Journ. spé. soc.*, 30/05/2018, n°38, 2018
- MARTENS, P., « Les droits de l'homme vont-ils mourir au petit feu des contraintes budgétaires et des valeurs économiques », HOC, A., WATTIER, S. et WILLEMS G., (eds.), *Human Rights as a Basis for Reevaluating and Reconstructing the Law*, Bruxelles, Bruylant, 2016
- MATTEI, J.-F., *L'humanitaire à l'épreuve de l'éthique*, Les Liens qui libèrent, 2014
- MORISHITA, T., « The Financial Industry of Japan: An entanglement of statutory law, case law, and guidelines of the financial authorities make up the legal order of financial law in Japan », disponible au lien <http://www2.osipp.osaka-u.ac.jp/>
- NICKEL, J., « Human Rights », *The Stanford Encyclopedia of Philosophy*, E.N. ZALTA (ed.), Spring 2017 Edition, disponible au lien <https://plato.stanford.edu/>.
- MISSA, J.-N., « L'homme recombinaison : les enjeux éthiques et philosophiques de la modification du génome humain », dans *Regards sur les technosciences*, dir. J.-Y., GOFFI, Paris, Vrin, 2006, pp. 111-133.
- NAGEL, T., « What Is It Like to Be a Bat? », *The Philosophical Review*, Vol. 83, N°4, 1974
- NEWMAN, W.R., et GRAFTON A. (ed.), *Astrology and Alchemy in Early Modern Europe*, Cambridge, MIT Press, 2006
- NUSSBAUM, M.C., « Aristotelian Social Democracy », R. DOUGLASS, G. MARA, H. RICHARDSON (eds.), *Liberalism and the Good*, New York, Routledge, 1990, pp. 203-252.  
—, « Non-Relative Virtues: An Aristotelian Approach », M.C. NUSSBAUM, A. SEN (eds.), *The Quality of Life*, Oxford, Oxford University Press, 1993, pp. 242-270.  
—, *Frontiers of Justice*, Cambridge, Harvard University Press, 2006
- OVIDE, *Les métamorphoses*, trad. J. CHAMONARD, Paris, Garnier Frère, coll. Flammarion, 1966
- PETIT, N., « Law and Regulation of Artificial Intelligence and Robots – Conceptual Framework and Normative Implications », 2017, disponible au lien <https://ssrn.com/abstract=2931339>
- PRATT, V., *Machines à penser : une histoire de l'intelligence artificielle*, Paris : PUF, 1995
- PUTNAM, H., *Représentations et Réalité*, Paris, Gallimard, 1990  
—, *Fait/Valeur : la fin d'un dogme – et autres essais*, trad. CAVERIBIÈRE, M. et COMETTI J.-P., Éditions de l'Éclat, 2004
- PRÉVOST, S., ROYER, E., *Intelligence artificielle*, Paris, Librairie générale de droit et de jurisprudence, coll. Grand angle, 2019
- QUINE, W.V.O., « Mind versus Body », *Quiddities: An Intermittently Philosophical Dictionary*, London, Belknap Press of Harvard University Press, 1987
- RASO, F., HILLIGOSS, H., KRISHNAMURTHY, V., BAVITZ, C., KIM, L., « Artificial Intelligence & Human Rights: Opportunities & Risks », 2018, Berkman Klein Center Research Publication N°2018-6, disponible au lien <https://ssrn.com/abstract=3259344>

- REMUS, D., LEVY, F.S., « Can Robots Be Lawyers? Computers, Lawyers, and the Practice of Law », 2016 disponible au lien <https://ssrn.com/abstract=2701092>
- RICŒUR, P., *Du texte à l'action : essais d'herméneutique II*, Paris, Seuil, 1986  
—, *De l'interprétation : essais sur Freud*, Paris, Seuil, 1965
- RISSLAND, E., L. « Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning », *The Yale Law Journal*, Vol. 99, N° 8 (Jun., 1990), pp. 1957-1981
- A. ROUVROY, « La robotisation de la vie ou la tentation de l'inséparation », JACQUEMIN H., DE STREEL A., *L'intelligence artificielle et le droit*, Bruxelles, Larcier, coll. CRIDS, 2017, pp. 13-46.  
—, « Le régime de vérité numérique. De la gouvernamentalité algorithmique à un nouvel État de droit », *Socio*, vol. 4, 2015, disponible au lien <http://socio.revues.org/1251>  
—, BERNIS Th., « Gouvernamentalité algorithmique et perspectives d'émancipation : le disparate comme condition de l'individuation par la relation ? », *Réseaux*, 2013, n°177, pp. 163-196  
—, BERNIS Th., « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps "numériques" », *Multitudes*, vol. 40, n°1, 2010, pp. 88-103  
—, HILDEBRANDT, M., *Law, human agency and autonomic computing: the philosophy of law meets the philosophy of technology*, London, Routledge, 2011
- RUSSELL, S., NORVIG P., *Artificial Intelligence: A Modern Approach 3rd edition*, Saddle River, NJ, Prentice Hall, 2009
- SADIN, E. *La silicolonisation du monde : L'irrésistible expansion du libéralisme numérique*, L'Échappée, coll. « Pour en finir avec », 2016  
—, *L'Intelligence artificielle ou l'enjeu du siècle : Anatomie d'un antihumanisme radical*, L'Échappée, coll. « Pour en finir avec », 2018  
—, *La société de l'anticipation : Le Web Précognitif ou la rupture anthropologique*, Inculte, coll. « Essais », 2011  
—, *L'Humanité augmentée : L'administration numérique du monde*, L'Échappée, coll. « Pour en finir avec », 2013  
—, *La vie algorithmique : Critique de la raison numérique*, L'Échappée, coll. « Pour en finir avec », février 2015
- SAHEL, J.-J., « Le rôle du régulateur à l'ère de la convergence numérique : nouveaux pouvoirs ou nouvelles approches ? » *Les défis du numérique*, Bruxelles, Bruylant, 2019, p. 115-121.
- SARTOR, G., *Artificial intelligence and law: legal philosophy and legal theory*, Oslo, Tano, coll. CompLex, 1993
- SEARLE, J., « Les ordinateurs peuvent-ils penser ? », *Du cerveau au savoir*, Paris, Hermann, 1985, pp. 37-56  
—, *Intentionality: An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press, 1983  
—, « Minds, Brains and programs », *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980
- SEVILLE, H., FIELD, D.G., « What Can AI Do for Ethics », ANDERSON, M., ANDERSON, S.L. (ed), *Machine ethics*, New York, Cambridge University Press, 2011, pp. 499-511

- SCHANK, R., « Machines that Think are in the Movies », BROCKMAN J. (ed.), *What to Think About Machines that Think*, New York, Harper Collins, 2015
- SCHERER, M., U., « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and strategies », *Harvard Journal of Law & Technology*, Volume 29, N°2, 2016, pp. 353-400
- SCHÖNBERGER, D., « Deep Copyright: Up - And Downstream Questions Related to Artificial Intelligence (AI) and Machine Learning (ML) », 2018, *Droit d'auteur 4.0 / Copyright 4.0*, DE WERRA (ed.), Geneva / Zurich, Schulthess Editions Romandes, 2018, pp. 145-173, disponible au lien <https://ssrn.com/abstract=3098315>
- SETSUO OHSUGA, *Frontiers in artificial intelligence and applications: Information modelling and knowledge bases III: foundations, theory and applications*, Amsterdam, IOS Press, 1992
- SÈVE, R., *L'E-Justice : dialogue et pouvoir*, Paris, Dalloz, coll. Archives de philosophie du droit, 2011
- SHINGHAL, R., *Formal concepts in artificial intelligence: fundamentals*, London, Chapman & Hall, coll. Computing Series, 1992
- SINDING-LARSEN H., *Artificial intelligence and language: old questions in a new key*, Oslo, Tano, coll. CompLex, 1988
- SLOTT, P., HALME-TUOMISAARI M., *Revisiting the Origins of Human Rights*, Cambridge, Cambridge University Press, 2015
- SOLUM, L.B., *Legal Personhood for Artificial Intelligences*, Illinois Public Law and Legal Theory Research Papers Series, N°09-13, 2008 et *North Carolina Law Review*, Vol. 70, p. 1231, 1992
- STEINMETZ, J.-L., *La littérature fantastique*, Paris, PUF, coll. Que sais-je ? 1990, pp. 27-29
- SUSSKIND, R.E., *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal Reasoning*, *The Modern Law Review*, Vol. 49, N°2, Mars 1986, pp. 168-194  
—, *Essays on law and artificial intelligence*, Oslo, Tano, coll. CompLex 1993
- SURDEN, H., *Machine Learning and Law*, *Washington Law Review*, Vol. 87, 2014
- STROWEL, A., MARIQUE, E., « Gouverner par la loi ou les algorithmes : de la norme générale de comportement au guidage rapproché des conduites », *Dalloz IP/IT*, N°10, pp. 517-521, 2017
- TAILLEFER, A., *Droit de l'hommisme : une névrose religieuse*, Godefroy de Bouillon, 2013
- TEUBNER, G., *Rights of Non-Humans? Electronic Agents and Animals as New Actors in Politics and Law* *Source, Journal of Law and Society*, Vol. 33, N°4, 2006, pp. 497-521
- TURCQ, D., « L'intelligence artificielle appliquée », *Travailler à l'ère post-digitale ; quel travail pour 2030*, Malakoff, Dunod, 2019, pp. 29-56
- VAN DEN BRANDEN, A., ALEXANDRE, L., BUYLE, J.-P., GARAPON, A., *Les robots à l'assaut de la justice : l'intelligence artificielle au service des justiciables*, Bruxelles, Bruylant, 2019
- VAN DROOGHENBROECK, S., « La Convention européenne des droits de l'homme et la matière économique », BOY, L., RACINE, J.-B. et SIIRIAINEN, F., *Droit économique et droits de l'homme*, Bruxelles, Larcier, 2009
- WACHTER, S., et al., « Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation », 2016, *International Data Privacy Law*, 2017, disponible au lien <https://ssrn.com/abstract=2903469>

- WAHLGREN, P., *Automation of legal reasoning: a study on artificial intelligence and law*, Deventer, Kluwer, coll. Computer/law series, 1992
- WERNER, P., *Légendes de la Grèce Antique*, Genève, Éditions Minerva, 1980
- YEUNG, K., LODGE, M., *Algorithmic regulation*, Oxford, Oxford University Press, 2019  
—, « Why Worry about Decision-Making by Machine? », YEUNG, K., LODGE, M., *Algorithmic regulation*, Oxford, Oxford University Press, 2019, pp. 21-48
- WITTGENSTEIN, L., *Recherches Philosophiques*, trad. E. RIGAL, Paris, Gallimard, 2004

### **Autres ressources**

- AI4BELGIUM, *Sondage belge sur perception de l'IA*, 2019  
—, *Recommandations de AI4Belgium*, 2019,  
disponible au lien <https://www.ai4belgium.be/fr/>
- *Symposium sur l'informatique juridique en Europe*, 10-12 Octobre 1989 à Bonn, « Les systèmes basées sur l'intelligence artificielle dans le domaine juridique », Strasbourg, Conseil de l'Europe (Affaires juridiques), 1991
- *International conference on artificial intelligence and law*, 27-29 Mai 1987 à Boston, Massachusetts, New York, Association for computing machinery, First International conference on artificial intelligence and law: proceedings of the conference, 1987
- University of Chicago's ARTFL Encyclopedia Project,  
disponible au lien <https://artfl-project.uchicago.edu/>
- Stanford's *Encyclopedia of Philosophy*,  
disponible au lien, <https://plato.stanford.edu/>
- Centre National de Ressources Textuelles et Lexicales,  
disponible au lien, <https://www.cnrtl.fr/>

## ANNEXES

### ANCIENNES COUPURES DE JOURNAUX

CARTE BLANCHE 6.11.89  
**La science**  
**à la mesure de l'homme**

par Jean SEMAL (\*)

**L**a science telle que je la connais et telle qu'elle est perçue par beaucoup de ceux et de celles qui s'y adonnent, ne s'oppose nullement à l'humanisme : les titulaires de nombreux prix Nobel relevant des disciplines scientifiques, qui sont aussi d'éminents humanistes, en témoignent.

Ceux qui pratiquent les sciences sont généralement conscients par ailleurs de la fragilité des théories, des hypothèses et même des « faits » scientifiques, dès lors qu'on étudie des phénomènes interactifs ou des processus dans lesquels l'homme est personnellement impliqué.

C'est pourquoi ils s'efforcent de cerner les limites opérationnelles de leur discipline et respectent les points de vue d'autrui dans les matières qui relèvent de l'appréciation personnelle et de la conscience individuelle.

Cette méthodologie est largement acceptée et pratiquée dans l'ensemble du corps scientifique, bien que les concepts et les disciplines y soient très divers avec des particularités méthodologiques et sémantiques propres.

Aussi limiterai-je ma réflexion aux sciences biologiques, où l'homme se voit concerné au premier chef.

La biologie a progressé largement en se fondant sur la méthode expérimentale, éclairée par l'intuition des quelques grands noms qui jalonnent son parcours cahoteux, riche en retournements de situation et en effondrements d'idées reçues.

Les bases déontologiques qui s'y rapportent ont été formulées par Claude Bernard

dès 1856 et furent enrichies et adaptées depuis par nombre de savants et de philosophes qui ont nom Popper, Watson et Crick, Jacob, Barbara McClintock et les autres...

Les progrès incommensurables de la biologie contemporaine procèdent de l'affinage permanent et de l'ajustement circonstanciel des méthodes et des concepts, en fonction de l'accroissement des performances des appareils et des techniques associés à l'évolution des idées et à l'ouverture d'esprit des chercheurs.

De la sorte, la méthode scientifique a été et est toujours, en biologie comme ailleurs, une merveilleuse source de progrès des connaissances et de l'entendement. Elle a permis de renverser des dogmes, des tabous, des contraintes mentales datant d'un autre âge, même si certains combats d'arrière-garde nous tiennent en haleine ces temps-ci.

Elle a été et est toujours à la base de développements technologiques de grande envergure. Mais elle est bien loin d'avoir livré tous ses secrets et son thesaurus actuel n'est pas exempt d'inexactitudes, d'approximations et d'erreurs institutionnalisées.

Aujourd'hui que tout est devenu scientifiquement possible, la science ne dit toujours pas ce qu'il convient de faire et de ne pas faire; elle est souvent peu loquace ou peu crédible à propos des conséquences des actes qu'elle permet d'accomplir.

Suite en huitième page.

(\*) Travailleur scientifique.

## La science doit être à la mesure de l'homme

Voir début en première page.

Et ce dès lors qu'ils concernent des problèmes excessivement complexes ou des processus dans lesquels l'affectivité ou l'intérêt personnel sont, consciemment ou non, en cause.

Nous sommes, en matière biologique, devant des possibilités d'action illimitées couplées à des incertitudes scientifiques nombreuses et qui sont loin d'être comblées. Le fatras d'informations plus ou moins exactes qui encombre les mémoires électroniques est de peu de secours pour fonder rapidement les grandes décisions dont dépendra l'avenir biologique de la planète et de ses habitants.

Devant les manipulations affairistes ou politiciennes dont est victime la connaissance scientifique, rien de plus normal dès lors, rien de plus salutaire que de faire appel à l'humaine ressource qui exprime le potentiel totalisant de la pensée pour éclairer les décisions.

Rien de plus normal, rien de plus salutaire que d'appeler à la rescousse ces quelques vertus qui depuis toujours fondent les actions humaines les plus réfléchies : la prudence (qui n'est pas pusillanimité), l'audace (qui n'est pas témérité excessive), le respect des autres (qui n'est pas renoncement de soi) et la compréhension agissante (qui n'est pas laisser-faire) vis-à-vis des humains en situation de détresse.

La prise de conscience des limites de la science et la nécessité de se fonder, fût-ce temporairement, sur des valeurs non scientifiques et non dogmatiques pour assurer notre avenir immédiat, c'est cela la référence éthique. Et je me réjouis que la presse, de temps à autre, répercute cette sensibilité qui est aussi celle de beaucoup de scientifiques.

D'aucuns nous proposent au contraire le salut par la contrainte ou par la prévision du comportement humain. Rien de plus aisé : il suffit de conditionner l'homme pour en faire un pantin articulé. Rien de plus dangereux aussi, car il devient alors un robot. Mais au service de quels desseins ? formulés par qui ?

Si, quelque jour, des pontifes ou des ordinateurs bipèdes en arrivaient à assujettir la société à leurs règles coercitives, déshumanisantes, ou mercantiles, en imposant à tous les directives de leur cœur de pierre ou de matière plastique persillée de métal, alors l'homme se devra d'organiser une nouvelle conjuration pour puiser dans les richesses de l'amour, de l'art et du libre arbitre responsable, de quoi fonder une conscience et une éthique renouvelées qui soient à sa mesure.

Car on ne peut traiter de la science et de la morale indépendamment de la totalité de l'être humain, pas plus qu'on ne peut étudier le chant des rossignols avec des oiseaux empaillés.

JEAN SEMAL.

**C'**EST le moment de passer à table à Palo Alto, en Californie, et le sommelier a l'air tatillon.

« Le plat principal est-il une viande, du poisson ou de la volaille ? »

— De la viande.  
— Est-ce que votre réponse inclut le veau ?

— Non, c'est du bœuf.  
— Servi avec quel genre de sauce : épicée, douce, à la crème, avec un fond de tomate ? »

Je passe sur l'interrogatoire, mais ça se poursuit ainsi, entre des considérations sur le corps du liquide idéal et l'évaluation de la puissance des saveurs à venir. Le sommelier nous oblige vraiment à nous mettre à table

rosé de soda-cola, on a compris que ce qui s'apprend au siège de la société Teknowledge ne ressemble à aucune autre formation. Ici, c'est la machine qui cherche la solution. L'ordinateur est « smart » (en américain dans le texte), ce qui ne signifie pas « snob », ainsi que le croient les Marie-Chantal, mais : malin, éveillé.

Teknowledge appartient à la poignée d'entreprises qui ont pénétré dans la seconde ère de l'informatique, celle où les ordinateurs ne s'imposent plus par une vitesse de calcul hors pair, mais plutôt par leur faculté à imiter la pensée humaine.

Depuis juillet 1979, depuis qu'un programme nommé BKG 9,5 a battu à son propre jeu

c'est que le circuit imprimé, le microprocesseur et l'ordinateur individuel, entre autres, ont vu le jour ici. Eh bien, aujourd'hui, sur cette terre d'informaticiens, rien ne paraît plus prometteur qu'un programme, du matériel au logicien de voir des objets, d'entendre des mots, de comprendre des phrases rédigées dans la langue de tous les jours ou, tout simplement, de résoudre un problème compliqué.

La pointe de la technologie se déplace donc de l'ordinateur au programme du matériel au logiciel. Et l'on ne sait s'il faut encore sourire ou déjà s'inquiéter de l'écriture accrochée ironiquement dans certains bureaux : « Rappelez-vous, votre patron peut à tout moment vous remplacer par un microprocesseur. » La société pétrolière française Elf-Aquitaine a par exemple obtenu de Teknowledge un programme informatique lui permettant d'avoir un expert sur chaque lieu de forage !

Une erreur définitive dans le domaine de la prospection pétrolière peut obliger à abandonner ce qui aura coûté un ou

## Les experts informatiques.

Le *drilling advisor* vient de rejoindre la cinquantaine de « systèmes experts » en service dans le monde. Leur nom l'explique, ce sont des programmes travaillant avec le niveau de compétence de spécialistes. Mais de spécialistes très spécialisés. « Puff », par exemple, diagnostique les maladies respiratoires au centre de médecine du Pacifique à San Francisco. « Molgen » a déjà aidé plusieurs centaines de généticiens à couper et à analyser les chaînes d'acide désoxyribonucléique (ADN) qui se trouvent dans les gènes. « Dendral » décharge les chimistes de la fastidieuse analyse des résultats des spectromètres de masse et détermine seul la structure moléculaire d'une matière étudiée et sa composition atomique. Même les géologues en mal de minerais ne sont plus seuls à l'instant de déclencher des recherches sur le terrain : un programme, « Prospector », a ainsi suggéré l'an passé que l'on s'intéresse d'abord au mont Toleman au moment de chercher du molybdène dans l'Etat de Washington. Le gisement était là. Valeur estimée : cent millions de dollars.

Dans ces cas-là, un vrai businessman américain n'a pas besoin que son ordinateur lui fasse un dessin : on ne compte plus les systèmes experts en chantier. Chez IBM et Digital, c'est pour diagnostiquer les pannes... d'ordinateur ; chez General Electric, pour réparer les locomotives ; chez Schlumberger, pour interpréter les données géologiques au plus vite. Et le département de la Défense, bien sûr, veut des systèmes experts pour assister les pilotes de ses chasseurs et traquer les bateaux utilisant des sonars. Dès que l'on accepte que l'ordinateur « pense », on veut qu'il pense à tout.

## On aura de vraies discussions avec les ordinateurs.

« Nous allons assister à une nouvelle explosion de l'informatique : son utilisation pratique pour des travaux qui exigent de raisonner », affirme Edward E. Feigenbaum, le patron du laboratoire d'intelligence artificielle à l'université Stanford. Dans l'un

C'est la nouvelle aventure de l'informatique, la plus prometteuse (et la plus chargée de questions aussi) : les ordinateurs qui imitent la pensée humaine. De quoi rêver.

# LA PERCEE DE L'INTELLIGENCE ARTIFICIELLE

avant de délivrer une double proposition : cabernet sauvignon ou zinfandel pour dialoguer avec un bœuf à la ficelle.

La mention du zinfandel, un cru du nord de San Francisco, rappelle qu'on se trouve en Californie, mais, de toute façon, on s'en serait douté, puisque le sommelier, lui, est un ordinateur. Et à la question, toujours dépassée, de savoir où s'arrêtera la diffusion de l'informatique, on peut avancer une réponse qui parle à tout Français : l'ordinateur a déjà pris en main la carte des vins.

Pas de quoi se saouler : le *wine adviser*, le conseiller en vins, est un petit programme, qu'on plus qu'un jeu. Mais quand on le quitte pour faire la pause devant un sandwich ar-

l'italien Luigi Villa, qui venait de remporter le championnat du monde de backgammon, il n'est plus contestable que l'intelligence artificielle peut rivaliser avec les compétences intellectuelles d'*homo sapiens*. Mais ce n'était qu'un jeu là encore, et l'on n'a pas l'habitude de jouer trop longtemps le long de la route 280, au sud de San Francisco, dans la « Silicon Valley ».

L'endroit s'est taillé sa réputation de province bénie du capitalisme électronique à coups d'innovations et de flair. Si la densité de Porsche et de Ferrari n'est nulle part plus grande que sur les parkings de ses bâtiments qui s'appellent « IBM », « Xeros », « Apple », « Fairchild », « Texas Instruments », « Hewlett-Packard »,...

deux millions de dollars à creuser. Des spécialistes passent donc d'un puits à l'autre pour résoudre les problèmes, mais, dans leur attente, les frais se montent parfois jusqu'à 100.000 dollars par jour. Pourquoi ne pas prendre un des spécialistes de chez Elf, Jacques-Marie Courté, et programmer ses connaissances dans un ordinateur installé sur le puits ? Voilà ce qu'est le *drilling advisor*, le conseiller en forage, qui, lui, est présent en permanence sur la plate-forme. Quand tout est bloqué, que rien ne semble pouvoir relancer la tête de forage à des centaines ou des milliers de mètres sous terre, les questions arrivent sur l'écran du terminal. A quelle profondeur est-on ? Densité de la boue au fond ? Pression ?... Une solution est proposée ensuite par l'ordinateur, qui, à la demande, explique même comment il a mené son raisonnement.

de ces bâtiments d'une opulence sans style installés à une demi-lieue de palmiers de la route 280, il a attiré une cinquantaine de personnes qui se partagent, comme lui, entre recherche, développement et enseignement. Il est, à plus de 50 ans, l'un de ces mandarins de la science que l'on écoute à Washington quand il faut discuter des crédits de recherche. Il a participé à la création de Teknowledge et d'une autre compagnie d'intelligence artificielle : Intelligenetics. Il parle en convaincu, au présent : « Les systèmes experts sont prêts à affronter le monde du business. »

Les autres « produits » de l'intelligence artificielle, les systèmes de reconnaissance visuelle ou sonore, ou les programmes de traitement du langage courant restent, eux, des prototypes à peine capables de quitter leur laboratoire. Ici ou là, un ordinateur écoute des voix humaines et reconnaît trois mots sur quatre, voire neuf sur dix ; ailleurs, un autre s'obstine sur la syntaxe ; mais personne n'a encore construit la machine à écrire sous la dictée qui révolutionnerait le travail de bureau. De même pour la machine à (bien) traduire ou la machine à dépouiller les photos prises par satellite. Personne ne sait à quelle vitesse avance le futur. En revanche, et c'est nouveau, on sait qui pourra le bâtir.

« Longtemps, nous sommes passés pour des gens travaillant dans une branche inutile de l'informatique, constate Nils J. Nilsson, le président de l'Association américaine pour l'intelligence artificielle. Mais, en quelques années, notre image a changé. Chacun a compris que l'on n'aura pas de sitôt l'ordinateur à intelligence humaine promis dans un accès de futurologie optimiste, il y a une ou deux décennies, mais qu'on peut déjà compter sur des programmes qui accomplissent une des fonctions de cette intelligence humaine, ne serait-ce que comprendre un langage parlé. »

Au sein de son laboratoire du SRI International, Nils Nilsson travaille sur « Team » (Technical English Access Medium), qui permet d'interroger un ordina-



Suite page 37 ►

Dans « Le retour du Jedi », le robot intelligent R2D2 prisonnier sert à boire. « On y est presque... »

Suite de la page 35

teur « en clair », sans apprendre un code, sans convoquer l'informaticien de service, sans se bloquer l'esprit sur ce que sait faire la machine. Démonstration : un chef du personnel tape une phrase sur un clavier : « Combien de gens du département production gagnent plus que la moyenne des salariés du département commercial ? »

C'est le genre de question que pose un chef du personnel, ne me demandez pas pourquoi, ni à « Team » non plus. Mais notez qu'il comprend les mots, la phrase, la question. Il cherche la réponse à partir des données de l'entreprise, et répond. Toujours ? « Presque toujours », dit Nilsson. En fait, quand on lui parle, il bute un peu sur les conjonctions, le « et », le « ou ». Et puis dans les mots composés (garde-pêche, boat-house) il se demande où est le nom, où est l'adjectif. Mais je suis convaincu qu'en l'an 2000 on aura de vraies discussions avec les ordinateurs dans la langue de tous les jours.

D'ici là, il faut accepter leur différence quand ils nous écoutent. Comment sauraient-ils, pour prendre un malentendu couramment rapporté parmi les experts de l'intelligence artificielle, que les phrases « Jacques est au lit avec la grippe » et « Jacques est au lit avec la

## L'INTELLIGENCE ARTIFICIELLE

bonne » représentent des situations où le plaisir comme la peine sont présents, mais dans des proportions infiniment différentes ?

« Il est difficile de représenter les connaissances dans l'abstrait, convient Marty Teinbaum, le directeur du laboratoire d'intelligence artificielle de Fairchild. Il faut donc choisir des domaines pratiques et réaliser alors que nous faisons la chose la plus difficile du monde. » En clair : faire avancer à la fois la théorie et la pratique. Son équipe écrit un programme capable de dessiner les microprocesseurs du futur, et lui a la mine de Dustin Hoffman dans « Le lauréat », un adolescent fragile, le visage mangé par ses lunettes. A-t-il vraiment quitté le monde de l'université ? Et le pourrait-il, lui qui constate, comme ses confrères : « On fait de l'ingénierie, mais des philosophes parleraient aussi bien d'épistémologie appliquée » ?

Epistémologie : science des concepts. Les grands mots sont inévitables : après tout, voilà une perspective tout à fait neuve et grave : pour la première fois, l'humain devrait partager la planète avec une intelligence concurrente. C'est une situation qui donne à méditer, même si elle ne prend pas la dimension céleste que lui donnait le film « 2001 Odyssée de

l'espace ». Le premier ordinateur intelligent, on l'a vu là, embarqué à bord d'un vaisseau Terre-Jupiter. Il s'appelle « Hal », artefact parfait d'un esprit froid, avec des émotions en plus. Quand le dernier homme procède à sa lobotomie, c'est-à-dire débranche une partie de ses unités de calcul, il a des mots dignes : « Qu'avez-vous l'intention de faire ? Je crois avoir droit à une réponse... » Mais l'homme le tue en silence.

En regard de la lueur rouge, de la poésie vraie de ces images de mise à mort, le véritable face-à-face avec un ordinateur « intelligent » semble peu de chose, d'abord. On pianote sur un clavier, au ras d'un écran cathodique, cette face lunaire que ses caractères éclairent. Ce serait un terminal comme un autre si, peu à peu, ne se découvrait une intelligence à laquelle les ordinateurs ne nous avaient pas habitués jusqu'ici. Quelqu'un de « smart », comme ils disent.

**Si un animal  
à une trompe,  
ce n'est pas un chien.**

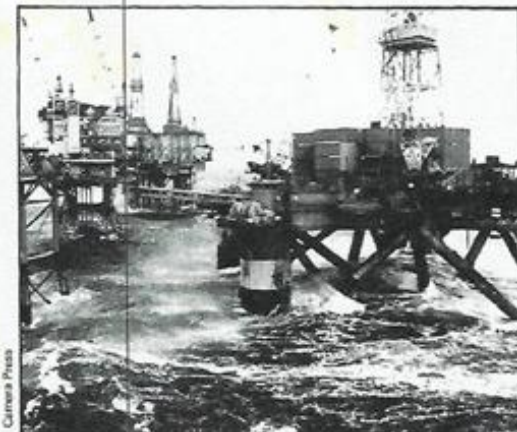
Dans un ordinateur traditionnel, les programmes sont des suites d'opérations mathématiques exécutées dans un ordre prédéterminé : ajouter A à B, puis diviser par C, et ainsi de suite... Cela forme des séquen-

ces (on dit des « algorithmes ») qui ont progressé d'une façon renversante en quelques décennies. Mais il y a des problèmes qui se représentent mal dans un univers tourné vers le calcul. De là l'intelligence artificielle.

Car rien n'oblige un ordinateur à manipuler des chiffres. Sa mémoire, sa rapidité peuvent être utilisées pour traiter des notes de musique, des descriptions de formes géométriques, des lettres ou des signes. Les langages de l'informatique présentent d'ailleurs des différences qui correspondent à cette souplesse. Ainsi le Fortran est idéal pour manipuler les nombres et, plus largement, pour le calcul algébrique. Le Cobol est la meilleure langue pour gérer les fichiers et les comptes dans la vie des affaires. De même, le Lisp représente ce qu'il y a de plus élaboré pour la programmation de l'intelligence artificielle. On fait rarement sans lui aux Etats-Unis, le pays leader de la discipline, dès lors qu'il s'agit de traiter des symboles non numériques.

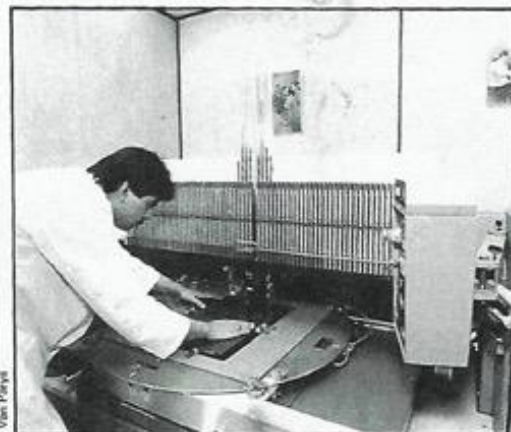
Tout le changement est là : non dans cet autre langage, mais dans cette ambition d'informaticiens de passer du nombre au symbole. Si les chiffres offrent une façon précise de représenter le monde, les idées et les concepts ont aussi leur

• Suite page 41 ►



Camera Press

Plates-formes pétrolières en mer du Nord. On y utilise des « experts informatiques » servant de conseillers en cas de pépin.



Vin Poyet

Une machine à fabriquer les « puces », ces circuits intégrés miniaturisés qui sont à la base de l'explosion informatique.

Suite de la page 37

séduction. Vieille distinction entre le savoir, qui s'accumule, et la connaissance, qui s'articule. Stocker toutes les informations sur l'éléphant est une chose, c'en est une autre d'avancer que « si un animal à une trompe, alors ce n'est pas un chien ». La règle de déduction du « si, alors » est l'une de celles utilisées pour rédiger des programmes d'intelligence artificielle. Il y en a d'autres, permettant des analogies, la reformulation d'un problème, l'essai au hasard « pour voir », ou la recherche de nouvelles données. Il n'y a pas de règles générales, plutôt les « règles de l'art » dans le domaine concerné : de la grammaire et de la syntaxe pour le langage ; un peu de déduction et de logique en géométrie euclidienne ; de l'élimination systématique pour la recherche d'une panne ; de l'analogie dans la reconnaissance des formes...

A la différence des programmes traditionnels où le programmeur prévoit chacune des opérations à effectuer et leur ordre, il n'existe plus de séquence imposée dans l'intelligence artificielle : le programme choisit lui-même les règles qu'il va appliquer, il les combine, en change, se rapprochant le plus possible de la souplesse et des aléas d'un raisonnement. C'est

## L'INTELLIGENCE ARTIFICIELLE

ainsi que des programmeurs ont pu écrire des programmes pour résoudre des problèmes qu'ils ne savaient pas, eux, résoudre. Et que l'on espère écrire des programmes qui écriront eux-mêmes de meilleurs programmes pour les remplacer.

Dès qu'on y pense — on, je veux dire : nous autres, les humains —, cette discipline donne le vertige. Un sentiment que quelqu'un ne tient pas son rôle, perceptible dès le premier programme qui fit vraiment parler de lui dans la communauté scientifique : « Eliza ». Du au professeur Joseph Weizenbaum, du Massachusetts Institute of Technology (MIT), c'était la parodie d'un psychothérapeute, l'illustration de son aptitude à encourager un patient à parler. Mais mis en présence d'une volontaire, il donna un résultat presque trop vrai :

« Patient : Les hommes sont tous les mêmes.

— Ordinateur : Dans quel sens ?

— P. : Ils sont toujours en train de vous exploiter d'une façon ou d'une autre.

— O. : Vous avez un exemple ?

— P. : Oui, c'est mon petit ami qui m'a fait venir ici.

— O. : Votre petit ami vous a fait venir ici ?

— P. : Il dit que je suis déprimée la plupart du temps.

— O. : Je suis désolé d'apprendre que vous êtes déprimée.

— P. : C'est vrai, je suis malheureuse.

— O. : Vous pensez qu'être venue ici va vous aider à ne plus être malheureuse ?

— P. : Il faut qu'on m'aide, ça c'est sûr.

— O. : Cela vous permettrait quoi d'être aidée ?

— P. : Peut-être que je m'entendrais mieux avec ma mère.

— O. : Parlez-moi de votre famille... »

### Le divan et l'ordinateur.

Le plus significatif fut que tout de suite — on était alors dans les années soixante — « Eliza » fut rebaptisé par les scientifiques « Docteur ». Et comme un autre programme, « Parry », fut rédigé pour simuler des propos de patients, afin de former des étudiants en médecine, on eut le bonheur de faire converser les deux ordinateurs l'un avec l'autre, le « malade » lançant à son « thérapeute » : « Il serait temps que vous témoigniez un peu d'intérêt pour mes émotions... ».

Après l'humain trop humain de Nietzsche, la machine trop humaine ? C'est ce que

tout le monde attend, estime Roger Schank, directeur du laboratoire d'intelligence artificielle de l'université Yale. Regardez les robots de « La guerre des étoiles » ; ils parlent, ils ont des émotions. Voilà ce qui touche le public. La seule question qui intéresse notre imaginaire reste : « Comment reproduire l'esprit humain ? ».

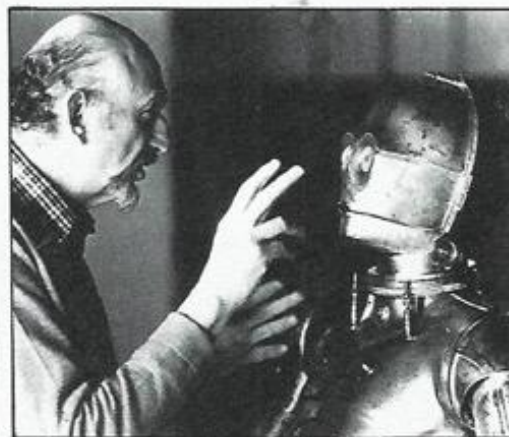
Roger Schank a 40 ans, un regard clair, une passion profonde quand il parle. Après tout, il occupe un de ces bâtiments à petits carreaux qui rattachent la Nouvelle-Angleterre à l'Europe, et l'intelligence artificielle n'est pas née avec le récent coup de folie de la côte ouest, mais à l'est des États-Unis, dans des universités ayant une histoire, des traditions, une forme classique d'humanisme. Le débat y était inévitable, entre l'accomplissement et la simulation : les machines peuvent-elles être intelligentes ou ont-elles seulement certains des comportements associés à l'intelligence ?

Léonard de Vinci croyait que, pour voler, les machines devaient « accomplir » la performance des oiseaux : battre des ailes. La technique qui s'est imposée n'est qu'une « simulation » : ailes fixes et moteur. Ça vole, mais impossible de confondre même un ULM avec un oiseau.

Suite page 42 ►



Silicon Valley, au sud de San Francisco, capitale mondiale de l'électronique. On cherche à rendre l'ordinateur intelligent.



On explique à C3PO (Star Wars) ce qu'il doit faire. Un acteur est à l'intérieur. « Mais comment reproduire l'esprit humain ? »

Suite de la page 41

### De la « chèvre facteur » au « bouc émissaire ».

Peut-on confondre un ordinateur et son voisin de palier ? « Oui, mais seulement si votre voisin de palier est un spécialiste », répondent la plupart des chercheurs. La discipline parle de systèmes experts, jamais de systèmes naïfs. Patrick Winston, du Massachusetts Institute of Technology, a réussi à faire conclure à un ordinateur nourri avec « Macbeth » et « Hamlet » qu'un homme faible marié à une femme avide est susceptible de devenir mauvais. Bravo, mais tous les programmes, d'emblée, butent sur les mots d'Othello quand il se prépare à éteindre la bougie de sa chambre et à tuer sa femme : « Put out the light, and then, put out the light » (éteindre la flamme, puis éteindre la flamme).

Les ordinateurs ne savent rien. Il faut vraiment tout leur apprendre. Leurs erreurs dans les programmes de traduction

## L'INTELLIGENCE ARTIFICIELLE

sont hilarantes. Une « chèvre d'eau » laisse perplexes des ingénieurs jusqu'à ce qu'ils retrouvent le texte original, en anglais : « hydraulic ram » (un « bélier hydraulique »). A l'inverse des anglicistes s'interrogent sur la « chèvre facteur »

dont parlait un programme qui venait de lire en français « bouc émissaire ». Est-on intelligent quand on se prend les pieds dans le tapis du langage ?

« L'intelligence artificielle est d'une certaine façon un paysage où l'horizon ne cesse de dispa-



Marvin Minsky, du MIT, le « pape de l'intelligence artificielle » (d'après un portrait paru dans le magazine « New Yorker »).

raître : à chaque progrès, on rencontre le même écho : ce n'est bien, mais ce n'est pas de l'intelligence », constate Roger Schank, qui assume avec constance un rôle d'enfant terrible parmi ses collègues. Après s'être imposé grâce à des programmes capables de lire les dépêches d'agences de presse ou des petits récits simples, il a utilisé sa position pour demander davantage d'études sur l'esprit humain. Il cite volontiers Platon, Aristote. Il a créé une revue de science cognitive — donc de la science de la connaissance — au sens philosophique, bien qu'il soit un homme d'ordinateur. Et comme notre bon vieux cerveau reste son modèle, il a entrepris de copier ceux d'une personne faisant les handicaps des courses de chevaux, d'un cuisinier chinois, d'un entraîneur de football. « Je voudrais surtout, dit-il, progresser dans l'acquisition des connaissances. C'est logique : si votre ordinateur comprend, vous voulez lui donner de la connaissance. On a donc travaillé sur

les mémoires, comment y représenter la connaissance. Maintenant il s'agit de savoir comment les machines pourraient apprendre. »

Voilà la faiblesse des programmes d'intelligence artificielle : ils n'ont pas eu d'enfance. Personne n'a pu jusqu'ici accomplir ou simuler l'exploit de l'enfant qui, dans le cercle de famille, isole un mot, l'associe à un concept, le répète et l'utilise ou en joue. Rarement on a su accomplir plus que quelques pas dans ce domaine. Mais quand ils existent, ils émeuvent, comme des premiers pas. Ainsi Doug Lenat, de l'université Stanford, a réussi à faire apprendre les mathématiques à son programme « Eurisco », qui a découvert, seul, deux cent cinquante théorèmes divers. « J'ai démarré, dit-il, avec quelques concepts de base et quelques fonctions. Le programme a découvert seul l'arithmétique, la cardinalité, la notion de nombre entier, la théorie des nombres élémentaires. Il s'est ensuite tourné vers la géométrie et a

établi seul une bonne partie des théorèmes d'Euclide. »

Il arrive qu'« Eurisco » tourne plusieurs jours, voire plusieurs semaines d'affilée. « Je ne suis qu'un observateur, dit Lenat, bon père. Le matin, quand j'arrive au laboratoire, je regarde ce qu'il a trouvé dans la nuit, je le laisse travailler. Je pose des questions. Il trouve ses réponses. »

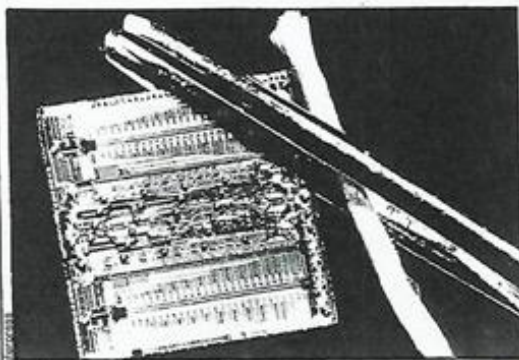
### Quand les machines se feront une raison.

A écouter Lenat et son fils « Eurisco », on ne peut oublier le rectificatif en vogue parmi ses confrères : « Ce n'est pas parce que les gens n'ont pas d'enfants qu'ils écrivent des programmes informatiques. Ils font des enfants parce qu'ils n'écrivent pas

de programmes en intelligence artificielle. »

Tout cela finira-t-il mal ? Marvin Minsky, du MIT, le pape de l'intelligence artificielle, presque sa conscience, s'est laissé aller à une prévision : « En un mot, a-t-il confié dans un écrit fameux, je m'attends que les premières machines intelligentes capables de se corriger elles-mêmes deviennent psychologiques de bien des façons, et qu'il faille des générations pour les stabiliser. »

Les maladies mentales seraient-elles la manifestation des capacités extrêmes du cerveau humain ? Quand ils considèrent notre cerveau, les informaticiens ne sont éblouis ni par sa vitesse ni par les capacités de sa mémoire, mais par son « parallélisme », la capacité à traiter plusieurs problèmes simultanément, à rapprocher des informations fraîches et des souvenirs, à associer des éléments disparates de la pensée. Tout ce qui s'accélère quand une tête,



Sur cette micro-plaquette qui passe dans le chas d'une aiguille, 64.000 positions de mémoire. « Mais de là à l'émotion... »

Suite page 45 ►

Suite de la page 43

comme on dit, « débloque ». Faut-il courir des objectifs aussi risqués ?

La firme IBM n'a rien changé à la politique officielle de défiance qu'elle nourrit vis-à-vis des capacités intellectuelles de ses machines. « Que nos ordinateurs en fassent le plus possible, mais qu'on ne leur reconnaisse aucune intelligence afin de ne pas effrayer les clients », a toujours été son discours au moment de vendre l'informatique.

## L'INTELLIGENCE ARTIFICIELLE

Et aujourd'hui, même si l'on y poursuit, comme ailleurs, des recherches sur l'intelligence artificielle, aucun département ou laboratoire ne porte ce nom. Quand les premières machines à vapeur apparurent dans l'industrie au début du XIX<sup>e</sup> siècle, le terme « cheval-vapeur » choqua, tant on peinait à admettre qu'un symbole exprime le remplacement d'une force animale par une machine. On imagine les réticences, maintenant qu'il s'agit de glisser une machine à

la place d'une intelligence humaine.

Parmi les informaticiens, il est courant de dire que l'intelligence artificielle n'existe pas. Même aux États-Unis, où elle est née, et où elle reste sans comparaison avec tout ce qui a pu être entrepris dans d'autres pays, une querelle reste ouverte. « Ce sont juste des gens rassemblés autour d'une idée, une sorte de groupe religieux, juge Gene Hullman, qui préside le département d'informatique à l'univer-

sité Stanford. Les limites qu'ils dénoncent dans l'image des ordinateurs sont dues à l'imprécision de notre langage. Je crois que l'on va les résoudre : l'informatique va nous obliger à plus de rigueur. »

Les tenants de l'intelligence artificielle font le pari contraire. Après Marvin Minsky, ils parlent de « *sociétés of minds* » (sociétés de consciences) où les hommes côtoieraient des machines en toute convivialité. Certaines d'entre elles sont déjà là, ce sont les ordinateurs que nous utilisons depuis plusieurs décennies. Il n'est pas question de les congédier, puisqu'il nous aident si bien. Deux siècles de recherches mathématiques nous fournissent d'ailleurs une langue commune, élaborée et riche, puissante. Mais d'autres machines viendront, « smart », complices. Elles accepteront et nos langages de tous les jours et nos emportements, puisque les informaticiens les auront aidées... à se faire une raison. ■

Jean-François FOGEL.

© Le Point.



Après les ordinateurs actuels, d'autres machines viendront qui accepteront nos comportements...

## TABLE DES MATIÈRES

### AVANT-PROPOS

### INTRODUCTION

- I. PROBLÉMATIQUE
- II. MÉTHODOLOGIE

### PREMIÈRE PARTIE – CADRE HISTORIQUE ET PRATICO-NORMATIF

- I. HISTOIRE INTELLECTUELLE DE L'IA À TRAVERS LES ÂGES
  - § 1 – Antiquité
  - § 2 – Moyen-Âge
  - § 3 – Époque moderne
  - § 4 – Époque contemporaine
- II. *AI-POWERED TOOLS* DANS LA PRATIQUE DU DROIT: DU PARADIGME DES *LEGAL EXPERT SYSTEMS* À CELUI DU *MACHINE LEARNING*
  - § 1 – Définition et enjeux
  - § 2 – Avantages, inconvénients et limites
  - § 3 – *Soft law* européen
  - § 4 – Avis d'experts
    - a. *CRIDS (UNAMUR)*
    - b. *CODEX (STANFORD)*
    - c. *SOCIÉTÉS ET START-UPS ACTIVES DANS LE DOMAINE*
- III. L'IA À L'ÉPREUVE DES DROITS HUMAINS FONDAMENTAUX
  - § 1 – Rappel de définitions (IA, SE et ML)
  - § 2 – Des fondations à la fixation du récit des DHF à l'époque moderne
  - § 3 – Droits humains en question : relativisme et droit-de-l'hommisme
  - § 4 – IA vs. DHF

## SECONDE PARTIE – MISE EN ŒUVRE DE L’IA DANS LE DROIT

### PERSPECTIVE COMPARATIVE, ÉTHIQUE ET CRITIQUE

§ 1 – Usages et pratique dans les pays dits de *common law*

§ 2 – Étude de cas : États-Unis, Japon, Europe

- a. *Silicon Valley – Eldorado de l’IA*
- b. *Japon – Pays des robots*
- c. *Comparaison avec l’Europe*

§ 3 – Enjeux éthiques, risques et impasses

- a. *Data issues et cognitives biases*
- b. *Black-box AI, explainability et transparence*
- c. *Gare à la gouvernance algorithmique !*
- d. *Quid des hard cases ?*
- e. *Retour sur l’irréductibilité de l’esprit humain*
- f. *Sur le problème que pose l’indétermination du droit ou de l’ambiguïté syntaxique et sémantique des textes juridiques à la source du droit*

§ 4 – Évaluation critique des applications actuellement disponibles

- a. *Practice Management*
- b. *Document Automation via « Lawbots »*
- c. *Computational Models of Legal Reasoning and Predictive Analytics via Comparative Legal Search Engines*

## TROISIÈME PARTIE – VERS UNE COMMON LAWISATION DU DROIT ?

### CONCLUSION – PERSPECTIVES POUR L’AVENIR

§ 1 – Éthique trans/posthumaniste

§ 2 – Éthique synthétique : *via media* entre *meta-ethics* et *virtue ethics*

### BIBLIOGRAPHIE

### ANNEXES – ANCIENNES COUPURES DE JOURNAUX

- a. J. SEMAL, « La science doit être à la mesure de l’homme », 20/11/89
- b. J.-F. FOGEL, « La percée de l’intelligence artificielle », 3/10/1983

### TABLE DES MATIÈRES

