



HAL
open science

CONNAÎTRE LA PRODUCTION DES JURIDICTIONS OU PREDIRE LES DECISIONS DE JUSTICE ?

Isabelle Sayn

► **To cite this version:**

Isabelle Sayn. CONNAÎTRE LA PRODUCTION DES JURIDICTIONS OU PREDIRE LES DECISIONS DE JUSTICE ?. 2018. halshs-01745381

HAL Id: halshs-01745381

<https://shs.hal.science/halshs-01745381>

Preprint submitted on 28 Mar 2018

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.

CONNAÎTRE LA PRODUCTION DES JURIDICTIONS OU PREDIRE LES DECISIONS DE JUSTICE ?

Isabelle SAYN
Directrice de recherche au CNRS
Centre Max Weber (UMR 5283, CNRS,
Université de Lyon)

(à paraître, Les cahiers de la justice, 2018, Dossier droit et numérique)

Connaissance des activités de justice et connaissance de la production des juridictions. La connaissance des activités non pénales de justice est documentée via le répertoire des affaires civiles qui, à partir d'une nomenclature des affaires, permet depuis longtemps déjà de connaître la production des juridictions. Cette nomenclature liste l'ensemble des demandes susceptibles d'être formées et leur devenir procédural. Les données, codées à l'occasion de l'activité de gestion des greffes, font l'objet d'une remontée systématique auprès de l'administration centrale, permettant ainsi de produire des statistiques sur les décisions prononcées, par nature de contentieux¹. Le mécanisme est comparable s'agissant des affaires pénales, à partir d'une autre nomenclature (NatInf), via l'appli Cassiopée et le Système d'information décisionnel (SID) crée à cet effet. Il permet d'obtenir des statistiques pénales sur les natures d'affaires, les orientations et les décisions, ainsi que sur les auteurs et les victimes.

Ces données ont longtemps alimenté l'annuaire statistique de la justice, remplacé en 2016 par les « références statistiques justice »². Les statistiques ainsi produites permettent de connaître l'issue des procédures mais pas le contenu des décisions, pour lesquels il faut disposer d'une base de décisions et procéder à leur analyse. Des études de ce type viennent ponctuellement compléter ce premier jeu de données. On peut ainsi, par exemple, savoir que les décisions rendues par les JAF fixent des pensions alimentaires d'un montant moyen de 200 euros et d'un montant médian de 266, dont 10% sont inférieures à 80 euros et 10% supérieures à 516 euros³.

Connaître ainsi, de façon plus précise, non plus seulement les activités de justice mais la production des juridictions requiert un investissement lourd. Ce type d'étude suppose avant tout de recueillir un échantillon représentatif de décisions de justice produites dans le champ d'analyse retenu. Cette première étape peut s'avérer assez coûteuse, les décisions du

¹ On peut d'ailleurs imaginer, dans un monde idéal où l'on pourrait faire l'impasse sur les difficultés matérielles de mise en œuvre d'un tel projet, que ces analyses résultent à l'avenir d'une analyse automatisée de la totalité des décisions rendues : l'ensemble des données recueillies via le registre des affaires civiles est en effet contenu dans les décisions.

² <http://www.justice.gouv.fr/statistiques-10054/references-statistiques-justice-12837/>

³ Montant des pensions alimentaires : pratiques et barème, Infostat Justice 116 - Avril 2012, <http://www.justice.gouv.fr/statistiques-10054/infostats-justice-10057/montant-des-pensions-alimentaires-pratiques-et-bareme-23976.html>.

fond n'étant pas immédiatement disponibles, particulièrement s'agissant des décisions de première instance⁴. Elle suppose ensuite de construire un outil d'analyse statistique via l'élaboration préalable d'une grille de lecture de ces décisions et de (re)lire une à une ces décisions pour renseigner cette grille avant de pouvoir, enfin, procéder aux analyses que la base de données ainsi constituée permet.

Parallèlement à la faible appétence des juristes français pour les analyses empiriques du droit, la lourdeur du processus et l'obligation corrélative d'un travail en équipe limitent de fait les analyses produites à quelques contentieux spécifiques, au gré des intérêts des chercheurs qui s'y consacrent ou de ceux du ministère de la justice qui procède également à ce type d'études seul⁵ ou en collaboration avec des équipes de recherche. Ce type d'activité s'est modestement développé dans la période récente, à l'initiative de quelques chercheurs venant d'horizons disciplinaires variés (droit, sociologie, économie).

Connaissance de la production des juridictions, accès aux données et analyse automatisée des décisions. La conjonction de l'open Data des décisions de justice, prévue par la loi du 7 octobre 2016 pour une République numérique, et des technologies d'analyse de la langue, de fouille de données et d'extraction de connaissances devrait permettre de renouveler les pratiques de recherche dans ce domaine et de produire des connaissances nouvelles sur la production des juridictions, soit près de 4 millions de décisions par an (France). Cette évolution constitue un enjeu majeur pour la recherche publique comme pour le monde de la justice et ses usagers : une connaissance beaucoup plus large de la production des juridictions devient accessible, non seulement au regard du nombre de contentieux susceptibles d'être analysés mais peut être aussi au regard de la finesse des analyses produites, le traitement automatisé de ces décisions permettant d'entrevoir des enseignements nouveaux susceptibles d'être tirés de ces analyses.

Les enjeux politiques de ces développements sont importants, au rang desquels se trouve la possibilité de renforcer la confiance des justiciables dans leur justice, parce que celle-ci se donnerait à voir. Mais l'accessibilité à une masse considérable de décisions préalablement anonymisées ou pseudonymisées ne permet pas, à elle seule, un accès à la connaissance de la production de ces juridictions : les décisions doivent être analysées pour fournir des informations utilisables. Des travaux de recherche y contribuent ponctuellement, sur la base de conventions passées avec le service des études et du rapport de la Cour de cassation, mettant à leur disposition telle ou telle extraction de la base JuriCA, dorénavant constituée de plus de 700 000 décisions d'appel. C'est également sur ce terrain que se positionne la legaltech, constituée d'un ensemble de start-up qui se proposent notamment de produire des connaissances à partir des décisions rendues, avant même la mise en œuvre de la loi de 2016.

Les équipes de recherche qui se sont intéressés à la production des décisions de justice et plus largement à la fabrication de ces décisions ont, sur ce terrain, une expérience

⁴ Même si Légifrance, pour les décisions des juridictions supérieures, ou Jurica et Jurinet pour les décisions d'appel, offrent dorénavant une quasi exhaustivité, en matière non pénale. Pour une présentation des différentes bases de données de décisions existantes, leur contenu, et leur disponibilité, voir L. Cadet, Rapport à Madame la garde des Sceaux, ministre de la Justice, L'open data des décisions de justice, Mission d'étude et de préfiguration sur l'ouverture au public des décisions de justice, nov. 2017, p. 61 s. < 049/remise-du-rapport-sur-[open-data-des-decisions-de-justice-31165.html](https://www.ladocumentationfrancaise.fr/etudes/publications/2017/049/remise-du-rapport-sur-
open-data-des-decisions-de-justice-31165.html)>

incontestable. Elle doit être mobilisée, d'autant que ces travaux se sont développés avec le seul objectif de produire des connaissances nouvelles, en mettant en œuvre des méthodes qui permettent, comme il est d'usage dans une démarche scientifique, la mise en débat des résultats proposés.

Ces approches justifient de se donner les moyens, en travaillant avec des chercheurs en informatique, d'une connaissance suffisante des technologies utilisées et des évolutions en cours, afin de mieux savoir ce que l'on peut en attendre. Elles supposent également de pointer les questions théoriques que ces évolutions soulèvent.

L'un d'elles est la qualification que l'on peut donner à la fois à la masse de décisions que l'on se promet de diffuser et aux analyses qui en sont issues. Cet ensemble peut-il prendre une place dans la production normative ? Une autre est liée à l'indétermination du droit : cet ensemble peut-il réduire les conséquences de l'indétermination du droit sur la dispersion des décisions de justice, en permettant une « justice prédictive » capable de réduire l'incertitude, voire d'anticiper les décisions de justice ? Cette question peut être abordée à partir des enseignements tirés de l'usage des barèmes dans la fabrication des décisions de justice : les barèmes constatés issus d'une analyse automatisée des décisions, tels qu'ils sont aujourd'hui proposés par certaines start-up, permettent-ils de distinguer les déterminants légitimes des montants proposés des déterminants qui ne le seraient pas ?

Les lignes qui suivent ne prétendent pas répondre de façon exhaustive à l'ensemble de ces questions ; elles proposent simplement de se focaliser sur quelques points susceptibles de prendre part au débat.

1. Une nouvelle modalité de production normative ?

L'objectif de mettre à la disposition du public l'ensemble des décisions de justice rendues par l'ensemble des juridictions françaises étant posé par la loi, il reste à définir les modalités pratiques de cet open data, modalités sur lesquelles s'est penché le rapport Cadiet. Ce rapport préconise notamment de respecter le périmètre des décisions actuellement diffusables via les greffes⁶ et s'interroge sur les limites de la pseudonymisation⁷. Le principe même de la pseudonymisation, qui rompt le lien entre les parties, quelles qu'elles soient (personnes privées ou publiques, entreprises) et les affaires juridictionnelles dans lesquelles elles ont pu apparaître, montre que les bases de décisions attendues ne constitueront pas une mémoire des espèces ou, autrement dit, une mémoire des histoires individuelles, mais bien une mémoire de la production juridictionnelle.

⁶ Pour une présentation synthétique des dispositions établissant des exceptions à la publicité des décisions rendues

en matière civile et pénale, voir L. Cadiet, L'open data des décisions de justice, préc., Annexe 1 et 2.

⁷ Sur la distinction entre pseudonymisation et anonymisation, voir le Livre blanc sur l'open data jurisprudentiel, Synthèse des travaux du programme Open « Case » Law, Janvier 2017, dans le cadre des travaux de l'association OpenLaw* Le droit ouvert. Dans les deux cas, il s'agit d'éviter toute identification a posteriori, ou bien en supprimant les données identifiantes (anonymisation), ou bien en remplaçant les données identifiantes par des données nouvelles, autorisant ainsi la réversibilité du processus (pseudonymisation). La suppression (ou le remplacement) des principaux éléments identifiants contenus dans les décisions de justice n'écarte pas tout risque de ré-identification ultérieure, par rapprochement avec des données extérieures aux décisions, mais elle le minimise fortement et elle s'ajoute à l'interdiction de procéder à des ré-identifications faite aux utilisateurs de ces décisions.

<http://openlaw.fr/sites/default/files/2017-05/Livret_blan_c_interactif21_04_0.pdf>

Il ne s'agit pas pour autant d'une mémoire du droit, au sens d'un ensemble de normes qui s'imposerait à l'occasion des opérations de jugement auxquels procède le juge lorsqu'il doit trancher un litige⁸. Le droit français est un droit écrit qui édicte des normes générales et abstraites conçues par d'autres que l'autorité judiciaire et qui s'imposent au juge. Certes, la jurisprudence s'incorpore à ces normes, offrant ainsi aux juridictions supérieures la possibilité de prendre leur part à la construction du sens des règles, mais ce pouvoir créateur – parfois contesté – n'est pas de même nature que l'activité quotidienne des juges du fond. En l'état actuel du droit positif, ces décisions ne participent pas à la construction de la règle de droit⁹. Autrement dit, il n'y a pas de raison que ces décisions soient mises à disposition du public sur la base Légifrance : il ne s'agit pas ici de faciliter l'accès au droit, entendu comme l'accessibilité aux textes qui contiennent les règles de droit applicables¹⁰.

La simple possibilité d'accéder à ces décisions et à leur contenu pourrait cependant modifier cette analyse. On a utilisé à leur propos la notion de « jurisprudence concrète »¹¹ ou encore celle de « jurisprudence-appréciation »¹². Ces expressions manifestent la difficulté à nommer cet objet nouveau et à apprécier la place qu'il pourrait prendre dans le raisonnement juridique. Même s'il n'est pas question de lui prêter un pouvoir normatif, sur le modèle du précédent, sa simple existence peut aboutir à lui assurer un pouvoir de fait. Déjà, les professionnels savent qu'ils peuvent se fonder sur une décision antérieure comparable pour entraîner la conviction du juge ou, étant juge, élaborer leur propre conviction. Simplement, cet usage traditionnel d'un argument rhétorique est facilité par l'accès simplifié à une banque de données dans laquelle la recherche d'une ou plusieurs décisions comparables est rapide.

Mais l'usage de l'intelligence artificielle pour produire une analyse automatisée du contenu d'une grande masse de décisions va permettre de proposer des solutions « moyennes », ou « probables », spécialement s'agissant de quantum. Indépendamment de la justesse de ces analyses, se pose alors la question du pouvoir performatif de ces propositions. En effet, de façon comparable à ce qui se passe pour les barèmes indicatifs proposés aux magistrats, mais avec une force accrue liée à la technicité de la méthode et à son apparente neutralité, on doit se demander dans quelle mesure la « proposition » ainsi faite par l'outil ne va pas systématiquement être retenue par les magistrats qui y verront, au final, la meilleure des solutions. Nous serions ainsi confrontés à une construction horizontale du droit, non pas sur le modèle du précédent (trouver une décision comparable et s'y conformer), mais sur un modèle de type probabiliste (trouver les solutions habituellement données à ce type d'affaire et s'y conformer). Il faudra alors non seulement préserver le principe même du

⁸ Sur le droit comme norme de jugement des actions, Voir A. Jeammaud, La règle de droit comme modèle, D. 1990, p. 199.

⁹ Sur la distinction entre contentieux et jurisprudence, voir E. Serverin, Juridiction et jurisprudence : deux aspects des activités de justice, Droit et Société 1993, n°25, pp. 339-349.

¹⁰ C'est d'ailleurs la raison pour laquelle le rapport Cadiet préconise un accès différencié aux bases de décision qui s'annoncent. La question se pose d'ailleurs de l'opportunité de laisser dans la base Légifrance les quelques décisions du fond qui s'y trouvent déjà, sauf à considérer qu'elles constituent dorénavant une règle jurisprudentielle – statue ce que la seule diffusion sur Légifrance ne permet pas d'acquiescer.

¹¹ A. Garapon et alii, La prudence et l'autorité, L'office du juge au XXI^{ème} siècle, rapport de l'IHEJ, mai 2013 <http://www.ihej.org/wpcontent/uploads/2013/07/rapport_office_du_juge_mai_2013.pdf>

¹² P. Deumier, qui l'utilise par opposition à la jurisprudence, au sens classique du terme, qualifiée par l'auteur de jurisprudence-interprétation, Contribution écrite in L. Cadiet, Rapport sur l'Open data des décisions de justice, préc.

pouvoir d'appréciation du juge mais aussi régler la balance entre ce pouvoir, réaffirmé, et les usages de ce type d'outils. En effet, les données issues de ces décisions, bien que n'ayant pas la nature des normes juridiques si l'on en retient une définition positiviste, pourraient exercer un pouvoir de fait qu'il ne faut pas négliger. Ainsi, si les bases de décisions ne constituent pas a priori une mémoire du droit, parce que ni ces décisions ni leur agrégation ne sont « du droit », ces bases de décisions et les analyses qu'elles vont permettre pourraient devenir une nouvelle source de normativité.

2. De l'indétermination de droit à la réduction des incertitudes ?

L'analyse de corpus de décisions de justice montre, s'il en était besoin, que la fabrication des décisions de justice ne se réduit pas à l'élaboration normative, via l'élaboration d'une interprétation qui fera jurisprudence. La connaissance de la production des juridictions permet de connaître ce que les parties et les juridictions font avec le droit, si et comment elles le mobilisent, dans quelles circonstances et comment les magistrats usent de leur pouvoir souverain d'appréciation, ensemble d'éléments que l'analyse de la jurisprudence ne permet pas d'appréhender. C'est la raison pour laquelle, ici, un travail de hiérarchisation des décisions¹³ en fonction de leur utilité n'a pas lieu d'être : la question n'est pas de connaître l'état du droit en retenant les seules décisions qui apportent un élément nouveau à cette construction normative. La connaissance de l'état du droit, qui justifie une telle hiérarchisation, n'est pas la connaissance des modalités d'application de ces normes, y compris interprétées, à des situations d'espèces. Ainsi, lorsqu'il s'agit d'apprécier comment les juges fixent la résidence habituelle des enfants, l'obligation d'entretien et d'éducation de leurs parents séparés, il n'y a pas de décisions qui soient moins importantes que d'autres. Il en est de même s'agissant de l'évaluation de dommages et intérêts par exemple ou encore des sanctions retenues pour tel ou tel type d'infraction ou de manquement. Cela étant, dès lors que l'on s'intéresse à la fabrication des décisions de justice, surgit la question de leurs déterminants : l'analyse des décisions produites permet-elle de déceler les déterminants de la décision, y compris les déterminants « illégitimes », au sens où ils ne résultent pas de l'application attendue des règles de droit ?

Indétermination du droit et dispersion des décisions de justice.

Les recherches ayant pour objet la connaissance de la production des juridictions partent de l'hypothèse de l'indétermination du droit et la conforte. Très rapidement exposée, celle-ci postule qu'un ensemble de règles de droit positif ne permet pas de connaître ce que seront les décisions rendues sous son autorité. Entre la règle générale et abstraite, même bien écrite, et la décision, se glisse tout un monde de discussion sur les faits qui seront retenus comme établis et de choix des textes auxquels ils seront soumis, en contemplation des intérêts défendus par chacune des parties et du résultat recherché par le juge. Le magistrat dispose ainsi d'un espace de liberté, plus ou moins large selon les espèces et les règles mobilisées, et qui fera l'objet d'un contrôle des juridictions supérieures plus ou moins resserré selon le pouvoir d'appréciation dont il dispose.

L'analyse de corpus de décisions de justice permet d'approcher cet espace de liberté et de le décrire, en mettant éventuellement en évidence la dispersion des décisions de justice, c'est

¹³ Hiérarchisation d'ailleurs introduite par les juridictions elles-mêmes dans les classifications qu'elles donnent des décisions qu'elles rendent, renforçant ainsi le caractère normatif de leur activité

à dire le fait que des situations comparables reçoivent des solutions différentes au sein du territoire de la République, que ce soit entre deux juridictions ou au sein d'une même juridiction. Cette disparité, a priori inhérente au fonctionnement du droit, peut être réduite par une approche transversale de l'activité des magistrats qui se doteraient d'outils communs de décision. L'analyse de l'expérience des barèmes utilisés au sein des juridictions montre cependant qu'une telle approche, défendue par certains, heurte pour d'autres une conception plus individualiste du travail du juge, seul maître des décisions qu'il rend. Pour ceux-là, la comparaison des décisions qu'ils rendent avec celles d'autres magistrats constituerait plutôt une forme de manquement à leur fonction.

Cette disparité des décisions est jusqu'à ce jour restée essentiellement invisible, notamment du grand public. Quant au monde du droit, s'il l'a pressenti, il semble s'en être d'autant plus facilement accommodé que sa démonstration restait très confidentielle. Mais cette disparité peut être considérée comme un grave défaut de fonctionnement de la justice dès lors qu'elle est plus largement connue. Or l'accès aux décisions de justice et les analyses qui sont produites contribuent à cette visibilité et la mise en évidence de la disparité des décisions impose à la justice de répondre à une injonction nouvelle d'égalité entre les justiciables sur l'ensemble du territoire national – ou de justification de cette disparité. La recherche de solutions pour la réduire et la proposition d'une « justice prédictive » s'inscrit dans ce mouvement. Reste à savoir dans quelle mesure les analyses qui sont produites et les conclusions qui en sont tirées sont non seulement susceptibles de réduire la disparité mais aussi et surtout de la réduire sur une base solide, en rapprochant des décisions rendues dans des situations effectivement comparables. Cette ambition présume que l'on puisse associer les solutions retenues dans les décisions aux éléments de fait et de droit qui les justifient. Plusieurs obstacles rendent ce projet sans doute trop ambitieux, qu'il s'agisse de la difficulté à identifier les déterminants des décisions de justice ou qu'il s'agisse de la place occupée par la construction juridique du différend, en amont de la saisine du juge. Il n'en resta pas moins que ces analyses peuvent faire avancer la connaissance et fournir des éléments supplémentaires d'analyses aux praticiens.

A la recherche des déterminants (même illégitimes) des décisions.

Pour reprendre le même exemple que précédemment en matière de pension alimentaire, on a vu que l'analyse des décisions permettait de fournir des données telles que les montants de pensions fixées par les juges. Une étape ultérieure de ce type de travaux est la recherche des déterminants de ces montants. Posant l'hypothèse d'un lien entre ceux-ci et les ressources du seul parent débiteur ou des deux parents, on peut confronter ces montants avec les ressources disponibles des parents. On constate alors que ces informations ne sont pas systématiquement détaillées dans la décision, ne serait-ce que pour des raisons de preuve. D'ailleurs, la formulation des décisions peut montrer que le juge est dubitatif sur les revenus déclarés par tel ou tel des parents et cette incertitude, si elle n'empêche bien évidemment pas le juge de fixer un montant de pension, peut influencer le montant qu'il fixera¹⁴. D'autres hypothèses doivent être testées, permettant d'apprécier la part d'autres critères dans la décision finale, tels que, dans notre exemple, l'âge des enfants, leur nombre

¹⁴ On mentionnera également que, dans le cadre d'une procédure écrite, la décision doit « exposer succinctement les prétentions respectives des parties et leurs moyens » mais que « cet exposé peut revêtir la forme d'un visa des conclusions des parties » (article 455 C. Pro. Civ), réduisant d'autant les informations disponibles pour comprendre les décisions rendues.

ou encore le coût de l'exercice du droit de visite et d'hébergement, voire la faute exclusive dans le cadre d'une procédure de divorce.

Les résultats qui sont produits sont ainsi nécessairement soumis aux hypothèses testées ou, autrement dit, aux déterminants qui auront pu être retenus par le chercheur comme pouvant avoir un effet sur les montants fixés. Or ces déterminants ne sont pas forcément attendus, ni légitimes. On peut ainsi déceler un lien entre le sexe du juge et le montant des prestations compensatoires qu'il fixe¹⁵, entre le sexe des enfants et le montant des pensions alimentaires qui leurs sont allouées, ou encore entre la présence d'un avocat et la générosité des montants obtenus¹⁶. L'expérience montre que la recherche traditionnelle (manuelle) des déterminants des décisions via les données récupérées au sein des décisions a ses limites et qu'elle ne permet pas d'expliquer de façon entièrement satisfaisante les résultats obtenus¹⁷. Il sera intéressant de tester si l'analyse automatisée peut permettre d'aller au-delà, en particulier par l'utilisation d'apprentissage non supervisé¹⁸. La finesse des analyses automatisées pourrait permettre de mieux déceler les déterminants illégitimes des décisions du justice, pour une meilleure connaissance du fonctionnement du droit et de la justice.

Du point de vue de la connaissance, on remarquera que les résultats qui sont produits sont nécessairement soumis aux informations disponibles dans le libellé même de la décision : la recherche des déterminants de la décision à partir du texte de cette décision atteint ses limites si l'on considère que certains de ces déterminants peuvent ne pas être explicites, particulièrement lorsqu'ils ne correspondent pas à l'application d'un critère légal de décision, voire qu'ils sont illégitimes ou mêmes involontaires. Tout dépend alors de la présence ou pas de l'information recherchée au sein de la décision. Du point de vue des usages opérationnels de ces connaissances, on notera que l'existence de déterminants illégitimes plaide pour que les analyses automatisées ne soient pas, en l'état, utilisées comme un outil de justice prédictive.

La construction juridique du différend.

On sait que du conflit qui oppose deux entités (personnes, entreprises, administrations ...) peut naître un litige, juridiquement circonscrit et audible par l'appareil de justice¹⁹. Cet accès

¹⁵ Il s'agit ici des résultats tirés d'une autre recherche, portant cette fois sur la prestation compensatoire, Ouvrage à paraître, C. Bourreau-Dubois, I. Sayn (dir), Larcier, 2018.

¹⁶ La présence d'un avocat parmi les déterminants illégitimes des décisions peut surprendre. Elle démontre logiquement que l'assistance d'un avocat est utile. Mais elle démontre aussi, a contrario, la dispersion des solutions retenues à l'issue du processus juridictionnel sur un critère qui n'est ni celui du droit applicable, ni celui des faits de l'espèce, remettant ainsi en cause les dispositions qui, au prétexte de favoriser de l'accès au droit en facilitant l'accès au juge, autorisent à saisir la justice sans l'aide d'un professionnel compétent.

¹⁷ Montant des pensions alimentaires : pratiques et barème, Infostat Justice 116 - Avril 2012, préc.

¹⁸ Pour une définition, voir CNIL, Comment permettre à l'homme de garde la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle, déc. 2017, p. 16 : « L'apprentissage profond (Deep learning) est le socle des avancées récentes de l'apprentissage automatique, dont il est l'une des branches³. On distingue apprentissage automatique supervisé (des données d'entrées qualifiées par des humains sont fournies à l'algorithme qui définit donc des règles à partir d'exemples qui sont autant de cas validés) et non supervisé (les données sont fournies brutes à l'algorithme qui élabore sa propre classification et est libre d'évoluer vers n'importe quel état final lorsqu'un motif ou un élément lui est présenté). L'apprentissage supervisé nécessite que des instructeurs apprennent à la machine les résultats qu'elle doit fournir, qu'ils l'entraînent ».

¹⁹ Sur la distinction entre conflit (« une relation antagonique que réalise ou révèle une opposition de prétentions ou aspirations souvent complexes, plus ou moins clairement formulées, entre deux ou plusieurs groupes ou individus, et qui peut connaître une succession d'épisodes, d'actions, d'affrontements ») et litige (« Une opposition de prétentions juridiques soumise à une juridiction civile, pénale, administrative ou arbitrale, appelée

à la juridicité suppose un travail préalable de décryptage des termes du conflit pour y déceler en quoi celui-ci peut être transcrit pour entrer dans un débat juridique judiciairement admissible. Ce processus est préalable à la saisine du juge²⁰, de sorte que reste nécessairement absent du libellé de la décision ce que l'on peut appeler la construction juridique du différend, c'est à dire les choix qui ont été fait en amont, par les parties, dans la sélection des faits exposés, des qualifications proposées et des conséquences juridiques à en tirer développées devant le juge. En particulier, seront accessibles les seuls développements autour de ce qui aura été demandé, en laissant dans l'ombre ce qui aurait pu l'être mais que les parties ont préféré laisser de côté, pour des raisons variées. Pour ne prendre qu'un exemple, dans le domaine du droit de la famille, on sait qu'une prestation compensatoire n'est pas toujours demandée dans des situations de fait où elle aurait pu être retenue par le juge. On ne peut alors que proposer des hypothèses sur les motifs de ce non recours, depuis un choix stratégique (par exemple négocier parallèlement sur la liquidation du régime matrimonial) jusqu'à une conception de l'égalité qui conduit les femmes à ne pas revendiquer la reconnaissance de leur sous-investissement professionnel bien qu'il soit lié à leur surinvestissement domestique.

La justice dite prédictive peut alors être envisagée comme une aide à la décision pour les parties et leurs avocats. Ils sauront plus précisément quels types de qualifications ont le plus de chances de prospérer devant les juges et pour quelles conséquences juridiques, leur permettant de développer les arguments adéquats. Mais cet outil ne saurait remplacer – et il n'en est d'ailleurs pas question – le travail de construction du différend réalisé d'abord par les parties, ensuite par les juges, qui retiennent ou non tels ou tels faits, telles ou telles qualifications, telle ou telle interprétation du droit qui leur sont proposés par l'une et l'autre partie à l'instance. C'est seulement sur la base des qualifications retenues dans les décisions elles-mêmes que les analyses proposées sont conduites. Ainsi, le fait qu'à des qualifications identiques correspondent tels types de décisions ne suffit pas à démontrer la pertinence d'une justice prédictive qui serait entendue comme les prémices d'un juge-robot.

3. Réduire la marge d'incertitude : des barèmes à la justice prédictive ?

Si l'analyse automatisée de grande masse de décisions de justice ne semble pas pouvoir tenir les promesses d'une justice prédictive, sauf à ignorer tout un pan de l'activité de construction du litige, il n'en reste pas moins qu'elle peut permettre de tenir les promesses d'une justice « probabiliste », en fournissant une décision « moyenne », ou « probable » à partir des demandes finalement formulées par les parties et des qualifications finalement retenues par les juges.

S'agissant de quantum, le rapprochement avec la problématique des barèmes déjà explorée s'impose. Un barème permet de formaliser le raisonnement suivi : il sélectionne les critères jugés pertinents auxquels il associe un résultat. Ce faisant, il hiérarchise les critères retenus

à la trancher par une décision »), voir A. Jeammaud, « Conflit, différend, litige », *Droits* 2001/2 (n° 34), p. 15-20.

²⁰ Même si les professionnels qui reçoivent des demandes formulées par des parties non représentées ni accompagnées, spécialement dans le domaine de la protection sociale, savent que les demande peuvent ne pas être juridiquement formulées et par conséquent pas toujours recevables, mettant ainsi à l'épreuve le rôle plus ou moins actif du juge dans la conduite du procès.

et délaisse d'autres critères, jugés secondaires. Le barème, sous réserve qu'il soit largement utilisé, est ainsi un outil qui tend à garantir une meilleure répétabilité et donc une meilleure prévisibilité des décisions. La construction même du barème relève donc d'une décision « politique » : des choix de construction du barème sont nécessaires au moment de son élaboration, comme une forme renouvelée d'interprétation du droit, et cela y compris s'agissant d'un barème élaboré à droit constant. Le barème propose des interprétations des textes mobilisés et remplit les silences des textes. L'explicitation des choix de construction des outils est donc essentielle pour apprécier l'opportunité d'utiliser l'outil dans tel ou tel cas d'espèce.

Cette fonction politique est *a priori* inexistante dans des barèmes constatés²¹, issus d'une analyse des décisions préexistantes, et cela que cette analyse soit réalisée de façon « artisanale » ou qu'elle utilise des techniques de *data mining* : ils ne font *a priori* que retranscrire des données chiffrées tirées de ces analyses. Cette neutralité doit cependant être vérifiée dans la mesure où ces barèmes constatés retiennent un nombre de critères de décision fini, déterminés *a priori* par les éléments d'information qui ont été fournis aux algorithmes pour les « former » à l'analyse, sauf à utiliser des techniques d'apprentissage profond non supervisé. Mais dans les deux cas, l'utilisation d'un algorithme ne permet pas de faire l'économie d'une réflexion sur la façon dont la mesure a été construite.

On touche ici la différence entre des analyses conduites dans un but opérationnel et des analyses conduites dans un objectif de connaissance. Lorsqu'ils s'intéressent aux déterminants des décisions de justice, les chercheurs espèrent les détecter tous, quel que soit leur poids relatif dans la décision et quel que soit leur statut (légitime ou illégitime, conforme au rôle du juge de dire le droit ou pas), indépendamment de toute approche normative. Les analyses conduites dans un but opérationnel ne peuvent, quant à elles, pas aboutir à intégrer dans leurs préconisations des critères de décision marginaux et encore moins illégitimes. Les outils proposés peuvent donc aboutir à évacuer certains des déterminants de la décision. Ils peuvent aussi reproduire de façon involontaire des critères de décision qui n'ont pas été décelés comme tels. On songe ici à l'expérience américaine où il apparaît finalement que l'outil de prédiction de la récidive qui a été mis en place aboutit à reproduire une discrimination raciale qui ressortait de fait des décisions antérieures sur lesquelles il se fondait²². Pour être pleinement utilisés – donc écartés si nécessaire – les barèmes doivent être accompagnés d'une explication des choix qui ont présidé à leur construction. L'utilisation d'algorithmes ne doit pas faire échapper à cette obligation dès lors qu'ils conduisent à proposer des barèmes aux professionnels. Mais cette condition est plus délicate à respecter pour les barèmes construits que pour les barèmes dits constatés : les analyses produites automatiquement peuvent intégrer des déterminants illégitimes non détectés, d'autant que l'usage d'algorithmes non dirigés rend l'analyse *ex-post* plus difficile, sinon impossible²³. Cette « boîte noire » et les risques qu'elle suppose sont écartées avec l'utilisation de barèmes construits, dont les choix sont explicités et qui, par construction, ne

²¹ Sur la distinction entre « barème constaté » et « barème construit », voir E. Serverin, V° « Barèmes », in L. Cadet, (dir.) Dictionnaire de la Justice, PUF, 2004. Voir également Le droit mis en barèmes ? I. Sayn (dir.), Dalloz, 2014.

²² Ce débat est rappelé dans CNIL, Comment permettre à l'homme de garde la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle, déc. 2017, p. 32.

²³ Sur la distinction entre algorithmes dirigés et algorithmes non dirigés expliquée au non spécialistes, voir également CNIL, Comment permettre à l'homme de garde la main ? préc.

peuvent pas contenir de déterminants inappropriés. Ils peuvent même être considérés comme permettant d'écarter le poids de critères non légitimes intégrés de fait par les magistrats.

Conclusion

Parallèlement au monde du droit, le monde de la recherche se mobilise autour de ces questions, à partir d'une approche nécessairement pluridisciplinaire permettant de confronter les compétences des juristes avec celles des informaticiens bien entendu, mais aussi celle des économistes ou des sociologues et plus largement des sciences humaines et sociales. Parallèlement à l'évaluation et à la progression attendue des outils d'analyse appliqués aux différents contentieux, il est nécessaire d'avancer dans la réflexion sur les usages de ces outils par les professionnels et les conséquences que l'on peut en attendre, notamment sur leur effet performatif ou leurs conséquences réelles ou supposées sur le développement des modes alternatifs de règlement des différends.

On peut aller plus loin encore et envisager les résultats qui pourraient être obtenus à partir de l'enrichissement des décisions de justice avec des informations collectées par ailleurs. L'open data des décisions de justice et les outils d'analyse susceptibles de lui être associés pourraient en effet permettre de dépasser les limites liées à la règle de l'unicité de l'instance, en permettant des recherches longitudinales, par exemple sur les « carrières judiciaires » des majeurs protégés, des enfants en danger ou des chefs d'entreprises, à partir de l'agrégation des décisions juridictionnelles les concernant, ou sur « l'histoire juridictionnelle » de telle ou telle entreprise publique ou privée. On peut également envisager de travailler à partir de décisions de justice enrichies, complétées d'informations collectées par ailleurs, notamment dans les bases existantes auprès de l'administration des impôts ou des organismes de protection sociale. Il s'agit, du point de vue de la recherche, d'un objectif de connaissance, indifférent à l'identification des personnes concernées, mais on mesure les précautions nécessaires pour respecter la vie privée. Il n'en reste pas moins que ces perspectives sont enthousiasmantes.