



HAL
open science

Les algorithmes pourraient-ils remplacer les juges ?

Serge Abiteboul, Florence G'Sell

► **To cite this version:**

Serge Abiteboul, Florence G'Sell. Les algorithmes pourraient-ils remplacer les juges ?. Le Big Data et le droit, Dalloz, 2019, Thèmes et Commentaire. hal-02304016v2

HAL Id: hal-02304016

<https://inria.hal.science/hal-02304016v2>

Submitted on 26 Nov 2019

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

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

Les algorithmes pourraient-ils remplacer les juges ?

Serge Abiteboul

Directeur de recherche, Inria et École normale supérieure, Paris

Florence G'sell

Professeur à l'Université de Lorraine, chercheur associé à l'IHEJ

En mars 2019, le ministère estonien de la Justice a annoncé avoir décidé de travailler à la création d'un « juge-robot » chargé de trancher les petits litiges de moins de 7 000 euros, notamment en matière contractuelle¹. L'idée est de permettre aux deux parties à un litige civil de télécharger les informations et les documents pertinents sur la plateforme du tribunal, à partir de quoi un outil d'intelligence artificielle rendra une décision. Sont principalement visées la fixation des indemnités de licenciement et les créances impayées. Ce projet s'inscrit dans un flot continu de réformes ayant conduit l'Estonie à devenir le fer de lance du recours aux machines intelligentes dans l'administration². Par la création d'un « juge-robot », l'Estonie fait toutefois un pas supplémentaire.

Certes, des projets recourant à l'intelligence artificielle en matière judiciaire existent déjà. Au Québec, la Régie du logement a décidé de mettre en place un « *justice bot*³ » chargé d'éclairer les justiciables sur l'opportunité d'une contestation. L'idée est de développer un agent conversationnel capable, après un échange de plusieurs questions et réponses avec un justiciable, de se prononcer sur l'issue éventuelle d'un litige, par exemple de préciser une possible augmentation de loyer à la lumière d'affaires comparables. Au justiciable de décider, au regard de cette évaluation, s'il souhaite introduire une contestation. Dans d'autres pays, comme la France ou le Royaume-Uni, des processus automatisés permettent de traiter certaines contraventions, telles que les contraventions de stationnement et les excès de vitesse. Pour autant, la délégation à une machine de l'entier pouvoir de rendre une décision judiciaire en matière civile restait, à ce jour, inédite. En envisageant de confier un tel pouvoir à une intelligence artificielle et en la justifiant par le besoin d'une justice plus rapide, plus efficace, dépourvue des biais et fluctuations qui affectent les juges humains, l'Estonie franchit une étape majeure. Le ministère estonien insiste sur le fait que les tribunaux auront ainsi plus de

¹E. Niiler, « Can AI be a fair judge in court ? Estonia thinks so », *Wired*, 25 mars 2019 ; L.-M. Gaveriaux, « Les robots débarquent au tribunal », *Les Échos*, n° 174, 19 juin 2019.

²Les citoyens estoniens disposent déjà d'une carte d'identité numérique et de la possibilité de voter ou déclarer leurs impôts en ligne. Ils bénéficient de l'automatisation de nombreuses démarches administratives, telle l'inscription de leurs enfants à l'école, qui est effectuée automatiquement par une machine à partir du registre des naissances. Des outils intelligents ont été développés pour assister les chômeurs dans leur recherche d'emploi ou pour réaliser certains contrôles administratifs, comme l'inspection des champs de foin. Les bases de données gouvernementales sont interconnectées sans que les administrés, qui sont tenus informés des accès aux informations les concernant, s'en émeuvent particulièrement.

³C'est le laboratoire de cyberjustice de l'Université de Montréal et l'Institut des algorithmes d'apprentissage de Montréal qui sont chargés de développer ce *justice bot*. V. <https://www.cyberjustice.ca/projets/justicebot/>

temps à consacrer aux cas les plus compliqués. Il garantit, par ailleurs, qu'il sera toujours possible d'exercer un recours contre la décision de la machine devant un juge humain.

L'idylle de la justice et de l'informatique était prévisible, malgré toutes les tensions auxquelles peuvent conduire les transformations violentes d'une vénérable institution. Mais l'exemple estonien réactive des questions maintes fois soulevées par l'utilisation de l'informatique dans de nombreux domaines. On peut notamment se demander si les machines sont réellement en mesure de mener des raisonnements juridiques et d'adopter des décisions exécutoires. Autrement dit, dès lors que l'intelligence artificielle permet d'ores et déjà de réaliser, à ce jour, un certain nombre de tâches juridiques, la perspective du « juge-robot » est-elle crédible ? De fait, s'il est incontestable que l'informatique est en train de bouleverser en profondeur le processus judiciaire (I), l'hypothèse du « juge-robot » ne peut que susciter une réaction partagée (II).

I. L'apport de l'informatique au processus judiciaire

Un algorithme est avant tout une méthode pour trouver une solution à un problème⁴. Quand on est capable d'articuler une telle solution, on peut la décrire sous forme d'algorithme, puis l'écrire sous forme de programme informatique, de logiciel, pour faire réaliser l'algorithme par un ordinateur. C'est le cas pour un logiciel de l'Agence spatiale européenne qui envoie une fusée dans l'espace. Mais c'est aussi ce qui se passe quand on « dématérialise » un processus judiciaire : on « numérise l'information » et on « code les traitements » sous forme de programmes informatiques. Aujourd'hui, les praticiens du droit, avocats et juristes d'entreprise, ont d'ores et déjà recours quotidiennement à des outils informatiques, par exemple pour l'analyse de document (*document review*) et l'audit de contrat, pour lequel la machine relève les clauses contractuelles faisant difficulté⁵. Mais il est probable que, dans un futur proche, ce soit l'ensemble des professionnels du droit, y compris les juges, qui aient recours à de telles technologies. L'informatique permet non seulement la dématérialisation des procédures (A) et un accès efficace à l'information juridique (B) mais également de mener des raisonnements, en servant alors d'outil d'aide à la décision (C).

A. La dématérialisation des procédures et le partage d'information

La procédure judiciaire met en jeu de nombreux échanges d'informations : entre les parties, les avocats, les greffes, le ministère public, les juges... Or l'informatique constitue une aide fantastique pour échanger de l'information. Les logiciels peuvent faciliter le travail de la justice en simplifiant les échanges (courriels, vidéoconférences, etc.) et en permettant la préparation et le suivi des documents (actes, notifications, etc.). De fait, depuis quelques années, des services de cette nature sont proposés aux justiciables par différentes *Legaltechs*, des entreprises mettant la technologie au service de la justice. Le site demanderjustice.com offre, par exemple, un service de saisine en ligne des tribunaux d'instance et conseils des prud'hommes, pour lesquels la représentation par avocat n'est pas obligatoire. À partir des éléments saisis en ligne par le demandeur, la plateforme se charge de mettre en forme la demande puis de l'envoyer par la poste à la juridiction compétente⁶. Aux États-Unis, la start-

⁴S. Abiteboul, G. Dowek, *Le temps des algorithmes*, Paris, Le Pommier, 2017.

⁵En matière d'outils d'analyse de contrats, on peut ainsi citer les start-up Softlaw et Hyperlex, par ex., v. E. Barthe, « Les outils de l'intelligence artificielle pour le droit français », *JCP* 2019. Doctr. 381.

⁶www.demanderjustice.com

up DoNotPay propose un service encore plus sophistiqué⁷. À l'origine, l'application a été développée pour générer automatiquement et gratuitement des réclamations contestant des contraventions de stationnement. Le succès a été immédiat : en 2016, des milliers de contraventions ont été annulées à Londres et à New York. Par la suite, le programme a été amélioré de manière à pouvoir engager d'autres types de procédure aux États-Unis et au Royaume-Uni : conflits relatifs aux baux d'habitation, bagages perdus par les compagnies aériennes, demandes d'asile, licenciement, litiges de consommation, divorce, réparation d'un dommage, plainte au pénal⁸, etc.

Révélatrice de ce que la technologie est aujourd'hui capable de proposer, l'application DoNotPay ne connaît pas, à ce jour, d'équivalent en France. Il existe cependant des *Legaltechs* françaises comme LegalPlace, Legalstart ou Captain Contrat offrant des services de rédaction automatisée de documents juridiques comme des contrats usuels ou des statuts de sociétés. On se pose d'ailleurs parfois la question de savoir si les tâches réalisées aujourd'hui par les applications développées par les *Legaltechs* correspondent à des activités de consultation juridique. Ce n'est, pour l'instant, pas le cas. Ces applications, qui assument pour l'heure des tâches relativement simples et répétitives, ne réalisent pas encore de raisonnements juridiques élaborés, et notamment ne se livrent pas au raisonnement bien connu des juristes, qui consiste à rapprocher une règle générale d'un cas concret particulier pour donner un avis juridique personnalisé⁹. Les choses devraient, en revanche, évoluer dans un avenir proche. Aux États-Unis, l'outil d'intelligence artificielle Alexsei¹⁰ propose, par exemple, de répondre en 24 heures aux questions juridiques qui lui sont posées en produisant, de manière automatisée, un mémorandum réalisé à partir de l'analyse des données de jurisprudence.

La mise à disposition d'outils technologiques sophistiqués n'a toutefois pas vocation à être exclusivement le fait des *Legaltechs*. Il revient également au service public de se saisir de la technologie pour proposer aux citoyens des services innovants et efficaces. Dans cette perspective, la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice comporte des mesures permettant la numérisation des procédures civiles et pénales, même si l'on peut regretter que cette transformation numérique reste encore trop limitée¹¹. En matière pénale, la loi prévoit la possibilité de déposer plainte¹² ou de se constituer partie civile¹³ en ligne. Tous les actes de procédures pourront être établis ou convertis sous format numérique¹⁴ dans le cadre d'un dossier de procédure numérique. En matière civile, certaines procédures, comme celles relatives aux petits litiges¹⁵ ou les injonctions de payer¹⁶, seront entièrement dématérialisées. Les justiciables pourront suivre en ligne l'avancement de la procédure sur le « portail du justiciable » où pourront être transmis

⁷www.donotpay.com.

⁸V., J. Porter, « Robot lawyer DoNotPay now lets you “sue anyone” via an app », *The Verge*, 10 oct. 2018 ; S. Wilkins, « DoNotPay is the Latest *LegalTech* Darling, but Some Are Saying Do Not Click », *Above the Law*, 12 oct. 2018, <https://abovethelaw.com/legal-innovation-center/2018/10/12/donotpay-is-the-latest-legal-tech-darling-but-some-are-saying-do-not-click/>

⁹F. G'ssell, « La protection du périmètre du droit face aux *Legaltechs* », *D. avocats* mars 2019.

¹⁰<https://www.alexsei.com/>

¹¹« Sur les 110 articles de la loi, seuls 8 traitent expressément des outils numériques », J.-B. Thierry, « La loi n° 2019-222 du 23 mars 2019, loi de réforme pour la justice numérique ? », *JCP* 2019. Doctr. 524. V. égal. C. Bléry, « Loi du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice : aspects numériques », *D.* 2019. 1069 ; A. Lecourt, « Loi de programmation et numérique : un dialogue fructueux ? », *Dalloz IP/IT* 2019, p. 514.

¹²Nouv. art. 15-3-1 C. pr. pén.

¹³Art. 420-1 C. pr. pén.

¹⁴Art. 8°1-1 C. pr. pén.

¹⁵Nouv. art. L. 212-5-2 COJ.

¹⁶Nouv. art. L. 211-18 COJ.

les avis, convocations et récépissés émis par les greffes des tribunaux civils et des cours d'appel. La possibilité de recourir à la visioconférence est également prévue¹⁷.

En prévoyant un plus large recours aux outils numériques dans la procédure, la France suit un mouvement général que l'on retrouve dans d'autres pays¹⁸. Par exemple, dans l'Ontario, au Canada, la Cour supérieure de justice (*Ontario Superior Court of Justice*), a engagé un projet pilote en février 2019, appelé le « *Digital Hearing Workspace (DHW)* », actuellement appliqué à tous les litiges commerciaux. Il s'agit de mettre à disposition des juges une application permettant de stocker, rechercher et établir électroniquement tous les documents relatifs à une affaire¹⁹.

B. La recherche documentaire

L'informatique facilite considérablement la recherche documentaire. Les logiciels peuvent assister les juristes dans la recherche de jurisprudence, leur faisant ainsi gagner un temps considérable ; on évite de se déplacer, d'avoir à parcourir des volumes de textes inutiles. Les bases de données équipées de moteurs de recherche perfectionnés jouent ainsi le rôle de mémoire auxiliaire pour les juristes, qui n'ont plus à mémoriser le grand nombre de cas et à se plonger dans des archives de documents papiers. Les bases de données existantes, qu'il s'agisse de la plateforme Légifrance ou de celles proposées par les éditeurs juridiques²⁰ ou les *Legaltechs* comme Doctrine, seront prochainement enrichies en France grâce à la publication de l'ensemble des décisions rendues sur le territoire français en vertu de la politique d'open data des décisions de justice prévue par la loi « pour une République numérique » n° 2016-1321 du 7 octobre 2016 et la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice. L'existence de telles bases de données ne va pas seulement permettre un accès plus simple et rapide aux décisions et à l'information juridique. Elle va également permettre la quantification des risques pour aider à décider de stratégies grâce, précisément, à l'apprentissage automatique.

Déjà présent dans la vie des juristes, le recours à des outils intelligents en matière de recherche juridique devrait de plus en plus s'imposer. On peut relever, à cet égard, la décision intéressante d'une juridiction canadienne ayant refusé à la partie victorieuse le remboursement des frais de procès relatifs à la recherche documentaire. L'un des arguments mobilisés par le juge était, en l'occurrence, que l'avocat de la partie l'ayant emporté n'avait non seulement pas eu recours aux outils de recherche juridique gratuits mais aussi qu'il s'était abstenu d'utiliser des outils d'intelligence artificielle qui auraient précisément permis de réduire significativement le montant des honoraires facturés, donc des frais encourus²¹.

C. L'apport de l'intelligence artificielle au raisonnement juridique

¹⁷Le Conseil constitutionnel a toutefois censuré les dispositions écartant la possibilité de refuser la visioconférence pour l'audience relative à la prolongation de la détention provisoire, Décis. n° 2019-778 DC du 21 mars 2019, spéc. pts 231-234.

¹⁸G. Canivet (dir.), *Justice, faites entrer le numérique*, Rapport de l'Institut Montaigne, nov. 2017.

¹⁹<https://www.flexlegalnetwork.com/blog/the-commercial-lists-new-digital-hearing-workspaces-what-you-need-to-know>

²⁰Aux États-Unis, l'éditeur juridique LexisNexis a, par exemple, racheté Ravel Law, spécialiste de l'apprentissage automatique appliqué à la jurisprudence, afin d'incorporer la technologie Ravel à sa plateforme. V., E. Barthe, « Les outils de l'intelligence artificielle pour le droit français », art. cit.

²¹*Cass c/ 1410088 Ontario Inc.*, 2018 ONSC 6959.

Il y a quelques années, les juristes se sont tournés vers les systèmes experts, des outils capables de reproduire les mécanismes cognitifs d'un expert par des raisonnements logiques, à partir d'une base de connaissances et d'un moteur d'inférence. La base de connaissances contient des connaissances spécifiques du domaine sous une forme exploitable par un ordinateur. Des règles, elles aussi incluses dans la base de connaissance, sont appliquées à ces connaissances spécifiques par le moteur d'inférence afin de déduire de nouvelles connaissances. Le système expert est ainsi supposé résoudre un problème de la manière dont le ferait un expert humain. Dans le domaine juridique, les systèmes experts décomposent les règles de droit en les réécrivant en langage informatique, de manière à établir un arbre de décision constitué de ramifications successives associées à une logique conditionnelle²². On retrouve par exemple ces systèmes experts dans les logiciels de gestion de contrat traditionnels²³ ou pour le calcul du montant d'une prestation compensatoire²⁴. Les résultats ont toutefois été jugés généralement décevants, même à propos de questions très techniques où il ne semblait y avoir qu'à reproduire un raisonnement syllogistique relativement simple pour dégager des solutions correctes. Cet échec relatif s'explique par le raisonnement assez réducteur de systèmes experts, qui ne savent prendre en compte ni les présomptions ni les analogies et ne peuvent se livrer au va-et-vient permanent des faits au droit qui caractérise le raisonnement juridique²⁵. Cela pourrait être corrigé dans de futures générations de logiciels. Une difficulté est que les règles juridiques manquent parfois de la précision adaptée pour un raisonnement mathématique, et qu'elles peuvent conduire à des contradictions, de sérieux handicaps pour des logiciels prenant en compte des centaines de règles²⁶. Le développement de l'apprentissage automatique (*machine learning*) a ouvert de nouvelles perspectives. Il a non seulement permis des évolutions majeures dans le traitement du langage naturel (*natural language processing*, ou NLP) mais a aussi conduit à l'analyse, supervisée ou non, d'immenses gisements de données, composés des décisions judiciaires, de règles juridiques ou d'exemples de cas.

Comment fonctionnent ces « algorithmes d'apprentissage automatique » ? Dans la vie quotidienne, nous ne savons pas toujours expliquer comment nous résolvons certains problèmes qui nous paraissent pourtant simples. Par exemple, nous savons distinguer un chat d'un chien tout en étant incapables d'expliquer comment nous le faisons. En réalité, nous « apprenons » à partir d'exemples. On nous montre un chat en le qualifiant, puis un chien, et cela pour un certain nombre d'exemples. À partir de là, nous construisons des modèles mémoriels de ces deux espèces qui nous permettent de les distinguer. De même, un algorithme d'apprentissage automatique se nourrit d'exemples fournis par des humains, ou que l'algorithme découvrira en ligne, voire construira. La machine finit donc par être en mesure de distinguer un chat d'un chien. Ce genre de technique se révèle très prometteur, par exemple, pour la détection de cancers²⁷. Dans les sciences juridiques, et plus largement dans

²²D. Bourcier, *La Décision artificielle : Le Droit, la machine et l'humain*, Paris, PUF, coll. « Les voies du droit », 1995 ; « L'acte de juger est-il modélisable ? », *Archives Phil. dr.* 2011, t. 54, *De la logique à la justice*, p. 37 s.

²³E. Barthe, « Les outils de l'intelligence artificielle pour le droit français », art. cit.

²⁴L. Godefroy, « Encadrement juridique des Modes Algorithmiques d'Analyse des Décisions (MAAD) », in L. Godefroy, F. Lebaron, J. Lévy-Véhel, *Comment le numérique transforme le droit et la justice vers de nouveaux usages et un bouleversement de la prise de décision, Anticiper les évolutions pour les accompagner et les maîtriser*, Mission de recherche Droit et Justice, Rapport n° 16-42, juill. 2019, p. 53.

²⁵Y. Tanguy, « Le droit à l'épreuve de l'informatique juridique », *Politiques et management public* 1993, vol. 11, n° 4, p. 135-146. V. égal. C. Mouly, « Signification épistémologique de l'échec des systèmes experts juridiques », *LPA* 1989, n° 91, p. 4.

²⁶E. Barthe, « Les outils de l'intelligence artificielle pour le droit français », art. cit.

²⁷K. Kourou *et al.*, « Machine learning applications in cancer prognosis and prediction », *Computational and Structural Biotechnology Journal* 2015, vol. 13, p. 8-17. La start-up française Owkin a, par exemple, développé

les sciences humaines, il peut également arriver que l'on ne sache décrire précisément ou expliquer le cheminement mental vers une solution. Certes, l'argumentation juridique vise à démontrer pourquoi telle ou telle décision s'impose plutôt qu'une autre. Un juge explique ses décisions, un avocat appuie son plaidoyer sur un raisonnement. Mais les explications fournies peuvent être imprécises et sujettes à interprétation. Surtout, elles capturent de manière très incomplète les raisonnements suivis. Par exemple, l'avocat n'explique pas pourquoi il a suivi une stratégie plutôt qu'une autre, le juge ne détaille pas toujours les questionnements qui ont guidé sa décision. Par analogie avec les mathématiques, on pourrait donc dire que ce qui est fourni est une esquisse de preuve que l'on a du mal à suivre et vérifier, alors que l'enjeu est de comprendre le raisonnement qui a conduit le mathématicien au résultat dans sa complétude. Il ne sera donc pas évident, à partir des explications fournies, d'en tirer un algorithme au sens classique du terme, soit un ensemble d'instructions conçu pour résoudre un problème.

Si on dispose de nombreuses décisions de justice, dans un contexte restreint, un algorithme d'apprentissage pourra se former à proposer une solution en se fondant sur les décisions précédemment adoptées. Un exemple est donné par Anthony Casey et Anthony Niblett²⁸ à propos de l'affaire *Jacobellis c/ Ohio* dans laquelle le juge a estimé qu'il était très difficile de déterminer précisément le critère de l'image pornographique et s'est contenté d'écrire : « Je reconnais la pornographie quand je la vois. » Il suffirait aujourd'hui, disent-ils, de soumettre un certain nombre d'images jugées pornographiques à une machine, pour que celle-ci soit alors en mesure de déterminer ce qui est pornographique et ce qui ne l'est pas. L'exemple illustre le principal apport de l'apprentissage automatique : il est en effet possible d'examiner toutes les décisions antérieurement adoptées pour proposer une décision prise sur le même modèle. Autrement dit, l'apprentissage automatique permet de prédire ou de proposer des solutions en considérant les décisions antérieures. Le recours à l'apprentissage automatique ne revient donc pas à proposer une argumentation juridique justifiant la proposition de décision, qui est une tâche bien plus complexe. La distinction doit être particulièrement soulignée dans notre système juridique, qui ne repose pas sur la règle du précédent.

Dans les pays de *Common Law*, le recours à l'apprentissage automatique s'inscrit dans une démarche classique consistant à se référer aux précédents pour prendre parti ou anticiper la décision du juge. L'apprentissage automatique se révèle, à cet égard, particulièrement efficace. Par exemple, l'application Ross, fondée sur l'intelligence artificielle du logiciel Watson d'IBM, recourt à l'apprentissage automatique dans sa version supervisée. Le logiciel DoNotPay, évoqué plus haut, a récemment intégré cette technique d'IA afin de permettre la saisine en ligne des tribunaux dans tous les États des États-Unis. Il suffit à l'utilisateur d'écrire sur l'écran d'accueil la nature de la difficulté qu'il rencontre puis de répondre à un certain nombre de questions posées par la machine. La machine classe alors l'affaire parmi 15 différents types de contentieux prédéterminés (inexécution contractuelle, négligence, etc.). Elle produit alors, de manière entièrement automatisée et sans intervention humaine, les éléments nécessaires pour saisir le tribunal et est en mesure d'établir un argumentaire que le plaignant n'a qu'à lire à l'audience. La machine est en effet capable de préparer à l'avance une stratégie juridique de réponse aux arguments prévisibles du défendeur. DoNotPay revendique un taux de succès de 55 % pour les actions introduites avec son outil. L'application a été mise à la disposition, à titre gratuit, des victimes de la divulgation par Equifax des données personnelles de ses clients : les 25 000 recours engagés grâce à l'outil de DoNotPay auraient permis aux plaignants d'obtenir en moyenne 9 000 dollars de dommages-

un logiciel très performant de détection des mésothéliomes : Courtiol, P., Maussion, C., Moarii, M. *et al.* « Deep learning-based classification of mesothelioma improves prediction of patient outcome », *Nature Medicine*, vol. 25, 1519–1525 (2019).

²⁸V. la contribution de A. Casey et A. Niblett, « La fin des règles et des standards », dans cet ouvrage.

intérêts²⁹. Pour autant, l'entreprise souligne elle-même que l'application n'est conçue que pour les petits litiges répétitifs, et non pour des contentieux complexes³⁰. Il est évident que de tels outils ne peuvent remplacer les connaissances, l'expérience, les intuitions des juristes dans des cas complexes.

En France, plusieurs start-up, comme Case Law Analytics ou Predictice, proposent des outils prédictifs qui permettent de mieux évaluer les résultats possibles d'une procédure, notamment en matière de quantification des sommes allouées : dommages et intérêts, pensions alimentaires, prestations compensatoires³¹. Cependant, au-delà de cette « justice quantitative³² », il reste difficile de déterminer jusqu'à quel point les algorithmes d'apprentissage automatique seraient en mesure de produire des résultats pertinents dans notre système juridique. En droit français, les décisions judiciaires découlent en principe d'un raisonnement de type syllogistique et déductif, même si les tribunaux sont sensibles aux décisions antérieurement adoptées sur une question donnée et respectent les décisions passées des juridictions supérieures afin d'éviter une future censure. La décision judiciaire est généralement réputée se fonder sur une certaine compréhension de la règle générale qui justifie son application à une situation particulière. On suppose donc que les concepts juridiques ont un sens, voire même un seul sens, qui existerait *a priori*, et qu'il suffit d'interpréter la règle générale.

Prenons l'exemple de l'interprétation de la loi du 5 juillet 1985 qui s'applique aux accidents dans lesquels sont impliqués des « véhicules terrestres à moteur ». Déterminer le champ d'application de cette loi suppose de préciser ce qu'est un véhicule terrestre à moteur. La question s'est par exemple posée de savoir si une tondeuse à gazon est, ou non, un véhicule terrestre à moteur. Il a été jugé qu'une tondeuse à gazon « auto portée » est un véhicule terrestre à moteur dans la mesure où elle constitue « un engin à moteur doté de quatre roues lui permettant de circuler, équipé d'un siège sur lequel une personne prend place pour le piloter³³ ». En revanche, une tondeuse à gazon dirigée par une personne marchant à pied n'est pas considérée comme un véhicule terrestre à moteur. Un algorithme aurait-il pu faire le raisonnement qui conduit à faire une telle distinction ? En toute autonomie ? En l'absence de décision antérieure exploitable sur ce sujet précis, il aurait fallu pouvoir se référer à d'autres décisions relatives aux tondeuses à gazon et réfléchir à la sémantique du mot « véhicule terrestre à moteur ». Pour pouvoir parvenir à une réponse pertinente, l'algorithme aurait donc dû avoir recours à d'autres techniques d'intelligence artificielle que l'apprentissage automatique comme les bases de connaissances généralistes, l'analyse sémantique de texte, et en particulier le raisonnement logique. Certes, les algorithmes sont, par nature, bien adaptés au raisonnement. Par exemple, le « théorème des quatre couleurs » a été prouvé en utilisant des ordinateurs par Georges Gonthier et Benjamin Werner à l'aide du logiciel Coq en 2005. Le contexte judiciaire recèle toutefois des difficultés qui lui sont propres, qui tiennent, notamment, à l'imprécision des données dont on dispose, à leur incomplétude (tous les cas ne sont pas toujours prévus), voire aux contradictions qui y sont présentes, car l'on peut parfois

²⁹C. Haskins, « New App Lets You “Sue Anyone by Pressing a Button” », *Vice*, 10 oct. 2018 https://www.vice.com/en_us/article/bj43y8/donotpay-app-lets-you-sue-anyone-by-pressing-a-button

³⁰L'application est par ailleurs en voie de diversification, avec des nouveaux services, là encore destinés aux particuliers : évaluer les trajets Uber, contester les trop-perçus par les banques, vérifier le respect de la réglementation en vigueur par les plateformes, assurer des réservations d'hôtels et de vols, obtenir une indemnisation en cas de livraison, faire des démarches auprès de la Sécurité sociale, obtenir des visas, des passeports ou des cartes vertes (*green cards*), prendre rendez-vous par téléphone...

³¹Pour une présentation exhaustive des différents acteurs de ce domaine, v. L. Godefroy, F. Lebaron, J. Lévy-Véhel, *Comment le numérique transforme le droit et la justice...*, *op. cit.*, p. 8 s.

³²J. Dupré, « Du droit saisi par l'IA au droit saisissant l'IA, éléments de réflexion », *Archives Phil. dr.* 2018, t. 60, p. 103.

³³Civ. 2^e, 24 juin 2004, n° 02-20.208, *Bull. civ.* II, n° 308, p. 260.

trouver des lois ou des arrêts qui se contredisent. On s'est d'ailleurs aperçu que l'on ne pouvait traduire tout le droit sous forme de règles précises et exploitables par une machine³⁴. Pour toutes ces raisons, l'informatique, si elle s'appuie strictement sur des raisonnements logiques, présente des limites.

L'informatique est encore loin de pouvoir réaliser le raisonnement déductif propre au droit français réalisé par un avocat ou un juge, qui revient à faire le va-et-vient entre règle générale et cas particulier (raisonnement syllogistique³⁵). Dans un tel système juridique, où l'examen des décisions passées n'est pas forcément suffisant, l'informatique actuelle peut aider dans les cas simples, mais pas les plus complexes. Pour le dire autrement, l'intelligence artificielle n'est pas capable, à ce jour, de remplacer les juges et les avocats. Le sera-t-elle un jour ? Peut-être. Mais pour ce faire, il faudra alors convoquer une large gamme de techniques au-delà des techniques d'apprentissage actuellement considérées incluant notamment les bases de connaissance, la recherche automatique de preuve, l'analyse du langage naturel. Tout ceci conduit donc à s'interroger sur les contextes dans lesquels l'utilisation d'un juge-robot comme celui de l'Estonie est véritablement souhaitable.

II. Le « juge-robot », une perspective envisageable ?

La perspective du « juge-robot » ou de « l'avocat-robot » fait souvent réagir³⁶. L'on met en avant le fait que seuls les humains sont dotés des modes de pensée, de l'intuition, de l'empathie qui seraient nécessaires à l'analyse d'un dossier ou au jugement d'une affaire. Cet argument est toutefois combattu. Il a été qualifié de fallacieux (*the AI fallacy*) par Richard et Daniel Susskind³⁷, pour qui l'idée selon laquelle certaines activités ne pourront jamais être réalisées par des machines n'est pas tenable : les machines auraient, tout simplement, leur propre manière de faire les choses et d'être efficaces. S'agissant de la justice, s'il est incontestable qu'une justice purement humaine présente des imperfections (A), les algorithmes ont eux-mêmes leurs limites (B) qui font douter que l'on puisse, en l'état, purement et simplement remplacer les humains par des machines. Pour l'heure, la perspective d'une justice automatisée reste fort lointaine (C).

A. Les imperfections de la justice humaine

La justice humaine est loin d'être parfaite. Quand ils participent à rendre la justice, les humains peuvent être partiels. Des études menées aux États-Unis ont montré que, dans des conditions similaires, les montants exigés pour une remise en liberté sous caution sont plus élevés de 25 % lorsque le prévenu est noir que lorsqu'il est blanc³⁸. De même, les peines

³⁴D. Bourcier, « L'acte de juger est-il modélisable ? », art. préc., p. 45.

³⁵Aux yeux de certains, « il est depuis longtemps acquis que le raisonnement juridique ne peut être programmé », F. Rouvière, « Le raisonnement par algorithme : le fantasme du juge-robot », *RTD civ.* 2018. 530 s. V. égal. Y. Tanguy, art. préc. et D. Bourcier, « L'acte de juger est-il modélisable ? », art. cit., spéc. p. 39-40.

³⁶F. Rouvière, « Le raisonnement par algorithme : le fantasme du juge-robot », art. préc. ; D. Reiling, « Quelle place pour l'intelligence artificielle dans le processus de décision d'un juge ? », *Cah. just.* 2019. 221. V. égal. A. van den Branden, *Les robots à l'assaut de la justice, L'intelligence artificielle au service des justiciables*, Bruylant, 2019 ; E. Poinas, *Le tribunal des algorithmes*, Berger-Levrault, coll. « Au fil du débat », 2019.

³⁷R. et D. Susskind, *The Future of the Profession : How Technology Will Transform the Work of Human Experts*, Oxford Univ. Press, 2016 ; R. Susskind, *Tomorrow's lawyers : An Introduction to Your Future*, 2^e éd., Oxford Univ. Press, 2017.

³⁸I. Ayres et J. Waldfoegel, « A Market Test for Race Discrimination in Bail Setting », *Stanford Law Review* 1994, n° 46, p. 987.

prononcées sont plus longues de 12 % lorsqu'elles concernent un défendeur noir³⁹. En cas d'homicide, le prononcé de la peine de mort est plus probable lorsque l'accusé est noir et les peines plus sévères lorsque la victime est blanche⁴⁰. Le genre des parties au litige constitue également un biais fréquemment relevé⁴¹. En France, il apparaît que les femmes sont condamnées en moyenne à vingt jours de prison ferme de moins que les hommes, toutes choses égales par ailleurs⁴². De manière générale, il est acquis que les préférences politiques, l'origine ethnique, le genre et certaines caractéristiques démographiques ont des effets sur les décisions judiciaires⁴³. D'autres circonstances influencent les décisions des juges, comme la répugnance à prendre plusieurs décisions d'affilée allant dans le même sens⁴⁴, le contexte d'une campagne électorale⁴⁵, le contexte médiatique⁴⁶, les résultats d'une équipe de football locale⁴⁷ ou la date d'anniversaire du défendeur⁴⁸. On pourrait enfin citer le célèbre exemple de l'influence de l'heure de la journée sur les décisions. Il a été démontré qu'il est plus probable d'obtenir un jugement favorable en début de journée ou après un repas : la probabilité de jugement favorable décline, en effet, au fil des heures de travail⁴⁹. Bref, les biais des juges seraient si importants qu'il a même été suggéré d'intégrer ces biais dans les modèles prédictifs destinés à prédire les décisions de justice de manière à améliorer leur fiabilité⁵⁰ !

Au-delà de leur caractère parfois partial et subjectif, les décisions humaines peuvent être difficiles à justifier ou à expliquer. Certes, en matière judiciaire, les tribunaux ont l'obligation de motiver leurs décisions, en fait et en droit. Cependant, il a déjà été remarqué que la motivation judiciaire, destinée à accréditer l'idée que le juge est neutre et ne fait qu'appliquer le droit, ne serait finalement qu'un leurre⁵¹. « Au premier abord, [la motivation] semble destinée à montrer que l'auteur d'une décision était lié. Puis, l'analyse révèle qu'il avait un pouvoir bien réel, que la motivation avait pour fonction de masquer », écrivait Michel

³⁹D. B. Mustard, « Racial, Ethnic, and Gender Disparities in Sentencing : Evidence from the U.S. Federal Courts », *Journal of Law & Economics* 2001, n° 44, p. 285, 300.

⁴⁰R. R. Banks *et al.*, « Discrimination and Implicit Bias in a Racially Unequal Society », *California Law Review* 2006, n° 94, p. 1169, 1175. V. égal. J. J. Rachlinski, S. L. Johnson, A. J. Wistrich et C. Guthrie, « Does Unconscious Racial Bias Affect Trial Judges ? », *Notre Dame Law Review* 2009, n° 84, p. 1195.

⁴¹A. L. Miller, « Expertise fails to Attenuate Gendered Biases in Judicial Decision-Making », *Social Psychological and Personality Science* mars 2019, vol. 10, n° 2, p. 227-234.

⁴²A. Philippe, « Vous jurez de n'écouter ni la haine, ni la méchanceté... Les biais affectant les décisions de justice », *Cah. just.* 2015/4. 563-577. V. plus largement l'ensemble du dossier intitulé « Des juges sous influence », *Cah. just.* 2015/4. 499 s.

⁴³T. J. Miles et C. R. Sunstein, « The New Legal Realism », *Univ. of Chicago Law Review* 2008, n° 75, p. 831 ; v. égal. C. Jolls et C. R. Sunstein, « The Law of Implicit Bias », *California Law Review* 2006, n° 94, p. 969 qui évoquent les biais implicites relatifs aux groupes désavantagés.

⁴⁴D. L. Chen, T. J. Moskowitz, K. Shue, « Decision-Making under the Gambler's Fallacy : Evidence from Asylum Judges, Loan Officers, and Baseball Umpires », *NBER Working Paper* 2016, n° 22026.

⁴⁵D. L. Chen, « Priming Ideology : Why Presidential Elections Affect US Judges », *TSE Working Paper* 2016, n° 16-681.

⁴⁶A. Philippe et A. Ouss, « No hatred or malice, fear or affection : Media and sentencing », *Working Paper* 2015, cité par A. Philippe, « Vous jurez de n'écouter ni la haine, ni la méchanceté... », préc.

⁴⁷O. Eren et N. Mocan, « Emotional Judges and Unlucky Juveniles », *NBER Working Paper* sept. 2016, n° 22611.

⁴⁸D. L. Chen & Arnaud Philippe, « Clash of norms : Judicial leniency on defendant birthdays », *TSE Working Papers* 2018, n° 18-934, Toulouse School of Economics (TSE).

⁴⁹S. Danziger, J. Levav et L. Avnaim-Pesso, « Extraneous factors in judicial decisions », *PNAS* avr. 2011, n° 108 (17), p. 6889-92.

⁵⁰D. L. Chen, « Judicial Analytics and the great transformation of American Law », *Artificial Intelligence and Law* mars 2019, vol. 27, n°1, p. 15-42.

⁵¹W. Mastor, « Essai sur la motivation des décisions de justice. Pour une lecture simplifiée des décisions des Cours constitutionnelles », *AIJC* 2000. 35-63.

Troper⁵². Ainsi, « ce qui est exposé dans la motivation, ce n'est pas réellement le processus intellectuel qui a conduit à la décision. La question de savoir comment celle-ci a été réellement obtenue ne relève d'ailleurs pas de la théorie du droit, mais des sciences du comportement⁵³ ». S'il est incontestable que les juges français sont imprégnés d'une rationalité juridique par laquelle ils se sentent liés⁵⁴, la part de liberté dont ils jouissent et la faculté dans laquelle ils sont d'exercer leur pure subjectivité sont indéniables. On pourrait se référer ici au propos d'Antonio Cassese, ancien président du Tribunal pénal international pour l'ex-Yougoslavie :

« Nous avons tous rédigé des jugements. Nous savons que l'on pourrait considérer les juges comme des experts en « manipulation ». Les juges « manient » habilement les lois, les critères, les principes d'interprétation dans le but, bien sûr, de rendre justice dans un cas d'espèce. En particulier dans la Justice pénale, on sent intuitivement qu'un homme est coupable, que le sens commun devrait nous conduire à cette conclusion. La construction du magnifique raisonnement juridique qui le justifie est postérieure⁵⁵. »

Dans cette perspective, la motivation aurait finalement surtout pour fonction de convaincre les justiciables du bien-fondé de la décision⁵⁶. Cet objectif est-il, dans les faits, atteint ? Les arguments donnés sont-ils toujours suffisamment complets, précis, de nature à emporter l'adhésion ? En réalité, la clarté de la motivation des décisions est parfois mise en cause⁵⁷ malgré l'exigence régulièrement rappelée d'intelligibilité des décisions⁵⁸. Les juristes français ont, par le passé, fait part de leur insatisfaction sur ce point, et demandé que la motivation des décisions permette réellement « au juge de livrer le fond de sa pensée, d'expliquer vraiment pourquoi il statue dans un certain sens, sans rien cacher des éléments qu'il prend en considération⁵⁹ ». L'on ignore, pour l'heure, si les dispositions récemment prises par la Cour de cassation pour améliorer la motivation des décisions seront de nature à mieux éclairer les justiciables sur le cheminement du juge⁶⁰.

La partialité des humains chargés de juger leurs semblables ainsi que l'éventuelle opacité du processus décisionnel peuvent conduire à conclure qu'une intelligence artificielle programmée pour adopter des décisions en fonction de critères objectifs, de manière quasi mathématique, se révélerait, finalement, plus juste et légitime. Une appréciation algorithmique du risque de récidive d'un prévenu serait préférable à une perception humaine

⁵²M. Troper, « La motivation des décisions constitutionnelles », in Ch. Perelman et P. Foriers (dir.), *La motivation des décisions de justice*, Bruylant, Bruxelles, 1978, p. 287, spéc. p. 301 .

⁵³M. Troper, *op. cit.*, p. 294.

⁵⁴D. de Béchillon, « Sur la diversité des méthodes et manières de juger », conférence prononcée dans la Grand'Chambre de la Cour de cassation le 29 nov. 2004.

⁵⁵R. Badinter et S. Breyer (dir.), *Les entretiens de Provence, Le juge dans la société contemporaine*, Paris, Fayard, 2003, p. 44. Ce propos rendait Ronald Dworkin fort triste ! V. égal. D. de Béchillon, « Sur la diversité des méthodes et manières de juger », conférence préc.

⁵⁶« La motivation est la condition de la lisibilité du jugement, de sa légalité et de sa légitimité », N. Fricero, « La qualité des décisions de justice au sens de l'article 6 § 1^{er} de la Convention européenne des droits de l'homme », in P. Mbongo (dir.), *La qualité des décisions de justice*, Strasbourg, Éd. du Conseil de l'Europe, 2007, p. 49 ; v. égal. H. Colombet et A. Gouttefangeas, « La qualité des décisions de justice. Quel critère ? », *Droit et société* 2013/1, n° 83, p. 155-176.

⁵⁷W. Mastor, art. préc., p. 56-57. Le problème n'est pas nouveau et la rédaction lapidaire des arrêts de la Cour de cassation a par le passé fait l'objet de critique : v. A. Touffait et A. Tunc, « Pour une motivation plus explicite des décisions de justice, notamment celles de la Cour de cassation », *RTD civ.* 1974. 488.

⁵⁸J.-M. Sauvé, « Les critères de la qualité de la justice », discours prononcé à l'occasion de la célébration des vingt ans du Tribunal de première instance des Communautés européennes (TPICE), Luxembourg, 25 sept. 2009.

⁵⁹A. Touffait et A. Tunc, « Pour une motivation plus explicite des décisions de justice, notamment celles de la Cour de cassation », art. cit.

⁶⁰V. sur ce sujet N. Molfessis, « Le chameau vu pour la première fois », *JCP G* 2019, 528 et B. Pireyre, « L'arbre qui cachait la forêt », *JCP G* 2019, 656. V. égal. P. Deumier, « Motivation enrichie : bilan et perspectives », *D.* 2017, p. 1783.

forcément subjective de la dangerosité de l'intéressé⁶¹. Les juges eux-mêmes pourraient tirer profit des algorithmes de prédiction des décisions judiciaires afin de prendre conscience de leurs biais ou éventuelles erreurs⁶². Selon Daniel Chen, le fait que certaines décisions soient particulièrement faciles à anticiper -telles les décisions relatives aux demandes d'asile que des chercheurs américains ont pu prédire facilement à partir du nom du juge et de la nationalité du demandeur⁶³- montrerait que les décisions ne sont pas suffisamment mûrement réfléchies et fondées sur les particularités de chaque demande. La confrontation aux algorithmes prédictifs pourrait ainsi avoir pour effet d'inciter les juges à rendre de meilleures décisions⁶⁴. En outre, au contraire de l'humain qui statue de manière partielle et peut difficilement progresser, l'algorithme peut être amélioré, rendu plus équitable, plus cohérent, plus transparent. Lorsqu'un biais ethnique, sexiste, etc., est détecté dans un algorithme, typiquement parce que ce dernier a été entraîné sur des données biaisées, on peut « forcer » (même si cela reste complexe) la correction de ce biais.

On pourrait par ailleurs arguer que le fait que les algorithmes d'apprentissage ne puissent expliquer leur choix est finalement négligeable au regard du caractère parfois opaque et peu intelligible des décisions judiciaires. Alors qu'une décision judiciaire n'est pas expliquée au-delà de sa motivation officielle, le code d'un logiciel et les données sur lesquelles se sont basés ses choix peuvent être rendus publics. L'exemple d'APB, l'algorithme décidant de l'affectation des élèves dans l'enseignement supérieur, est, à cet égard, éclairant⁶⁵. Alors que cette affectation se déroulait autrefois dans un arbitraire relatif, le processus est devenu plus équitable lorsque cette tâche a été donnée à un algorithme. Des demandes d'associations ayant conduit le gouvernement à publier son code, le caractère désormais public de l'algorithme a permis aux citoyens de s'appropriier le sujet en critiquant, par exemple, le paramétrage d'APB, choisi (par des humains, rappelons-le) pour prendre en compte la hiérarchie des choix des candidats⁶⁶. Tout ceci montre que le système s'améliore dans la transparence. Si le Conseil d'Etat a finalement décidé que les caractéristiques de l'algorithme de Parcoursup n'avaient pas à être divulgués aux syndicats étudiants, il en a profité pour souligner que les textes les plus récents imposaient une telle divulgation⁶⁷.

B. Les limites des algorithmes

L'attractivité des algorithmes ne doit pas faire oublier que les questions appréhendées par la justice sont parfois d'une telle complexité que le recours aux machines ne peut qu'être d'une utilité limitée, en tout cas dans l'état actuel de la technologie. Si l'on peut trouver en pratique des contentieux répétitifs (infractions au Code de la route, chèques sans provision), une

⁶¹J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, « Human Decisions and Machine Predictions », *NBER Working Paper* 2017, n° 23180.

⁶²A. Chen, « How artificial intelligence can help us make judges less biased », *The Verge*, 17 janv. 2019. Les algorithmes pourraient en effet permettre de mettre davantage en évidence les problèmes de discrimination : v. J. Kleinberg, J. Ludwig, S. Mullainathan, C. R. Sunstein, « Discrimination in the Age of Algorithms », *Journal of Legal Analysis* 2018, vol. 10, <https://doi.org/10.1093/jla/laz001>

⁶³D. L. Chen, M. Dunn, L. Sagun, H. Sirin, « Early predictability of asylum court decisions », *TSE Working Paper* 2017, n° 17-781.

⁶⁴A. Chen, « How artificial intelligence can help us make judges less biased », préc.

⁶⁵V. sur ce sujet, le rapport rédigé par la Cour des comptes, « Admission post-bac et accès à l'enseignement supérieur, Un dispositif contesté à réformer », oct. 2017. V., D. Restrepo-Amariles, « Le droit algorithmique », dans cet ouvrage.

⁶⁶E. Untermaier-Kerléo, « Les nouveaux visages de la décision administrative : d'une administration assistée à une administration automatisée », *JCP Adm.* 2018. Étude 2339.

⁶⁷Conseil d'État, 12 juin 2019, n° 427916 et 427919. Le Conseil d'Etat a souligné que le décret du 26 mars 2019, postérieur au litige, oblige les établissements d'enseignement supérieur à publier les critères généraux utilisés dans leur procédure de sélection.

grande partie des contentieux présentent des éléments de singularité. Or la singularité des affaires à juger pose aux algorithmes des problèmes bien réels. Les algorithmes se basent sur des modèles (mathématiques ou algorithmiques). Ils proposent des solutions dans un cadre précis, des limites données. Les techniques d'apprentissage automatique sont bien adaptées pour des situations récurrentes, comme, par exemple, le traitement de contraventions de stationnement ou la gestion des conflits sur une plateforme de e-commerce. En revanche, dès que l'on s'écarte d'un cadre bien délimité, les réponses proposées perdent leur sens. L'apprentissage automatique n'est pas adapté pour les situations rares, les conflits commerciaux inhabituels, les affaires criminelles qui sont à chaque fois uniques. Si l'on souhaitait permettre à une machine d'appréhender des affaires singulières présentant une certaine complexité, il faudrait alors utiliser d'autres techniques d'intelligence artificielle beaucoup plus générales, résoudre des défis scientifiques aujourd'hui encore hors de notre portée.

Pour autant, il apparaît bien indispensable de recourir aux algorithmes d'apprentissage dès lors que les algorithmes « classiques » ne peuvent fournir de solutions dans tous les cas où nous ne savons expliquer suffisamment précisément comment nous parvenons à telle ou telle appréciation ou telle ou telle décision. Comme nous l'avons vu, il arrive fréquemment, dans la réalité, qu'un jugement humain intervienne de manière spontanée, instinctive, sans correspondre à un cheminement de pensée facilement explicable. Une approche efficace pour l'informatique est donc d'essayer d'imiter les décisions humaines grâce aux algorithmes d'apprentissage. Il reste que l'apprentissage automatique ne donne alors des résultats satisfaisants qu'à certaines conditions. La quantité et la qualité des données pour ces techniques d'apprentissage sont des prérequis absolument majeurs. Il faut de grandes quantités de données pour entraîner les algorithmes. Si les données sont de mauvaise qualité, mal retranscrites ou mal annotées, l'algorithme peut avoir des résultats décevants⁶⁸. Or les données humaines nécessaires pour faire travailler des algorithmes d'apprentissage sont, par nature, contradictoires, biaisées, imprécises, incomplètes. L'imprécision et l'incomplétude des données peuvent alors conduire les algorithmes sur des fausses pistes et leur faire accumuler les erreurs. De fait, le problème posé par la qualité des données se révèle particulièrement aigu en matière judiciaire. En particulier, les données peuvent se révéler contradictoires : si l'on compile les jugements de première instance, par exemple en matière prud'homale, on trouvera un grand nombre de variations et de contradictions dans les décisions adoptées. En outre, les données humaines comportent toutes sortes de biais : les algorithmes d'apprentissage vont travailler à partir de données qui comportent des discriminations, par exemple. Si on n'y prend pas garde, les algorithmes vont reproduire ces discriminations, voire les accentuer avec leur disposition naturelle à systématiser leurs choix. Certes, on peut essayer de détecter algorithmiquement ces biais et les corriger, mais cela complique considérablement le problème.

Croire qu'une décision algorithmique est juste par définition est une erreur commune, mais bien réelle. Un algorithme peut être mal conçu ou s'appuyer sur des données de mauvaise qualité et biaisées. Les algorithmes ne se comportent pas bien par nature. Il est intéressant d'évoquer ici l'étude réalisée par Pro Publica aux États-Unis, qui a montré par exemple que l'évaluation, par le logiciel COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), du risque de récidiver en commettant une nouvelle infraction dans un délai de deux ans avait un biais ethnique défavorable aux noirs⁶⁹. À partir de l'étude d'un

⁶⁸S. Abiteboul, V. Peugeot, *Terra data, qu'allons-nous faire des données numériques ?*, Paris, Le Pommier, 2018.

⁶⁹J. Angwin, J. Larson, S. Mattu, L. Kirchner, « Machine Bias », Pro Publica, 23 mai 2016. V. égal. C. O'Neil, *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016. V.

échantillon, Pro Publica est arrivé à la conclusion que l'algorithme formule une prédiction fiable dans seulement 65 % des cas. L'étude a également montré que lorsque l'algorithme commet des erreurs, celles-ci consistent le plus souvent à exagérer le risque de récidive pour les noirs et le sous-estimer pour les blancs. Plus précisément, les cas dans lesquels la machine identifie à tort un défendeur comme récidiviste potentiel concernent deux fois plus souvent les prévenus noirs que les prévenus blancs. Pourquoi un tel résultat ? Bien que COMPAS soit un logiciel protégé dont les caractéristiques ne sont pas divulguées, on sait qu'il repose sur 137 variables dont la couleur de peau ne fait pas partie. Le résultat a alors été expliqué par le fait que le logiciel tient compte de données relatives à la famille et aux relations du défendeur ainsi qu'à sa catégorie socioprofessionnelle, ce qui explique sans doute que les noirs, qui évoluent dans des environnements moins favorisés, soient discriminés. Ce cas particulier est intéressant parce que plusieurs critères d'équité étaient avancés et que Kleinberg a montré qu'on ne pouvait tous les satisfaire dans le même temps⁷⁰. On a aussi montré par la suite que le logiciel n'était finalement pas de bonne qualité. Tant la fiabilité de COMPAS que les conclusions de l'étude Pro Publica ont, en effet, été contestées. Une étude publiée en 2018⁷¹ a montré que les prédictions formulées par des humains choisis au hasard à partir de 7 variables (et non 137) présentaient un taux de fiabilité de 62 %, quasi-analogue à celui de COMPAS (65%). Elle a également conclu que les erreurs commises par les humains conduisaient à surévaluer le risque de récidive pour les noirs et le sous-estimer pour les blancs, dans une proportion analogue. Pourtant, les humains interrogés n'étaient informés ni de la couleur de peau des intéressés, ni de leur origine socioprofessionnelle⁷². L'exemple montre en tout cas que, s'il est simple de concevoir un algorithme, il est bien plus complexe d'en concevoir un de qualité, *a fortiori* pour résoudre des problèmes de la complexité de ceux qui se posent dans le cadre de la justice.

Enfin, il est incontestable que les techniques d'apprentissage automatique ne permettent pas d'expliquer précisément leurs résultats même si des recherches essaient de développer des techniques plus explicatives. Or, dans le cadre de la justice, le recours aux algorithmes d'aide à la décision n'est le plus souvent acceptable que si ceux-ci se révèlent suffisamment transparents, robustes et prévisibles pour ceux qui les programment⁷³. Les décisions qui ont des impacts sérieux sur des vies humaines doivent pouvoir être expliquées. Cependant, un réseau neuronal profond peut être entraîné sur un gros volume de données, ce qui permet de dégager un grand nombre de paramètres qui vont guider les décisions futures. Mais ces caractéristiques n'ont pas de sens pour des humains et ne peuvent donc pas être utilisées comme explications. Un juge serait nécessairement, et à juste titre, réticent à prononcer un jugement à partir de la suggestion d'une « boîte noire » qui n'explique pas ses choix.

Dans l'état actuel de la technologie, nous avons donc de sérieuses raisons de douter des décisions que pourraient prendre ou proposer des algorithmes dans des cas complexes. Même

EUBANKS, *Automating Inequality, How High-Tech Tools Profile, Police, and Punish the Poor*, St Martin's Press, 2018.

⁷⁰J. Kleinberg, « On Algorithms and fairness », Séminaire au Collège de France, 16 janv. 2018, <https://www.college-de-france.fr/site/claire-mathieu/seminar-2018-01-16-11h00.htm>. V. égal. J. Kleinberg, S. Mullainathan, M. Raghavan, « Inherent Trade-Offs in the Fair Determination of Risk Scores », *Innovations in Theoretical Computer Science* 2017, vol. 67; S. Tolan, « Fair and Unbiased Algorithmic Decision Making : Current State and Future Challenges », *JRC Digital Economy Working Paper* 2018-10.

⁷¹J. Dressel et H. Farid, « The Accuracy, Fairness and Limits of Predicting Recidivism », *Science Advances* 17 janv. 2018, vol. 4, n° 1.

⁷²Les 7 éléments divulgués aux personnes interrogées étaient : le genre, l'âge, la nature de l'infraction commise, les condamnations antérieures, les accusations antérieures de délits et crimes et les délits et crimes commis lorsque l'intéressé était mineur.

⁷³N. Bostrom et E. Yudkowsky, « The ethics of artificial intelligence », in *The Cambridge Handbook of Artificial Intelligence*, Cambridge Univ. Press, 2014, p. 316-334.

si la justice humaine est loin d'être parfaite, on peut douter qu'une justice algorithmique représente un progrès dans de tels cas. Et dans tous les cas, la perspective dans laquelle la machine prendrait la place de l'humain pour rendre la justice reste fort lointaine.

C. La justice automatisée, une réalité lointaine

Malgré les limites évoquées, rejeter de manière globale tout recours aux algorithmes dans le processus judiciaire serait sans doute une erreur. Il convient de trouver une forme d'équilibre dans lequel la technologie serait utilisée au mieux. L'on peut, par exemple, souhaiter que des outils intelligents permettant d'assister les juges seront à l'avenir adoptés par l'institution judiciaire, qu'il s'agisse du robot assistant qui convoque les parties ou produit des documents de manière automatisée, de l'algorithme capable, en quelques secondes, de rassembler toute l'information pertinente sur un point de droit donné, ou de la machine qui peut analyser le contenu d'un dossier et en extraire les éléments intéressants. L'application « Duplex », récemment lancée par Google, est par exemple en mesure de fixer des rendez-vous, de passer des coups de téléphone et de rédiger des documents simples. Malheureusement, pour l'instant, et malgré les efforts de modernisation de la justice dont témoigne la loi de programmation votée en mars 2019, la mise à la disposition des juges d'outils sophistiqués ou intelligents ne paraît pas envisagée. Il n'est par exemple pas prévu de développer des applications permettant d'automatiser les tâches répétitives ou la production de certains actes, tels les actes de saisine. Pour le reste, si le recours à la technologie dans le cadre de tâches répétitives et administratives se défend facilement, l'usage, par les juges, d'outils d'aide à la décision suscite une méfiance compréhensible. Les possibilités sont, pour l'heure, réduites. En droit français, les juges n'ont qu'une faculté limitée de pratiquer ce que les américains appellent l'« *evidence-based sentencing* », soit le recours aux outils du Big Data pour apprécier le comportement des prévenus, tel le risque de récidive évalué par le logiciel COMPAS⁷⁴. En effet, l'actuel article 47 de la loi « Informatique et libertés » pose le principe selon lequel « aucune décision de justice impliquant une appréciation sur le comportement d'une personne ne peut avoir pour fondement un traitement automatisé de données à caractère personnel destiné à évaluer certains aspects de la personnalité de cette personne⁷⁵ ». Cet article est à rapprocher de l'article 4-3 introduit par la loi n°2019-222 du 23 mars 2019 dans la loi n°2016-1547 du 18 novembre 2016 de modernisation de la justice du XXI^e siècle, qui prévoit que les services en ligne de conciliation, médiation ou arbitrage « ne peuvent avoir pour seul fondement un traitement algorithmique ou automatisé de données à caractère personnel ». Il en découle que le recours aux algorithmes pour apprécier les comportements humains n'est pas entièrement interdit, mais que les résultats des analyses provenant de traitements automatisés ne peuvent constituer le seul fondement de la décision, comme l'a, du reste, jugé la Cour Suprême du Wisconsin à propos de COMPAS⁷⁶. Faudrait-il aller plus loin, clarifier la législation sur ce point, voire prévoir des garanties supplémentaires? En réalité, la mise en place de tels outils au sein des tribunaux français ne semble pas envisagée pour l'instant. Tout dépend, pour le reste, de la fiabilité et de la pertinence des algorithmes concernés et de la qualité des données qu'ils traitent. Il serait certainement dommage que les tribunaux soient privés d'outils pouvant les éclairer s'il est avéré que ces outils sont en mesure de fournir une information fiable, pertinente et neutre. Pour le dire autrement, de tels algorithmes pourraient constituer une garantie, et non un danger, pour les justiciables à la condition d'être

⁷⁴A. Garapon et J. Lassègue, *Justice digitale*, Paris, PUF, 2018, p. 255-256.

⁷⁵Art. 47, al. 1^{er} de la loi n° 78-17 du 6 janv. 1978 relative à l'informatique, aux fichiers et aux libertés.

⁷⁶ *State c/ Loomis*, 881 N.W.2d 749 (Wis. 2016), *Harvard Law Review* 2017, vol. 130, p. 1530. V. F. G'sell, "Les décisions algorithmiques", dans cet ouvrage.

véritablement des outils de qualité pouvant servir à corroborer ou infirmer les autres éléments à la disposition du juge.

Se pose, par ailleurs, la question de la « justice prédictive⁷⁷ », à savoir la possibilité, à partir notamment des données relatives aux décisions antérieurement rendues, de prédire ce que pourrait être la décision d'un tribunal. Le terme peut induire en erreur car il ne s'agit pas, à proprement parler, de « prédire » ce que serait la décision de justice. La variance dans les décisions de plusieurs juges ou celle dans les décisions d'un seul juge sur des cas relativement voisins mettent en évidence qu'il n'est pas question de véritablement prédire une telle décision. Le but est plus modestement de réunir des décisions sur des cas voisins du cas examiné et de quantifier l'aléa⁷⁸. Il est ainsi possible d'évaluer les chances de succès d'une procédure, le quantum d'une future condamnation, voire d'élaborer des modèles capables de reproduire l'éventail des décisions judiciaires rendues sur un point donné⁷⁹. En France, le débat sur ce sujet a été largement alimenté par la politique de diffusion des décisions de justice entamée par la loi n° 2016-1321 du 7 octobre 2016 sur la République numérique qui a prévu la mise à disposition gratuite de toutes les données émanant des administrations publiques et donc des décisions de toutes les juridictions⁸⁰. La loi n° 2019-222 de programmation de la justice adoptée le 23 mars 2019 a légèrement modifié l'article L. 111-13 du Code de l'organisation judiciaire qui prévoit désormais que les décisions rendues par les juridictions françaises seront mises à la disposition du public à titre gratuit sous forme électronique, ce qui permettra de bénéficier de larges bases de données sur lesquelles travailleront des algorithmes d'apprentissage. Cette politique d'open data des décisions de justice a suscité un large débat quant à la protection de la vie privée des justiciables, des magistrats, voire même des tiers. Dans ce contexte, la loi du 23 mars 2019 a exclu toute possibilité de profilage nominatif des juges en disposant que « les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées », sous peine de sanctions pénales⁸¹. Ce texte traduit non seulement le souhait de protéger les magistrats mais aussi une forme de réticence face au risque de mise en évidence du manque de qualité dans les décisions de juges humains, alors même qu'il serait souhaitable d'améliorer cette qualité.

Dans tous les cas, cette nouvelle disposition exclut la production de toute statistique nominative relative aux juges mais n'empêche pas l'analyse des décisions des tribunaux et d'éventuelles prédictions. Si l'on suit la lettre du texte, qui est de nature pénale, donc d'interprétation stricte, il est non seulement possible de produire des prédictions des décisions à venir des juridictions à la lumière des décisions passées mais aussi de réaliser des comparaisons entre les juridictions⁸². En cela, la technologie constituera un outil d'aide à la décision utile aux justiciables, à l'image du *justice bot* canadien précédemment évoqué. Elle amènera très certainement les parties à transiger plus souvent en se fondant sur l'évaluation

⁷⁷S. Lebreton-Derrien, « La justice prédictive. Introduction à une justice “simplement” virtuelle », *Archives Phil. dr.* 2018, t. 60, *La justice prédictive*, p. 14 ; Y. Ménéceur, « Intelligence artificielle. Quel avenir pour la “justice prédictive” ? Enjeux et limites des algorithmes d'anticipation des décisions de justice », *JCP* 2018. Doctr. 190 ; F. Rouvière, « La justice prédictive : version moderne de la boule de cristal », *RTD civ.* 2017. 527 ; A. Garapon, « Justice. Les enjeux de la justice prédictive », *JCP* 2017. Doctr. 31.

⁷⁸Ce que montre précisément un récent rapport : v. L. Godefroy, F. Lebaron, J. Lévy-Véhel, *Comment la numérique transforme le droit et la justice...*, *op. cit.*, spéc. p. 7-8.

⁷⁹Th. Kirat et M. Sweeney, « Une comparaison d'application de « justice prédictive », le cas du contentieux de l'indemnisation du licenciement abusif », *JCP G* 2019, suppl. au n°44-45.

⁸⁰Art. L. 111-13 COJ.

⁸¹Art. L. 111-13, al. 3 COJ.

⁸²F. G'ssell, « Predicting court's decisions is lawful in France and will remain so », *Actualités du droit*, 2 juillet 2019.

produite par la machine⁸³. Elle permettra également de suivre et d'évaluer l'activité des tribunaux d'une manière à ce jour inédite, ce qui ira de pair avec l'évaluation des performances des algorithmes utilisés, dès lors qu'il est souhaitable d'exiger davantage des algorithmes, dont le code peut être amélioré, que des humains dont il nous faut bien tolérer les erreurs. Dans l'ensemble, les outils prédictifs pourront certainement offrir aux tribunaux une source d'information précieuse qui leur permettra d'adapter et rationaliser leur action⁸⁴. Il est, par ailleurs, possible que l'évolution technologique conduise à exiger du législateur qu'il précise davantage le droit positif : les algorithmes sont, en effet, plus à l'aise avec des règles précises, plus formalisées. La technologie nous mènera donc peut-être un jour à un droit plus lisible et moins incertain.

Pourra-t-on, par la suite, imaginer que des machines proposent, à l'image d'Aleksei, des solutions à des questions juridiques, donc le contenu d'un jugement ? Ceci n'est pas exclu. L'on pourrait imaginer des machines chargées de suggérer des décisions au juge, voire même des machines chargées de prendre des décisions simples comme l'envisage, à cette heure, le gouvernement estonien. Cependant, nous avons vu que les outils actuels disposent de capacités relativement limitées en matière de raisonnement juridique, *a fortiori* dans le contexte du droit français. Force est donc de reconnaître qu'il existe aujourd'hui une distance considérable entre les fonctions d'un juge en France et ce que les algorithmes savent réaliser. Il ne faudrait toutefois pas en conclure que cet état de fait ne changera pas. Le temps nous a appris à être prudent. Des activités qui, nous le pensions, étaient irréalisables par des algorithmes (les échecs, le Go) sont devenues réalisables. D'autres, comme la traduction automatique, semblent pratiquement l'être. Devant la rapidité des progrès, il est difficile de se représenter les capacités des machines dans dix ans, dans un siècle ou dans dix ! C'est ainsi qu'il est délicat d'apporter une réponse définitive à la question de savoir si des logiciels pourraient un jour proposer des décisions qui se révéleraient tout aussi pertinentes que celles d'un juge. Même si l'on ignore aujourd'hui si de tels algorithmes pourraient être conçus, aucune base scientifique ne permet d'affirmer que c'est impossible.

Si donc il existait un jour un algorithme capable d'élaborer une décision judiciaire, faudrait-il emprunter cette voie ? Après tout, l'acte de juger autrui est une responsabilité colossale, voire, en un sens, inhumaine. Et bien que les humains acceptent d'assumer cette responsabilité par nécessité, l'on pourrait admettre qu'il est préférable de les en décharger, dès lors que cela est possible, sur des logiciels commettant moins d'erreurs. D'un autre côté, l'on pourrait soutenir que la tâche de juger les humains est par essence humaine : une telle responsabilité reviendrait par principe aux humains malgré leur subjectivité et les erreurs qu'ils peuvent commettre. Laisser la responsabilité de la justice à des algorithmes serait abandonner une part de notre humanité. En 1976, Joseph Weizenbaum, créateur d'Eliza, l'un des tout premiers agents conversationnels, affirmait avec force que les décisions importantes ne doivent pas être laissées aux machines, à qui ces qualités humaines que sont la compassion et la sagesse font cruellement défaut⁸⁵. Sans doute faudrait-il poser la question de manière plus précise et nuancée, en distinguant selon les décisions et leur contexte. Il reste qu'une telle question, au-delà de son caractère certainement trop simpliste, relève à ce jour de la science-fiction dès lors que nous sommes loin de savoir aujourd'hui concevoir des algorithmes capables de remplacer les juges de manière pertinente et satisfaisante.

⁸³F. G'sell, « L'automatisation des décisions de justice, jusqu'où ? », *Ann. Mines – Enjeux Numériques*, sept. 2018, vol. 3.

⁸⁴P. Deumier, « La justice prédictive et les sources du droit : la jurisprudence du fond », *Archives Phil. dr.* 2018, t. 60, *La justice prédictive*, p. 49.

⁸⁵J. Weizenbaum, *Computer Power and Human Reason : from Judgement to Calculation*, W. H. Freeman and Company, 1976.

* *
*

Si l'on résume la situation à ce jour, il apparaît que les outils technologiques actuellement en développement vont venir améliorer la situation actuelle de la justice. Les tribunaux devraient être, sous peu, mieux équipés et formés au numérique. L'accès à la jurisprudence devrait être généralisé grâce à la politique d'open data. L'analyse de données pratiquée sur des décisions de justice largement disponibles permettra d'aider les justiciables et d'évaluer, de manière inédite, l'activité des tribunaux et la qualité de la justice. L'informatique va donc permettre d'améliorer le fonctionnement de la justice avec des règlements plus efficaces et plus rapides des différends, des réductions de coûts pour le justiciable comme pour la société, etc., même si les limites de la technologie actuelle sont évidentes quand des décisions complexes sont envisagées. Dans le même temps, l'évolution en cours est souvent surtout présentée et ressentie comme une source d'économie de ressources pour la collectivité, et ce, dans un contexte de pénurie. Cela explique peut-être, au moins en partie, les réactions de refus. Les réticences devant les changements tiennent aussi beaucoup à l'utilisation d'outils mal adaptés et au manque de formation de personnels qui sont tenus de changer leurs habitudes de travail sans comprendre la technologie sous-jacente et sans formation pour maîtriser les nouveaux outils. Il faut donc œuvrer à améliorer les outils informatiques au service de la justice par la technique, à mieux préparer les personnels de justice à leur utilisation par la formation, et à combattre les sentiments de perte de contrôle et de frustration par la pédagogie.