

Séance 5 - Intelligence artificielle, algorithmes et justice à l'heure de l'*open data* et du *big data* (10 février 2021)

Exposants invités : Isabelle Sayn, directrice de recherche au CNRS en sociologie et sciences du droit. Marc Clément, président de chambre au tribunal administratif de Lyon.

Points clefs de la présentation d'Isabelle Sayn

En guise de remarque préliminaire, Isabelle Sayn souligne que l'« **effet de mode** » autour de la justice prédictive est quelque peu retombé récemment. Il y a quelques années, on avait encore du mal à faire la part des choses entre **le discours marketing** et ce qui était **réellement possible et envisageable** à ce sujet. Les outils techniques continuent néanmoins à se développer, à être proposés et même à être commercialisés, et il faut donc continuer à les prendre au sérieux.

Justice, algorithmes et IA : une typologie des utilisations possibles

Isabelle Sayn distingue quatre grands types d'utilisations. Tout d'abord, il y a **(1) l'aide à la recherche d'informations juridiques**, qui existe depuis assez longtemps et qui est beaucoup utilisée par les juristes. Avec ces outils, il est possible d'accéder plus simplement aux textes de loi et aux décisions de justice - en tout cas à celles qui sont recensées dans les bases de données. L'aide à la recherche d'informations juridiques s'est beaucoup développée et est en partie pilotée par les pouvoirs publics, notamment à travers le site *Legifrance*, qui permet d'accéder à un ensemble conséquent de sources juridiques. Les éditeurs privés sont également des acteurs importants, en ce sens où ils permettent d'accéder à ce même type d'informations, en les enrichissant de commentaires et d'annotations. Ensuite, on peut citer **(2) la dématérialisation des procédures**, c'est-à-dire la possibilité de saisir la justice de façon numérique ainsi que d'avoir des échanges de façon dématérialisée. C'est un aspect qui est supposé faciliter l'accès à la justice (sous réserve d'accès à l'outil numérique), et qui est dorénavant rendu obligatoire pour saisir les juges. Un autre aspect concerne **(3) l'aide à la rédaction d'actes juridiques et de contrats**. Plus précisément, les outils techniques relatifs à cette utilisation permettent de rédiger des contrats de façon à s'assurer que les clauses soient à jour vis-à-vis de la dernière jurisprudence.

Aide à la décision et justice prédictive

Enfin, il y a **(4) les outils d'aide à la décision**, qui sont des outils créés à partir de **l'analyse des décisions de justice**. Ici, nous ne sommes plus dans l'accès aux juges via la dématérialisation des procédures, ni même dans l'accès à la doctrine au sens large, mais plutôt dans l'accès aux décisions de justice elles-mêmes et à ce qu'on peut en tirer dans le but d'aider les acteurs **à obtenir ou à rendre une décision de justice**. En particulier, il s'agit de prendre un corpus (*a priori* exhaustif, même si c'est très limité pour le moment) de décisions de justice **comparables**, d'en extraire des informations, et de fournir à partir de celles-ci des informations aux avocats, aux parties, aux magistrats ainsi qu'aux médiateurs, pour les aider dans la construction d'une décision de justice **à venir**. C'est là où on parle de **justice prédictive** ou **prévisible**. L'une des principales questions que pose

cette utilisation est la suivante : quel type d'informations pouvons-nous obtenir et à partir de quel type de décision ?

Ces outils d'aide à la décision, qui sont issus de l'analyse automatisée, concernent également la médiation. L'addition de la procédure dématérialisée d'un côté, et des outils d'aide à la décision de l'autre, a conduit à la construction **de systèmes de résolution en ligne des litiges par la médiation**. Ces modes de résolution sont ainsi enrichis par des informations sur des décisions qui ont été prises dans des situations comparables, pour aider les parties à se mettre d'accord sur telle ou telle solution. Mais, que l'on se situe au niveau de la médiation ou de l'aide à la décision des magistrats, ces pratiques renvoient à d'autres qui existent déjà, comme le recours aux **barèmes**, qui permettent de cadrer le raisonnement notamment lorsqu'il s'agit de *quantums* (e.g. combien indemniser, quelle prestation payer, combien de dommages et intérêts verser, etc.). Ces barèmes, qui peuvent être ou non issus de processus automatisés, sont des outils d'aide à la décision servant à obtenir d'éventuels accords ou à aider les juges dans leurs prises de décision.

Accéder à des connaissances nouvelles

L'analyse automatisée d'un corpus de décisions produit **des connaissances qui n'étaient pas encore accessibles**. Jusqu'à présent, en tant que juriste, on pouvait effectivement connaître la jurisprudence, c'est-à-dire l'ensemble des décisions normatives généralement prises par les juridictions supérieures, celles-ci permettant de savoir quel sens accorder à une norme ou à une règle. Par contre, on avait très peu souvent accès à ce qu'on appelle le contentieux, à savoir cette masse de décisions rendues par les juridictions de premier degré et d'appel, parce que même dans le cas où on avait accès à ces décisions, elles se révélaient trop nombreuses pour qu'une analyse manuelle puisse en être faite. Il est certain qu'à partir du moment où on va savoir ce qu'ont fait des juges dans des décisions comparables, cette connaissance va avoir des effets. L'enjeu principal reste de **déterminer la nature** de ces derniers et la manière dont il est possible de les **anticiper** et de les réguler.

Points clefs de la présentation de Marc Clément

Juridiction administrative et dématérialisation : une évolution fulgurante

Dans le cadre de la juridiction administrative, l'évolution du recours aux outils numériques est assez radicale. Ainsi, de 2000 à 2021, on passe d'une période où on n'utilisait pas internet dans la juridiction et où on ne disposait quasiment pas de bases de données, à une période où **95% des requêtes sont dématérialisées**. Cette évolution reste néanmoins peu étudiée en tant que telle, et les juristes ne sont pas les premiers à faire ces analyses, croyant que ces outils ne changent pas fondamentalement le métier.

Vers la mise en place d'un juge automate ?

Par ailleurs, pour bien comprendre et se positionner sur la question de la justice prédictive, il peut être intéressant de mettre en relation celle-ci avec une « méfiance » qui existe à l'égard du juge. D'après ce dernier point de vue, le juge n'est pas vu comme étant là pour interpréter ou créer du droit, mais plutôt comme un « **juge automate** ». Ce faisant, selon Marc Clément, la mécanique que l'on pourrait mettre en place dans le cadre d'un outil numérique intervient dans l'optique de créer un « juge parfait », se limitant à l'application d'un syllogisme juridique « pur ». La perspective du recours aux systèmes informatiques est donc à placer vis-à-vis d'une telle vision de la justice, celle-ci étant sous-jacente à beaucoup de discours en faveur de ces outils.

Différents niveaux d'automatisation

Tels que présentés par Isabelle Sayn, les systèmes informatiques correspondent à différents niveaux d'automatisation. Du plus simple au plus complexe nous avons : (1) **les bases de données juridiques**, (2) **les bases de données interactives**, (3) **la proposition de contextes juridiques**, (4) **la proposition de solutions juridiques et de scénarios**, et enfin (5) **la solution juridique automatisée**. Cette gradation permet de souligner que les outils juridiques les plus développés concernent le deuxième voire troisième niveau, très peu d'entre eux allant au-delà. Cela peut s'expliquer du fait que la mise à disposition des jugements est **encore assez peu développée et très incomplète** à ce jour en France. Le même constat peut être établi ailleurs en Europe, où on observe que des pays comme les Pays-Bas ou le Danemark ont respectivement mis 4% et 20% de leurs jugements à disposition dans des bases de données. Le travail reste donc largement à faire.

Approches statistiques et approches sémantiques

En ce qui concerne le traitement des données, Marc Clément distingue deux familles d'approches techniques. Premièrement, l'approche dite **statistique** (e.g. machine learning, etc.) consiste dans l'exploitation statistique d'une grande masse de données, le tout en visant une certaine performance. La deuxième approche consiste quant à elle en **une analyse sémantique** des décisions, à partir du **sens du texte**, dans le but de classifier des grandes catégories de décision. Cette distinction met en lumière le fait que nous sommes donc en présence de différents types d'outils, qui ont des niveaux de développement assez variés. Toutefois, jusqu'à présent, la majorité du recours aux outils numériques dans le domaine de la justice a concerné l'utilisation statistique des bases de données.

La motivation d'un jugement peut-elle reposer sur une simple corrélation ?

Il faut faire attention aux statistiques et à leur tendance à mettre en évidence non pas des causalités, mais **des corrélations**. S'agissant du recours aux statistiques dans le domaine du droit, le problème qui surgit rapidement est que le droit ne se contente pas simplement de la corrélation. En plus de cela, il faut une **motivation** du jugement ainsi que des éléments qui permettent de comprendre le raisonnement du juge. L'exercice de la fonction de juge passe par le développement de l'élément fondamental qu'est l'explication, qui représente un facteur d'acceptabilité important. D'après Marc Clément, un jugement ou un arrêt se rapportent toujours à la situation factuelle qui est spécifique à chaque personne. Or, entre la statistique et la situation individuelle, **il y a un écart très important**.

Absence de neutralité et prise de recul

Lorsqu'on parle d'IA, d'algorithmes et de justice, il y a un effet souvent mal perçu, qui est celui de **l'absence de neutralité** des outils. L'utilisation des bases de données constitue déjà aujourd'hui un élément de **modification** de la façon dont le juge va traiter l'information dont il dispose. Ensuite, même si on suppose que l'on puisse développer des outils statistiques qui seraient utiles, on a vu que des systèmes, à l'image de **COMPAS** aux États-Unis, posent des problèmes considérables. En réalité **discriminatoires**, les scores produits par ces outils peuvent être interprétés par le juge comme étant neutres. Il est ainsi difficile pour un juge de disposer de tels outils et d'être en même temps capable de prendre la distance suffisante pour traiter correctement les informations qui en sont issues.

Pour aller plus loin ...

Aletras, N., Tsarapatsanis, D., Preoțiu-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights : A Natural Language Processing perspective. *PeerJ Computer Science*, 2, e93.

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias : There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Le Métayer, D., & Desmoulin-Canselier, S. (2020). *Décider avec les algorithmes—1re ed. : Quelle place pour l'Homme, quelle place pour le droit ?* (1ère ed.). Dalloz.
- Sayn, I. (2019). Connaître la production des juridictions ou prédire les décisions de justice ? *Les Cahiers de la Justice*, 2(2), 229-242.