

L'éthique de la justice prédictive

Par Louis LARRET-CHAHINE

Cofondateur et directeur général de Predictice

Agir avec éthique, c'est se conformer à un ensemble de principes moraux élaborés pour protéger un idéal perçu comme universel. L'idéal de justice, qui est un des socles du pacte social d'une société démocratique, entre dans cette catégorie.

Cependant, l'idée de justice n'est pas tout à fait de ce monde et n'a pas à l'être. Les professionnels du droit, qui sont en prise quotidienne avec elle, le savent et ne mélangent pas l'idéal de justice avec les querelles juridiques ou politiques qui agitent le service public de la justice française.

Il est donc indispensable de dissocier deux impératifs. D'un côté, il y a l'idéal de justice, presque divin, et en réalité intouchable par la technologie – sauf dans certaines œuvres littéraires qui restent fort heureusement de simples fantasmes dystopiques. De l'autre, il y a l'objectif d'un service public efficient, intelligible et pragmatique pour le citoyen, nécessaire au maintien du pacte social. C'est celui-ci qui est largement perfectible ; et c'est uniquement en son sein que viendra l'hybridation fertile entre l'intelligence artificielle et le droit.

À écouter certains cris d'orfraie, l'irruption de la technologie dans le monde du droit aurait érodé la transcendance du procès ou de la justice. Mais si la justice a magnifiquement résisté à la mort de Dieu, elle résistera tout aussi bien à celle de la numérisation de la procédure ou à la réduction de l'aléa dans les contentieux sériels, par ailleurs déjà très largement barémisés.

À en écouter d'autres, l'époque vivrait une alliance entre juristes et ingénieurs dont un des buts, inavouable, serait d'automatiser la justice. Rien n'est plus faux. Les enjeux éthiques sont au cœur du débat depuis le premier jour, comme le montrent les multiples initiatives, collaborations et projets de recherche lancés sur ce thème. L'objectif a toujours été de contribuer à l'amélioration de la performance des professionnels du droit et à la transparence du service public de la justice, dans le respect de règles éthiques qui demeurent indispensables.

Dans ce contexte, la justice prédictive est devenue, à une vitesse fulgurante, un point central des débats du monde juridique. Amateurs comme détracteurs, tous ont eu leur mot à dire face au phénomène d'appropriation de ces outils d'aide à la décision par les professionnels du droit.

Mais quels que soient les préjugés, un point fait l'unanimité : la révolution de la justice prédictive – tant cette technologie lance un défi à la pratique de la justice – ne pourra pas s'achever sans le respect d'impératifs éthiques et déontologiques, dont certains doivent être renouvelés du fait même des possibilités qu'ouvre la technologie.

Cet article tente une synthèse des problématiques éthiques que la construction et le déploiement de la justice prédictive ont soulevées, mais aussi de la façon dont ces défis ont été abordés et résolus jusqu'à présent.

Les trois principaux points de croisement constatés entre l'éthique et la justice prédictive seront traités de façon chronologique : la conception des outils d'abord, le traitement des données ensuite, puis, pour finir, l'impact sur la norme et le jugement.

L'éthique dans la conception des technologies de justice prédictive

Inintelligibilité, erreur de conception : le manque de compréhension induit un syndrome de la boîte noire, en particulier dans le monde du droit, longtemps resté étanche aux technologies. Ces inquiétudes, légitimes, conduisent naturellement à proposer des solutions comme l'ouverture du code informatique ou la création d'une autorité de régulation.

Traitement automatique du langage juridique : cherchez l'erreur

Le manque d'intelligibilité des technologies induit de la méfiance. À la façon de certaines inventions qui étaient prises pour de la magie à leur début, comme ce fut le cas pour le cinématographe de Louis Lumière, l'enjeu est ici de faire un effort de pédagogie.

Réaliser des opérations statistiques sur un grand volume de décisions de justice nécessite l'utilisation du machine learning, ou apprentissage automatique. Le machine learning est un champ d'étude de l'intelligence artificielle avec laquelle il se confond parfois dans le discours adressé au grand public. Il désigne la conception, l'analyse, le développement et la mise en œuvre de méthodes permettant à un programme informatique d'évoluer par un processus systématique. Le *natural language processing* est une branche du machine learning dont le but est de permettre l'analyse, la compréhension, voire la création du langage humain. Ces technologies ont d'abord été déployées dans certains secteurs industriels, comme l'aéronautique, sans soulever des questions éthiques spécifiques.

Le monde du droit, resté longtemps étanche aux dernières évolutions technologiques, a moins bien réagi. Les professionnels du droit, d'abord peu concernés et formés à ces technologies, n'ont pendant longtemps pas fait d'effort d'adaptation. En conséquence, certains voient dans ces technologies une « boîte noire » dont le fonctionnement serait incompréhensible ou imprévisible. Les compétences pour démonter et remonter ces logiciels de leurs mains, mais aussi pour les auditer ou les contrôler leur manquent. Cette problématique d'intelligibilité induit généralement une inquiétude : celle de l'erreur de conception.

La crainte d'une erreur dans le code des outils de justice prédictive doit être surmontée par de la transparence : des équipes entières testent en permanence les solutions, généralement construites avec des briques qui sont rendues publiques par d'autres entreprises, comme ce fut le cas avec Syntaxnet qui a été ouvert par Google.

En outre, s'agissant du traitement du langage naturel, une fiabilité de 100 % est difficile à concevoir, même d'un point de vue théorique. En effet, le langage évoluant rapidement (et s'agissant des décisions de justice, les concepts juridiques aussi), les algorithmes peuvent toujours faire une erreur au moment de résoudre une ambiguïté syntaxique, même s'ils ont été parfaitement construits. En d'autres termes, s'appuyant sur une matière évolutive et ayant parfois plusieurs significations, les algorithmes ne comprennent pas toujours parfaitement le langage ou le sens caché qu'il revêt – c'est le cas avec des doubles négations particulièrement complexes dans les décisions de justice.

Face au manque d'intelligibilité et au risque d'erreur, les solutions sont limitées : ouvrir les algorithmes ou créer un organisme de contrôle ou de certification.

Transparence et régulation : solutions alternatives ou cumulatives ?

La question de la transparence des algorithmes est épineuse. Construire un algorithme nécessite de réunir une équipe de talents, et de les faire travailler ensemble pendant plusieurs mois, et même plusieurs années. Le processus est coûteux, mais être propriétaire de sa technologie représente *in fine* un avantage concurrentiel.

Certains des GAFA (Google, Apple, Facebook, Amazon) rendent libres une partie de leurs algorithmes, notamment pour stimuler un écosystème d'innovation et créer une logique de softpower autour de leur marque. Néanmoins, ils gardent toujours le contrôle sur les données qu'ils ont en leur possession, préservant un avantage concurrentiel impossible à compenser.

Alors, s'agissant des legaltechs, faut-il rendre le code source accessible à tous ? Selon certains auteurs, comme Michaël Benesty⁽¹⁾, la réponse est oui, au nom d'un impératif de confiance. Mais dans un monde où la donnée est libre – c'est le cas des décisions de justice –, les logiciels ne peuvent l'être complètement, au risque de tuer les avantages compétitifs acquis en construisant la technologie pour les exploiter. Contraindre les entreprises à mettre leurs algorithmes à la disposition de tous reviendrait à bloquer les cycles d'innovation.

Une solution alternative existe. Une autorité de régulation nationale indépendante pourrait avoir accès à l'intégralité du code source provenant de chaque projet utilisant des décisions de justice, afin de garantir l'intégrité des solutions développées tout en les protégeant.

Cette régulation des algorithmes n'est d'ailleurs pas inédite. Les algorithmes de trading à haute fréquence qui sont utilisés dans le domaine de la finance, par exemple, sont régulés par les directives et régulations européennes MiFID et MiFIR⁽²⁾.

Afin de ne pas créer une nouvelle autorité administrative indépendante, la responsabilité pourrait être confiée à la CNIL. Certaines questions devront alors être tranchées. Le code doit-il être mis à disposition une unique fois pour obtenir un agrément lors de son implémentation ? La transmission doit-elle être faite à une certaine fréquence, en continu ou encore sur demande ? L'algorithme nécessite-t-il un nouvel agrément à chaque mise à jour ? L'accord doit-il être obtenu *a priori* ou *a posteriori* ? Là est le problème de la régulation d'une technologie qui évolue en permanence. Il est parfois difficile de la superviser sans restreindre l'innovation⁽³⁾, et le processus sera gourmand en ressources.

En attendant une régulation institutionnelle, des initiatives spontanées ont vu le jour. Une charte éthique a été proposée par l'ADIJ et l'association Open Law⁽⁴⁾. Peu contraignante, elle a permis de rassembler la plupart des acteurs de la legaltech, mais c'est un premier pas qu'il faut encourager. De façon plus spécifique, un Comité éthique et scientifique de la justice prédictive, composé de neuf professionnels du droit, professeurs et praticiens, a été créé en 2017⁽⁵⁾. Ce comité, dont les travaux sont suivis par la Clinique de l'École de Droit de Sciences Po Paris, mène une réflexion éthique sur le fondement des travaux de la société Predictice, auxquels il a accès.

Comme dans toutes les industries exploitant des données, l'intérêt vient d'un croisement entre des algorithmes bien conçus et des données, en l'occurrence des décisions de justice, dont la collecte et le traitement soulèvent également des questions éthiques.

L'éthique dans la collecte et le traitement des décisions de justice

L'avènement des services gratuits dans lesquels le client échange ses informations personnelles contre un service a fait des données personnelles un bien précieux et éthiquement sensible. Les données personnelles des justiciables, et notamment des personnes physiques, le sont tout particulièrement. Cet enjeu se retrouve dans les articles 20 et 21 de la loi pour une République

(1) BENESTY M. (2017), « L'open data et l'open source, des soutiens nécessaires à une justice prédictive fiable ? », *Journal of Open Access to Law*, Vol. 5, n°1.

(2) Cabinet Norton Rose Fulbright (2014), MiFID/MiFIR Series, octobre.

(3) KAY J. (2017), "How do you regulate a self-improving algorithm?", *The Atlantic*, 25 octobre.

(4) <https://www.chartheethique.legal/charte-ethique>

(5) <https://predictice.com/ethics-commitee>

numérique, dans lesquels il est précisé que les décisions et jugements des diverses juridictions françaises seront « mis à la disposition du public à titre gratuit dans le respect de la vie privée des personnes concernées ». Cette disposition soulève deux principales difficultés éthiques : l'anonymisation et la gestion des biais et de la mise à disposition des données.

Anonymisation et ré-identification dans les décisions de justice

En l'absence de publication du décret d'application de la loi pour une République numérique à la date de la rédaction de cet article, il est difficile d'avoir une idée précise de l'étendue de l'anonymisation qui sera mise en place dans le cadre de l'open data. S'agira-t-il d'une simple opération de remplacement de noms de personnes privées par des Monsieur X et des Madame Y, sans apporter aux décisions d'autres modifications ? C'est indéniablement la solution qui, en termes de coût, de faisabilité et d'intérêt pour les professionnels du droit, présente le plus d'avantages. Le même système pourrait être reproduit, sans trop de difficulté ni travestissement de la logique de la loi, sur des informations comme le numéro de téléphone, l'adresse et autres identifiants.

D'un autre côté, une anonymisation plus extensive, notamment du contexte, est possible. Celle-ci, beaucoup plus coûteuse – au vu de l'état de l'art, elle nécessite une relecture manuelle pour garantir un résultat fiable –, réduirait drastiquement l'utilité de la loi et n'empêcherait pas forcément la ré-identification des personnes privées, comme a pu l'expliquer le chercheur Tristan Allard⁽⁶⁾.

Précisons que du point de vue de la compréhension de la justice, mais aussi du point de vue commercial, la ré-identification présente très peu d'intérêt. En outre, l'article 226-19 du Code pénal dispose qu'« est puni de cinq ans d'emprisonnement et de 300 000 euros d'amende (...) le fait, hors les cas prévus par la loi, de mettre ou de conserver en mémoire informatisée des données à caractère personnel concernant des infractions, des condamnations ou des mesures de sûreté ». Une sanction dissuasive est donc déjà prévue pour les impétrants...

La gestion des biais dans les bases judiciaires

Par définition, les algorithmes d'analyse statistique de décisions de justice requièrent des quantités massives de données, dans lesquelles peuvent se cacher des biais. Ces logiciels, comme tout autre algorithme du secteur des big data, se restreignent aux données qu'on leur donne et dont ils se nourrissent. Leur monde est fini, c'est celui de la décision de justice, mais celui-ci n'est pas exempt de biais, qui sont d'ailleurs parfois inconscients.

L'exemple le plus célèbre, s'agissant des biais que peuvent comporter les données d'apprentissage, est celui du logiciel Compas, développé par la société Northpoint dans le but d'anticiper le risque de récidive. Selon la chercheuse Laurel Eckouse, le logiciel reproduirait les préjugés raciaux, amenant à considérer les détenus noirs plus sévèrement⁽⁷⁾.

Plusieurs solutions pourraient être envisagées pour corriger ce type de biais.

La première est la modification du *dataset* pour le faire correspondre à une réalité souhaitée. Cette approche doit être rejetée d'emblée : elle est dangereuse, car elle rompt la confiance dans ces outils d'aide à la décision et ne permettra pas des analyses intéressantes.

Une deuxième possibilité est de pointer du doigt les biais pour que la réalité se corrige progressivement, influant ainsi le *dataset* en cours de constitution, et donc les statistiques qui seront réalisées dans le futur. Un processus d'alerte pourrait ainsi être mis en place : quand un juge est 90 % plus

(6) Tristan Allard, maître de conférences en informatique à l'Université de Rennes, a abordé le sujet de l'anonymisation des décisions de justice lors d'un colloque sur la justice prédictive qui s'est déroulé le 6 avril 2018 à Laval. Une publication est à venir.

(7) <https://www.lebigdata.fr/justice-big-data-prejudice-race-1302>

sévère ou plus clément que l'ensemble de la profession sur un sujet donné, une discussion publique pourrait s'engager pour comprendre les ressorts à l'œuvre derrière le processus décisionnel. Les biais disparaîtraient ainsi progressivement grâce à l'intelligence collective des acteurs et des auxiliaires du service public de la justice, chargés de détecter et de réfléchir aux variations sur le territoire.

Une autre manière de réduire le risque est de refuser de faire travailler les algorithmes sur certains domaines du droit. À la façon de la voiture autonome qui apprend à ne jamais franchir une ligne blanche continue, les algorithmes peuvent être programmés pour ne pas fonctionner sur certaines branches du droit. C'est d'ailleurs la décision qui a été prise par le Comité éthique de la Justice prédictive, s'agissant du fonctionnement de Predictice sur le droit pénal.

Enfin, un autre impératif s'impose : que tous les algorithmes travaillent sur la même base de données. Cela implique d'être très clair sur la provenance et le mode de collecte mais aussi de sanctionner sévèrement les collectes sauvages de données, qui biaisent très fortement les résultats et trompent les utilisateurs. En effet, si des données sont aspirées auprès d'une entreprise ou dans une région donnée, les analyses ne porteront plus sur un échantillon représentatif ou l'exhaustivité des décisions rendues par une juridiction, mais sur une base dont le mode de constitution est flou et la complétude imprécise...

L'éthique dans le fonctionnement de la justice

Une part non négligeable des débats sur la justice prédictive est alimentée par la crainte des effets que la technologie aura sur la justice. Sans tomber dans l'écueil fantasmagorique des juges-robots, concept totalement irréaliste, tant le processus de création d'une décision de justice est complexe et multifactoriel, il faut néanmoins s'interroger sur les impacts que la justice prédictive a sur l'accès au droit et l'homogénéisation des jugements.

Déjudiciarisation : un frein à l'accès au droit ?

Comme ils donnent une idée de ce qu'il sera possible d'obtenir en cas de contentieux, les outils d'analyse statistique des décisions de justice sont une arme de négociation étonnamment efficace, ce qui laisse prédire un bel avenir à la résolution extra-judiciaire des litiges, mais génère beaucoup d'inquiétude sur l'ampleur de la déjudiciarisation à venir.

L'accès à la justice pour les justiciables est centrale : la justice prédictive ne sera jugée positivement que si elle améliore l'accès et le fonctionnement de la justice pour les justiciables. Antoine Garapon et Jean Lassègue l'ont bien exprimé : « Que serait en effet une justice qui ne s'adresserait pas au sentiment de justice, à ce qu'il y a d'humain dans les humains⁽⁸⁾ ? »

Mais accéder à la justice, est-ce nécessairement passer devant une juridiction du fond, avant de faire un appel ou un pourvoi, jusqu'à finalement obtenir son dû – ou non – tout en perdant des années et des sommes phénoménales ? C'est précisément dans la machine judiciaire que les justiciables ont parfois la sensation de ne pas avoir ce « sentiment de justice ».

Dans un contexte où les magistrats sont débordés, où les délais d'attente croissent et où les justiciables ne font guère confiance à la justice, la déjudiciarisation doit nécessairement être envisagée, et de façon un peu plus ambitieuse que par la restriction des procédures d'appel ou la complexité des formulaires Cerfa pour saisir les prud'hommes.

(8) GARAPON A. & LASSÈGUE J. (2018), *Justice digitale*, PUF, p. 134.

L'intérêt des modes alternatifs de résolution de litiges (MARL) pour les justiciables est évident. Plus rapides, moins coûteux, moins traumatisants, ils permettent de trouver des solutions satisfaisantes dans une grande variété de cas. Déjudiciariser ne veut pas dire empêcher l'accès à la justice. Il s'agit au contraire de mieux orienter le flux de litiges, en s'assurant qu'une solution est trouvée quel que soit le procédé de résolution, judiciaire ou extra-judiciaire.

Risque performatif ou création d'un nouveau système de production normative ?

La performativité désigne un phénomène de prophétie auto-réalisatrice, dans lequel le simple fait d'avoir accès à l'information a un impact sur le réel. Il existe plusieurs exemples dans lesquels le fait même d'avoir prévu un événement influe fortement sur la réalisation de ce même événement, notamment en finance. Ce risque semble cependant moins probable s'agissant de la justice prédictive, principalement car l'analyse d'événements dépendant de facteurs irrationnels ou restant en dehors de la décision de justice (la chaleur dans la salle d'audience, la fatigue d'un magistrat, une plaidoirie particulièrement réussie), bref ce qui relève de la sensibilité humaine dans le processus de jugement ne peut pas être pris en compte par les algorithmes. Une étude est menée actuellement sur l'effet performatif par l'Université Paris-Dauphine pour approfondir le sujet.

S'agissant enfin de l'influence supposée des logiciels sur les magistrats (dans l'éventualité où ils y auraient accès), il ne faut pas la surestimer. Ce serait oublier la capacité intellectuelle et le courage des magistrats de supposer que, par crainte du regard des autres, ils modifieraient leurs décisions, d'autant plus que la motivation de leurs jugements continuera de les obliger et de les protéger.

Il est vrai, cependant, que la justice prédictive aura certainement pour effet de créer un renversement des normes juridiques. Le déploiement des solutions de justice prédictive dans le quotidien des professionnels du droit est en effet en train de modifier le concept même de jurisprudence. Grâce à la combinaison entre la justice prédictive et l'open data, les professionnels du droit n'auront plus à se limiter aux quelques litiges portés à la connaissance des hautes juridictions. Il sera possible d'étudier avec une grande finesse les millions de décisions du fond rendues chaque année.

Cette possibilité, vertigineuse, va entraîner la naissance d'un nouveau système juridique, dit isométrique, dans lequel l'importance normative des décisions des hautes juridictions ne disparaîtra pas, mais sera progressivement dissoute dans la mesure où toutes les décisions, de fond comme de hautes juridictions, peuvent se valoir dans une statistique donnant une idée extensive de l'interprétation du droit par les juges.

Ainsi, la technologie de justice prédictive et l'open data sont peut-être, de façon assez ironique, en train de rendre la justice plus transparente, mais aussi plus humaine et plus démocratique, ou du moins isométrique, en donnant une voix normative aux neuf mille magistrats français.