



DROIT ET CROISSANCE
FAIRE DU DROIT UN VECTEUR DE CROISSANCE

Réutilisation des données d'identité des magistrats : Quel avenir pour la justice prévisionnelle ?

LE JUGE FRANÇAIS FACE AUX LEGALTECH

par Eva Fourel
et Youcef Rahmani
membres de Droit & Croissance

La loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (la « Loi de programmation ») comprend trois nouveautés pour les *legaltech* développant des outils de justice prévisionnelle, c'est-à-dire les outils statistiques permettant de mieux cerner les enjeux et issues potentielles d'une loi ou d'une régulation lors d'un potentiel contentieux juridique. Premièrement, par le biais du contrôle a priori du Conseil constitutionnel, la Loi de programmation valide l'activité des plateformes de service en ligne d'aide à la résolution à l'amiable de litiges, sans que celles-ci aient besoin d'obtenir une certification par un organisme accrédité. Deuxièmement, la Loi de programmation vient préciser le régime d'*open data*, selon lequel toutes les décisions de justice doivent désormais être accessibles aux justiciables sous format électronique. Ce principe avait été consacré une première fois par la Loi pour une République numérique de 2016. Dans ce cadre, l'occultation des données nominatives des parties et des tiers au litige sera obligatoire¹. Enfin, la Loi de programmation, en son article 33, interdit la réutilisation des données d'identité des magistrats et membres du greffe présents dans les décisions de justice lorsque cette réutilisation a « *pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées* ».

¹ L'anonymisation des magistrats et auxiliaires de justice est également prévue lorsque la divulgation de leur identité est de nature à porter atteinte à la sécurité ou au respect de la vie privée de ces personnes ou de leur entourage, par exemple dans des affaires de terrorisme.

Après (i) une présentation de l'article 33 de la Loi de programmation, il s'agit d'envisager de manière sommaire, (ii) dans quelle mesure le développement de certaines *legaltech* est susceptible de contribuer à une meilleure efficacité économique du droit, et (iii) quelles sont les conditions relatives à l'anonymat souhaitables du point de vue de cet objectif d'efficacité, en tenant compte des éléments contextuels propres à la justice française.

Nous concluons par des propositions pour la mise en œuvre d'une politique de protection des données relatives aux magistrats qui soit compatible avec la préservation de l'égalité entre les justiciables ainsi que le développement d'outils juridiques innovants et des sociétés dites « de transparence », qui apportent avec elles de nouvelles exigences des citoyens en matière de contrôle de l'activité des juridictions.

1. Le champ d'application de l'interdiction de réutiliser les données d'identité des magistrats à des fins statistiques ou prévisionnelles

1.1 Sanctions et portée de l'interdiction : les praticiens doivent-ils s'inquiéter ?

L'article 33 de la Loi de programmation prévoit l'interdiction de la réutilisation des données d'identité des magistrats et des membres du greffe ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées. L'interdiction est sanctionnée à la fois sur le plan administratif, par le biais de la CNIL, et sur le plan pénal, par les peines prévues à l'article 226-18 du Code pénal de cinq ans d'emprisonnement et 300 000 euros d'amende. Les peines complémentaires des articles 226-24 et 226-31 du Code pénal s'appliquent également, et comportent principalement, pour ce qui concerne notre sujet, « l'interdiction d'exercer l'activité professionnelle ou sociale dans l'exercice ou à l'occasion de l'exercice de laquelle l'infraction a été commise ».

À titre préliminaire, relevons que les conditions d'application de l'article 226-18 du Code pénal ne peuvent, en principe, servir de base pour l'interprétation de l'article 33 de la Loi de programmation. En effet, cette dernière dispose que les articles du Code pénal précités ne gouvernent que les sanctions de la nouvelle infraction, et non son champ d'application². Ce régime ne pourra être précisé que par le biais de futurs contentieux nourrissant la jurisprudence sur cette question.

Si l'article 33 de la Loi de programmation semble avoir pour but de restreindre l'utilisation des données par certaines *legaltech* engagées dans la prévision algorithmique de litiges, elle impacte aussi, plus généralement, toute pratique de profilage des magistrats. Pourtant, ce type de pratique existe depuis de nombreuses années, en France comme à l'étranger, et est considérée par certains comme faisant partie d'une stratégie de contentieuse légitime, ce qui nous amène à nous interroger quant aux questions spécifiques que ce sujet soulève en France. On peut se demander, par exemple, si cette interdiction s'appliquera également aux praticiens du droit qui prennent en compte l'identité des magistrats pour déterminer leur stratégie de procès, sans aucun tempérament à la règle. Un avocat qui a l'habitude de travailler auprès d'un certain tribunal connaît précisément les tendances et pratiques jurisprudentielles de ses magistrats et en informe son client. Or, la frontière est en pratique très difficile à tracer, notamment en raison du manque d'éléments concrets permettant d'évaluer l'étendue de la pratique (absence de code, de documentation, voire de trace écrite) qui permettraient d'établir une présomption relative aux motifs des choix d'audiences et de magistrats. D'où une première question : comment délimiter précisément l'infraction ? Quelle doit être la frontière entre comportement légitime, du moins difficilement répréhensible, et comportement sanctionnable ?

On doit donc se demander pourquoi le législateur, qui a d'une part consacré une plus grande transparence au droit français en permettant la mise à disposition du public des décisions des juridictions judiciaires et administratives, a d'autre part imposé une telle interdiction, obstacle évident à la transparence, et ce d'autant plus que les noms des magistrats ne seront en principe pas occultés.

² La jurisprudence de l'article 226-18 du Code pénal, qui réprime la collecte déloyale, frauduleuse ou illicite de données, précise que le simple fait de détenir, dans un fichier, des données obtenues de manière déloyale, frauduleuse ou illicite suffit à constituer le délit, sans qu'aucune intention délictueuse soit nécessaire. En outre il importe peu que ce fichier soit automatisé ou non. V. Cass. crim., 3 nov. 1987, n° 87-83.429

Ainsi, si une décision de justice est désormais mise à disposition en *open data*, et que les noms des magistrats doivent y figurer en principe, pourquoi existe-t-il une restriction quant à la réutilisation de ces données publiques ?

1.2 Les arguments avancés en faveur de cette interdiction : une illustration de la « justice institutionnelle » à la française ?

1.2.1 Une contradiction apparente

Il existe une contradiction apparente entre, d'une part, le fait de publier les décisions de justice pour les rendre accessibles à tous les justiciables et, d'autre part, interdire la libre utilisation, par les justiciables, d'un élément présent dans ces publications.

Rappelons d'abord que la justice est un service public rendu au nom du peuple français, conformément aux articles L. 111-1 du Code de l'organisation judiciaire et L. 2 du Code de justice administrative. En outre, l'article 15 de la Déclaration des Droits de l'Homme et du Citoyen dispose que « *La Société a le droit de demander compte à tout Agent public de son administration* ». Ces dispositions devraient favoriser la transparence des décisions de justice et le contrôle démocratique, par les citoyens, de la mise en œuvre des lois, via la publication des décisions de justice. Notons que, si l'article 15 est bien applicable aux magistrats, dans le cadre de l'objectif à valeur constitutionnelle de bonne administration de la justice, il n'est pas à ce titre, invocable directement par le justiciable via une question prioritaire de constitutionnalité, car cet objectif n'est pas considéré comme faisant partie des droits et libertés que la Constitution garantit (au sens de l'article 61-1 de la Constitution).

Si les décisions sont publiques et doivent obligatoirement mentionner les noms des magistrats les ayant rendues afin d'assurer le respect d'une exigence constitutionnelle, pourquoi les justiciables devraient-ils être limités dans la réutilisation de ces données ?

1.2.2 Les craintes autour de l'individualisation du juge

Les députés ayant introduit cette disposition ont expliqué que la transparence inhérente à l'*open data* ne saurait conduire à une réutilisation des données d'identité des magistrats et auxiliaires de justice, en particulier pour le profilage ou le « *ranking* ». Il s'agit ainsi d'éviter, par exemple, un classement de juges et greffiers permettant de dire qu'un juge est plus répressif qu'un autre, ou qu'il donne plus souvent raison à un demandeur de sexe féminin qu'à un demandeur de sexe masculin³.

Ces arguments sont sous-tendus par un débat sur fond de culture juridique. En France, la justice est rendue de manière institutionnelle et non personnelle, et le principe du secret du délibéré impose qu'une juridiction statue sans qu'il n'y ait d'opinion dissidente exprimée. Ainsi, comme le fait remarquer le Vice-président du Conseil d'Etat, « *dans la conception française, la décision de justice transcende la personne des juges : celle-ci s'efface devant l'œuvre de justice* »⁴. En ouvrant la voie à

³ Exposé sommaire de l'amendement n°1425, 15 novembre 2018, [lien](#)

⁴ Ministère 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, novembre 2017, p.48

des études jurisprudentielles en fonction des magistrats auteurs de la décision, la Loi de programmation permettrait aux *legaltechs* d'identifier des pratiques juridictionnelles, voire des orientations jurisprudentielles personnalisées, ce qui mettrait effectivement fin au mythe fondateur de la justice républicaine du juge dénué de toute subjectivité. L'heure est pourtant plutôt à la reconnaissance de cette subjectivité et au contrôle de son exercice via une irruption sur la place publique des questions de politique juridique et d'accès aux données des juridictions.

Le gouvernement fait valoir plusieurs arguments contre la justice prévisionnelle sur la base des données d'identité des magistrats⁵.

D'abord, la justice prévisionnelle serait contestable sur le plan éthique, car elle permettrait de cibler les magistrats et individualiser les décisions des juridictions. Le problème est ici encore un refus du "réalisme juridique" difficilement discutable. Les algorithmes ne préjugent pas d'un cas (en lieu et place du juge), mais permettent de produire des estimations de coûts totaux d'un litige pour les parties, en considération des issues le plus probables en fonction des pratiques d'un juge en particulier. Le rejet du gouvernement se porte donc sur ce point. Du point de vue de l'analyse économique du droit, pourtant, l'estimation des coûts est plus fiable si elle est à même d'incorporer ces éléments individuels. En effet, la justice prévisionnelle permet aux parties d'estimer les coûts totaux d'une instance, et pourrait donc désengorger les tribunaux en encourageant les parties vers le règlement amiable de leur litige. Une prévision ne saurait en aucun cas influencer la décision d'un juge, à tout le moins lorsqu'il s'agit de modèles privés licenciés par des entreprises aux avocats et justiciables. En outre, plusieurs outils de justice prévisionnelle existent déjà sur le marché. Cela ne fait donc aucun sens de parler d'une menace pour la position du juge dans la mesure où ces algorithmes produiront des solutions propres à chaque litige et seraient uniquement accessibles aux parties qui les utilisent. Par ailleurs, il pourrait être souhaitable que de tels outils soient également développés à destination des magistrats, afin de simplifier le traitement des types de dossiers les plus récurrents.

1.2.3 La question du « *juge shopping* »

Ensuite, le gouvernement fait valoir que les outils de prévision déstabiliseraient les institutions juridictionnelles, notamment parce qu'ils encourageraient des comportements de *forum shopping*. Mais est-il vraiment possible d'avoir autant de flexibilité dans le choix de sa juridiction ? Dans la plupart des litiges, les règles de compétence sont claires, et l'on peut penser que, pour les litiges d'une plus grande ampleur, les parties en cause ont généralement suffisamment de moyens pour déjà disposer de bases de données et conseils stratégiques. La question qui est alors posée n'est pas celle du *forum shopping* stricto sensu, mais du « *juge shopping* ».

Dans un certain nombre de cas, l'on dépose un dossier ou bien l'on assigne devant un juge donné, par exemple devant le tribunal d'instance. De tels outils d'analyse auraient, il est vrai, une influence sur le choix des juges par les parties au sein des juridictions. Bien que les juges ne soient culturellement pas individualisés en France, de telles pratiques existent déjà informellement, et sont difficiles à quantifier et à contrôler. C'est leur systématisation et leur commercialisation qui posent un nouveau

⁵ Voir : Assemblée nationale, XV^e législature Session ordinaire de 2018-2019, Compte rendu intégral, Deuxième séance du jeudi 22 novembre 2018, [lien](#) ; Observations du Gouvernement sur la loi de programmation 2018-2022 et de réforme pour la justice, 14 mars 2019, p.10 ; Ministère 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, novembre 2017, p.48

problème. Pourtant, ne serait-ce pas donner à tous les justiciables les armes dont disposent les initiés ? De plus, si la justice française tient à maintenir sa conception institutionnelle de cette mission, interdire le « juge shopping » n'est-il pas justement un aveu de ses limites ?

Enfin, le gouvernement estime que l'exploitation des noms de magistrats permettrait aux acteurs de la justice prévisionnelle de se livrer à des corrélations pernicieuses entre le sens des décisions et les caractéristiques personnelles des justiciables concernés, telles que leur sexe, origine ou appartenance confessionnelle. En croisant ces données avec des informations publiquement disponibles sur les magistrats, davantage de corrélations seraient possibles, ouvrant la voie à davantage de récusations, ou à des critiques personnelles adressées par le public aux magistrats.

2. Les bénéfices économiques liés à la prévisibilité : l'argument économique des coûts transactionnels

Les coûts relatifs à la négociation, ou "coûts transactionnels", sont classiquement présentés par la littérature économique comme un obstacle majeur à l'obtention de résultats globalement optimaux par les acteurs. Depuis les travaux de Coase⁶, il est généralement admis que s'il n'existait pas de coûts transactionnels, les parties à un litige tendraient à négocier entre elles la solution la plus optimale globalement. La conséquence de cela est qu'une justice économiquement efficace est une justice qui cherche à réduire ces coûts transactionnels afin que les parties soient, le plus souvent possible, en situation de pouvoir négocier une telle issue. L'on peut dire que l'amélioration de la prévisibilité du droit agit en ce sens de deux manières.

2.1 Clarifier l'allocation des droits : entre prévisibilité et liberté du juge

D'abord, les coûts transactionnels sont d'autant plus importants que l'allocation des droits est incertaine : si les parties ne sont pas sûres de savoir qui a le droit de faire quoi, il leur sera plus difficile d'identifier les interlocuteurs avec lesquels négocier, de cerner l'objet de la négociation et ses limites. De plus, le nombre de litiges tendra généralement à augmenter, en raison d'appréciations non concordantes des parties quant à la réalité de leurs droits.

De telles incertitudes résultent habituellement notamment des divergences d'interprétation existantes. A cet égard, les outils prévisionnels seraient très utiles pour identifier des incohérences d'application du droit par les juridictions de fond et ainsi améliorer la lisibilité et l'intelligibilité du droit français. Chantal Arens, récemment nommée première présidente de la Cour de cassation, a déclaré au Monde que l'harmonisation de l'interprétation du droit sur l'ensemble du territoire de la République figurait parmi ses priorités. De manière intéressante, elle explique qu'elle souhaiterait que les nouveaux textes de loi soient « *immédiatement interprétés et appliqués de manière cohérente par l'ensemble des juridictions pour désamorcer de futurs contentieux* »⁷. Ce désir d'œuvrer pour rendre l'application de la loi plus transparente et cohérente en France, et ainsi désamorcer les contentieux en amont, pourrait peut-être se trouver en partie facilité par le développement d'outils algorithmiques de prévision établissant des statistiques concernant l'application de la loi par les juges.

Si les conséquences économiques de la prévisibilité des contentieux sont à présent connues, les outils d'harmonisation apparaissent comme contraire aux pratiques effectives des juges, voire, selon certains, comme contraires à leur mission, parce qu'ils sont perçus comme susceptibles de contraindre une harmonisation de leurs interprétations en conflit avec la réflexion individuelle qu'ils doivent conduire à l'occasion de chaque espèce. Cependant, il semble que de tels outils de prédiction soient de nature à concilier ces deux visions, en permettant de rétablir la prévisibilité des décisions sans passer par une harmonisation forcée.

⁶ Coase, R. (1960). The Problem of Social Cost. The Journal of Law & Economics, 3, 1-44.

<http://www.jstor.org/stable/724810>

⁷ Le Monde, « La Cour de cassation veut plus de prévisibilité dans les décisions de justice », 7 septembre 2019, [LIEN](#)

Du reste, les opinions divergent quant au bien-fondé de la tradition de rébellion des juges du fond. Certains estiment qu'il s'agit d'un phénomène consubstantiel au système français, apportant une diversité qui conduit, par un processus analogue au mécanisme de sélection naturelle, à l'adoption généralisée des meilleures pratiques. D'autres y voient un facteur d'instabilité juridique dont les conséquences économiques néfastes sont plus importantes que les bénéfiques liés à la diversité des pratiques au sein des juridictions.

2.2 Estimer les coûts relatifs au procès pour savoir quand transiger

Deuxièmement, la prévisibilité est liée à la possibilité d'anticiper de manière plus précise les décisions des juges saisis. Or, cela permet une estimation plus juste des coûts relatifs au procès, et donne l'opportunité aux parties de transiger pour obtenir une solution qui leur soit plus favorable⁸. La littérature économique a notamment mis en évidence que les différences d'estimation de l'issue d'un procès étaient de nature à empêcher les négociations d'aboutir⁹. Cela se comprend dans la mesure où les parties n'ont pas la même estimation de ce qu'elles peuvent espérer de l'issue d'un procès, il est donc plus difficile pour elles de s'accorder sur une solution au litige qui soit mutuellement plus avantageuse. Le risque est que ce qu'est prête à payer l'une des parties en considération du montant auquel elle pense être condamnée si l'affaire était portée devant un juge soit moins important que ce que l'autre partie espère gagner dans cette situation : dans ce cas aucune des parties n'acceptera de transiger, quand bien même cela aurait été finalement moins coûteux pour toutes les deux.

2.3 Des bénéfiques escomptés pour le fonctionnement de la justice

L'on comprend de ce qui précède que l'utilisation d'algorithmes de justice prévisionnelle est alors susceptible d'améliorer le fonctionnement de la justice.

D'une part, cela pourrait permettre d'encourager les modes alternatifs de règlement des différends, et ainsi désengorger les tribunaux. Une telle régulation de la demande de procès rendrait possible le financement d'un système judiciaire de meilleure qualité, dont les ressources seraient dirigées vers les cas pour lesquels il est plus difficile de réduire les coûts transactionnels (notamment ceux pour lesquels il n'existe pas de précédents), et donc pour lesquels la valeur ajoutée d'un juge est la plus importante.

⁸ V. Deffains, B. (1997). L'analyse économique de la résolution des conflits juridiques. *Revue française d'économie*, 12(3), 57–99. <https://doi.org/10.3406/rfec.1997.1027>

⁹ V. par exemple Cooter, R., & Rubinfeld, D. (1989). Economic Analysis of Legal Disputes and Their Resolution. *Journal of Economic Literature*, 27(3), 1067-1097. <http://www.jstor.org/stable/2726775>
“For a risk-neutral plaintiff, the subjective value of the possible damage award at trial, denoted T_p , equals the money value of expected damages, D_p , times the subjective probability of their award, denoted p_{vp} , that is, $T_p = p_{vp}D_p$. Similarly, for a risk-neutral defendant, $T_d = p_{vd}D_d$. When plaintiff and defendant have the same expectations about trial ($p_{vp} = p_{vd}$ and $D_p = D_d$), they concur about its expected Value, so that $T_p = T_d$. If the parties are relatively pessimistic about the prospects at trial ($p_{vp} < p_{vd}$ and $D_p < D_d$), plaintiff will expect to win less than defendant expects to lose, so that $T_p < T_d$. If they are relatively optimistic about their own prospects at trial ($p_{vp} > p_{vd}$ and $D_p > D_d$), plaintiff expects to win more than defendant expects to lose, so that $T_p > T_d$. If the surplus is negative, the disputants prefer a trial to any possible settlement, so trial is inevitable. If the surplus is positive, however, there is scope for settlement out of court.”

D'autre part, l'on peut également anticiper des améliorations pour l'organisation du service public de la justice. En effet, une prise en compte des données relatives aux décisions rendues, en ce compris celles relatives aux magistrats, pourrait doter les pouvoirs publics d'un outil sans précédent pour accompagner et évaluer la mise en œuvre des différents dispositifs déployés dans les tribunaux. Ces données peuvent offrir une meilleure vue d'ensemble de l'application du droit par les magistrats, faire prendre conscience aux magistrats eux-mêmes de leurs biais éventuels (donc constituer un outil d'analyse de leur propre pratique).

Au contraire de l'argument qui consiste à penser qu'une observation des pratiques individuelles des magistrats peut mettre en évidence les différences de traitement éventuellement injustifiées et ainsi les exacerber, voire cristalliser d'inutiles polémiques, l'on peut également penser que les outils de la justice prévisionnelle pourraient justement être utilisés pour corriger d'éventuelles différences de traitement discriminatoires et défendre l'égalité réelle des justiciables devant la loi. Les algorithmes pourraient orienter la politique juridique de manière plus efficace, en tenant compte des éventuelles « corrélations pernicieuses » constatées.

Cette collecte de données sur les magistrats permettrait en outre de nourrir la recherche sur l'aide à la prise de décision, ce qui constitue une source d'innovation très importante dans de multiples domaines, et qui pourrait par exemple se traduire par une optimisation du système judiciaire via l'introduction de technologies destinées à alléger la charge de travail des juges.

2.4 Un enjeu concurrentiel relatif au développement des industries liées à la donnée

Sur un plan plus macroéconomique, restreindre l'usage des données sur les magistrats est de nature à créer un désavantage concurrentiel de la France dans les secteurs d'avenir liés, en particulier le big data et l'intelligence artificielle. En particulier, si d'autres juridictions développent des bases de données structurées rapidement, limiter ce phénomène fait courir à la France un réel risque de diminution de l'attractivité de son droit, via une diminution de sa prévisibilité. Cela serait d'autant plus marqué dans l'hypothèse où les outils prévisionnels deviendraient la norme dans d'autres juridictions.

Il semble donc qu'il faille encourager le développement de telles technologies, dans la continuation des objectifs de développement de l'intelligence artificielle en France, et de manière plus strictement économique, en raison des différents bénéfices relatifs à la prévisibilité des décisions. Pourtant, une série de problèmes relatifs au contexte institutionnel français se posent, et il semble impératif d'adopter des modalités d'anonymisation qui soient les plus appropriées eu égard à l'objectif de concilier logique économique et protection de la justice française.

3. L'anonymat des magistrats dans le contexte français : arguments, contraintes, mise en œuvre

3.1 La démocratisation des algorithmes de justice prévisionnelle : un nouveau mode d'accès au droit ?

Le développement de la justice prévisionnelle crée de nouveaux moyens de connaissance du droit par les justiciables, et donc d'accès au juge. Il semble légitime de penser que le justiciable, en tant que sujet et consommateur vulnérable du droit, devrait avoir accès à la jurisprudence de la manière la plus simple et intuitive possible. Les algorithmes permettraient, de manière générale, de rendre la jurisprudence plus accessible puisqu'ils agrègent toute la jurisprudence sur un thème pour produire une opinion finale qui fait sens au justiciable non-initié au droit. Les algorithmes exploitant les données des magistrats, dans la mesure où ils permettent de produire des résultats plus précis, contribuent donc à cela.

Il convient, à cet égard, de signaler une demande du public d'accéder à ce type d'information, qui pourrait être comprise comme reflétant une évolution de la perception sociale du contrôle des juges. A l'heure de la transparence et de l'*open data*, l'image d'impartialité de la justice passe sans doute par l'acceptation de la démocratisation de l'utilisation de ces données et de l'élargissement de l'accès des citoyens aux coulisses des missions de service public assurées en leur nom.

3.2 Protéger la fonction de magistrat dans ce nouveau contexte

Il faudrait également accompagner une autorisation d'utiliser les données de nouvelles garanties d'indépendance pour les juges vis-à-vis de leur hiérarchie (procureur, président, dans la répartition des dossiers par exemple). Si l'indépendance statutaire des juges est en principe établie, il convient de réfléchir à la manière dont un contrôle démocratique pourrait être détourné en un contrôle politique. Nous invitons donc à repenser la place des magistrats plus globalement, dans un contexte de bouleversements technologiques.

Par exemple, des logiciels permettant de traiter de manière plus objective les questions de répartition de dossiers et de mutation, en proposant des suggestions sur la base des souhaits exprimés par les membres de la magistrature, pourrait être une piste à explorer. D'autres dispositifs envisageables pourraient comprendre un protocole de mobilité permettant de traiter les situations à risque (identifiées par exemple algorithmiquement), ou encore la mise en place d'une cellule indépendante ayant pour mission de vérifier que l'utilisation des données effectuée par l'administration ou les justiciables n'entrave pas *de facto* l'indépendance des juges.

3.3 Les modalités d'utilisation de ces données par les algorithmes : peut-on réellement protéger certaines caractéristiques ?

Compte tenu de la formulation de l'article 33 de la Loi de programmation, on peut se demander si le nom des magistrats peut être intégré dans un modèle algorithmique, sans avoir été préalablement stocké dans un fichier. Pour le moment, les algorithmes prévisionnels sont souvent basés sur des techniques de *machine learning*, dans lesquels des milliers de données, étiquetées en fonction de leur nature, sont analysées par un algorithme qui, via des techniques de régressions mathématiques, « apprend » à reconnaître certains résultats en se fondant sur un certain nombre de caractéristiques. Pour ce faire, l'algorithme fonctionne en identifiant des corrélations et autres motifs récurrents dans les données, pour ensuite « prédire » un résultat en fonction de futures caractéristiques qui lui sont soumises. Ainsi, plus les données sur lesquelles l'algorithme s'est entraîné sont affinées, plus l'algorithme sera précis dans ses prédictions.

Une autre manière d'aborder cette interdiction serait de considérer que l'identité des magistrats revêt désormais le statut de « caractéristique protégée », et ne peut donc jamais être utilisée comme une donnée d'entrée dans un algorithme prévisionnel, que ces noms figurent dans un fichier ou qu'ils soient utilisés directement depuis le document de la décision. Cette approche serait la plus prudente. Il devient alors nécessaire de sélectionner avec soin les données d'entrée (ou « *input* ») d'un algorithme pour analyser une décision de justice. Si la « prédiction » est suffisamment précise en n'utilisant que les faits, les sources textuelles de droit et le lieu de la juridiction en question, alors les *legaltechs* n'ont rien à craindre de cette loi, puisque ces données n'entrent pas sous le coup de l'interdiction de l'article 33.

Enfin, les algorithmes, de manière générale, s'entraînent déjà bien souvent sur des bases de données très fortement corrélées à certaines caractéristiques pourtant protégées par la loi. Par exemple, s'il est interdit de prendre en compte l'ethnie d'une personne pour prendre une décision professionnelle, le fait d'indiquer être, par exemple, propriétaire ou locataire de sa résidence, permet d'identifier de manière indirecte certaines caractéristiques pourtant protégées, puisque statistiquement la caractéristique de propriétaire est corrélée à certaines autres. Les anglophones désignent ces caractéristiques, qui agissent comme des caractéristiques protégées « par procuration », par l'expression « *proxy features* ». En combinant plusieurs caractéristiques fortement corrélées à une population ou personne, les *proxy features* dégagent des motifs suffisamment distincts pour être rattachés à cette population ou personne.

3.4 Identifier des formations juridictionnelles plutôt que des magistrats individuels

La collégialité inhérente à la plupart des juridictions signifie que l'identification de chaque magistrat individuel n'est pas forcément pertinente. Les modèles algorithmiques précédemment décrits basent leurs prédictions sur les éléments les plus déterminants. Il est donc tout à fait possible de se contenter d'identifier la combinaison des trois magistrats ayant rendu des décisions répondant à certains critères. Ainsi, par exemple, en combinant le lieu de la juridiction, sa compétence et les thèmes abordés, il sera assez simple pour l'algorithme de pouvoir réidentifier les magistrats, sans toutefois formellement les nommer. Si la loi protège l'identité des magistrats uniquement dans la mesure où leurs données d'identité sont réutilisées, alors un algorithme parvenant à distinguer une « composition juridictionnelle A » d'une « composition juridictionnelle B », et leur attribuant ensuite

des statistiques, pourrait être légal et accomplir à peu de choses près la même fonction. Il n'est pas clair pour le moment que la loi sanctionne la réidentification indirecte des magistrats, mais la possibilité d'établir des statistiques en fonction de la formation juridictionnelle ne semble pas interdite. Lors de leur audition, des producteurs de logiciels de justice prévisionnelle ont d'ailleurs indiqué que la connaissance du nom des juges ne leur était d'aucune utilité technique¹⁰.

3.5 L'interdiction d'utiliser les données d'identification des magistrats : un obstacle pour la recherche

Dans le domaine de la recherche, cela semble priver les travaux français de ces outils nouveaux, pour lesquels on ne mesure pas encore l'étendue de ce qu'ils pourraient nous apprendre, et à l'utilisation desquels il n'existe pas d'obstacle juridique dans d'autres pays. L'on peut notamment raisonnablement penser que l'analyse de ces données par les scientifiques, les chercheurs de tous bords et l'administration elle-même sont de nature à permettre des innovations majeures sur des sujets divers. Parmi ceux-ci, on peut citer les points d'intérêt suivants : les biais cognitifs, les schémas récurrents derrière les divergences d'interprétation, des approches novatrices en matière de sociologie juridique, ou encore d'analyse des coûts économiques de certaines caractéristiques du système judiciaire.

Il semble alors pertinent de distinguer d'une part l'utilisation des données d'identification des magistrats au bénéfice de la recherche académique et des pouvoirs publics (impliquant une anonymisation des résultats lors de la publication de travaux), qui peut bénéficier d'une autorisation d'utilisation des données en question, et d'autre part, l'utilisation de ces données par les justiciables, pour lesquels une interdiction de réutiliser l'identité des magistrats n'est pas économiquement dommageable si la production d'analyses statistiques sur la base de la composition juridictionnelle est autorisée.

3.6 Repenser l'économie de l'*open data*

Et si l'autorisation d'utiliser les données d'identification des magistrats (voire d'autres types de données fournies par les juridictions) aux fins prévues par l'article 33 dans un cadre commercial appelait une rétrocession aux juridictions ? Cela ne concernerait pas la recherche, sur laquelle pourrait peser une obligation d'anonymisation des données dans les travaux finaux présentés : on pourrait connaître la prévalence de certains biais chez les juges français de certaines juridictions, sans qu'il soit besoin de les citer nommément. Ainsi, les entreprises qui font du bénéfice sur cette base pourraient contribuer au financement de la justice – qui est un problème récurrent.

Deux points nous semblent fondamentaux : (i) le bénéfice de ces nouveaux outils ne doit pas s'arrêter aux portes des palais de justice. Les magistrats doivent être les premiers à pouvoir développer leur pratique grâce à ces logiciels, dans un souci d'amélioration de la qualité de service produisant des externalités économiques positives. (ii) Le rôle et la protection des magistrats doivent

¹⁰ Ministère 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, novembre 2017, p.86-87

être repensés, afin de leur donner de réelles garanties d'indépendance sans empêcher l'exercice d'un contrôle citoyen réclamé par nos « sociétés de transparence », au service de la crédibilité et de l'apparence d'impartialité de la justice.

Recommandations

Ainsi, nous invitons le législateur à préciser l'étendue des restrictions qu'il a entendu appliquer à l'utilisation des données d'identification des magistrats, afin de :

- Ne pas pénaliser la France sur le plan de l'innovation en matière de *legaltech*, qui facilitent le travail des professions du secteur et doivent aussi pouvoir être mises au service des magistrats ;
- Clarifier ce qui est permis ou non, afin notamment de permettre aux producteurs de logiciels d'orienter leurs pratiques vers des alternatives statistiquement viables qui ne soient pas préjudiciables au bon fonctionnement de la justice ;
- Réduire les coûts liés au financement de la justice, en réduisant les coûts transactionnels : permettre aux parties de mieux prévoir l'issue d'un potentiel jugement, c'est leur permettre de comprendre leur intérêt à transiger ;
- Rassurer les magistrats quant à leur protection contre toute tentative d'influence et d'intimidation, de la part de leur hiérarchie comme du grand public ;
- Tout en autorisant le développement de travaux de recherche d'un genre nouveau, susceptible de permettre d'optimiser les cas et modalités de recours à la justice, pour le ministère de la justice comme pour les justiciables, et pour offrir des perspectives inédites en matière de compréhension des processus de décision.

A l'heure de l'open data, il semble nécessaire de restreindre le moins possible l'utilisation de ces données – le traitement de données étant la ressource-clé du 21ème siècle – tout en tenant compte de certaines contraintes liées à l'organisation de la justice en France. A ce titre, nous proposons en particulier :

- *A minima* une exemption spéciale à l'interdiction prévue par l'article 33 de la loi de programmation à destination des chercheurs, assortie éventuellement d'une obligation d'anonymisation de leurs travaux finaux (qui ont une portée générale et ne sauraient porter sur tel ou tel juge) ;
- Le paiement de droits d'utilisation à la justice par les plateformes qui font un usage commercial de ces données (les utilisateurs professionnels de ces données se trouvant alors hors de l'open data, ce qui est de nature à exclure la qualification de service public, mais qui semble une hypothèse de fonctionnement plus réaliste pour l'avenir, étant donné la valeur économique toujours croissante liée aux données et à leur exploitation) ;
- Encourager le développement de technologies d'aide à la décision pour les magistrats et autres praticiens du droit, qui leur offrent de nouvelles manières d'envisager leur pratique, des outils de réflexion et de simplification des tâches répétitives ;
- Tenir compte de l'importance de l'image de justice impersonnelle sur laquelle est construite la culture judiciaire française, notamment en se concentrant sur les formations juridictionnelles, quitte à pouvoir les identifier mais en maintenant l'anonymat des magistrats ;
- Tenir compte de l'asymétrie de moyens entre justiciables y compris dans l'accès aux données permettant d'anticiper les décisions de justice et prévenir les risques juridiques, et réaliser qu'une autorisation d'utilisation dans un cadre algorithmique et statistique peut être – à

certaines conditions – de nature à rétablir l'égalité face à l'information entre les justiciables (ou plutôt, face à son traitement, dans un monde dont la complexité rend inabordables les données brutes), enjeu dont l'importance devrait se révéler dans les décennies à venir ;

- Tenir également compte de l'impact d'une société plus transparente sur les praticiens du droit et les magistrats en particulier, en prévoyant des mécanismes de protection spécifiques vis-à-vis de leur hiérarchie (procureur, président, dans la répartition des dossiers en particulier) et en intégrant à la réflexion les pratiques contemporaines en matière de réseaux sociaux et de campagnes de presse comme faisant partie des éléments de stratégie contentieuse.

En un mot, interdire tout court cette utilisation, sans même clarifier son étendue, semble préjudiciable à une bonne application de la rationalité économique des agents à la réduction des coûts liés au procès, et surtout, au développement de la recherche et d'outils innovants qui accompagneront demain la pratique du droit. Cependant, la mise à disposition du public de ces données sensibles ne peut se faire sans un encadrement approprié, issu d'une réflexion sur la place du juge et de la politique juridique dans nos sociétés de transparence, et sur la manière dont ces nouvelles technologies juridiques peuvent permettre de repenser les pratiques.