

Enjeux de l'open data des décisions de justice

Colloque dématérialisé organisé par la Compagnie nationale des experts médecins de justice (CNEMJ), le 22 octobre 2020.

Le 29 juin 2020, une étape importante a été franchie concernant l'open data des décisions de justice avec la publication du décret n°2020-797 relatif à la mise à disposition du public des décisions de juridictions judiciaires et administratives ; ce décret d'application de l'article 33 de la loi n° 2019-2022 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice pose les règles générales quant à l'accès aux décisions de justice et leur publicité.

Avant la publication de ce décret, plusieurs moments clés ont rendu possible l'émergence de cet open data des décisions de justice :

- le 7 octobre 2016, la promulgation de la loi n°2016-1321 pour une République numérique, dite loi Lemaire, a permis de poser la première pierre de l'open data. La loi avait, en effet, institué en ses articles 20 et 21 la mise à disposition du public à titre gratuit de l'ensemble des décisions de justice en précisant qu'elle devrait se faire « dans le respect de la vie privée des personnes concernées ».
- Puis, en novembre 2017, Loïc Cadiet, professeur à l'École de droit de la Sorbonne, à qui avait été confiée en mai 2017 une mission sur l'open data des décisions de justice, a déposé un rapport dans lequel sont formulées des recommandations pour la mise en œuvre de cet open data.
- Enfin, le 23 mars 2019, la promulgation de la loi de programmation 2018-2022 et de réforme pour la justice a permis que soient déterminés les grands principes de l'open data des décisions de justice.

« L'objectif de l'open data des décisions de justice est d'assurer la transparence de la justice, renforcer la confiance dans l'institution judiciaire et assurer une meilleure sécurité juridique¹ », a indiqué Jean-Michel Sommer, président de chambre à la Cour de cassation et directeur du Service de la documentation, des études et du rapport de la Cour de cassation. Ses propos ont ouvert le

colloque dématérialisé organisé par la Compagnie nationale des experts médecins de justice (CNEMJ), présidée par Mary-Hélène Bernard.

Selon le décret du 29 juin 2020, la Cour de cassation est responsable de la mise à la disposition du public, sous forme électronique et gratuitement, des décisions de justice rendues par les juridictions judiciaires². Il incombe, par ailleurs, à la haute juridiction de mettre en œuvre l'occultation de certains éléments des jugements et des arrêts mis en ligne : nom et prénom des personnes physiques, de manière automatique, et les éléments susceptibles de permettre une éventuelle réidentification de ces personnes et de porter atteinte à leur sécurité ou à leur vie privée ainsi qu'à celles de leur entourage, cela étant décidé par le président de la formation ayant rendu la décision.

« Au sein de la Cour de cassation, nous souhaitons accompagner les juridictions du fond dans cette action afin que des règles relativement harmonisées soient mises en place. Cela permettra d'occulter un nombre plus ou moins important d'entités dans les décisions judiciaires en fonction de la sensibilité des contentieux », a décrit Jean-Michel Sommer. « Des contentieux comme la réparation du préjudice corporel ou le droit des mineurs sont, par exemple, globalement bien plus sensibles que la construction ou d'autres sujets. La Cour de cassation s'emploie à rédiger un texte qui pourrait s'apparenter à un guide d'aide à l'anonymisation, destiné aux juridictions. »

S'agissant des nom et prénom des magistrats et membres du greffe, Jean-Michel Sommer a rappelé que « le parlement a décidé qu'ils ne seront pas occultés ; en revanche, il est possible d'obtenir cette occultation lorsque la révélation de leur identité serait de nature à créer un risque pour leur sécurité ou leur vie privée. Il faut par ailleurs signaler que le profilage des juges est puni par la loi, et qu'il n'est ainsi pas autorisé de donner aux usagers de la justice la possibilité de choisir leur juge ou de vérifier les pratiques des juges qui

pourraient être amenés à trancher leur litige ».

L'occultation de l'identité des magistrats et des membres du greffe ne sera pas décidée par le magistrat qui a rendu la décision mais par le président de la juridiction concernée.

Cette anonymisation est « le principal défi technique » de cet open data des décisions de justice, a signalé, lors de ce colloque dématérialisé, Estelle Jond-Necand, conseillère référendaire et directrice du projet Open data au Service de la documentation, des études et du rapport de la Cour de cassation.

Cette opération doit se dérouler en deux étapes :

- d'abord, un moteur de recherche automatique va anonymiser les décisions de justice grâce à des algorithmes utilisant l'intelligence artificielle ;
- ensuite, une relecture humaine des décisions anonymisées automatiquement sera effectuée par des agents du Service de la documentation, des études et du rapport de la Cour de cassation, afin de vérifier qu'il n'y a pas d'erreurs.

Estelle Jond-Necand précise : « Nous avons déjà identifié les catégories qui sont repérées automatiquement par le moteur d'anonymisation : nom et prénom des personnes physiques (parties ou tiers) ; nom et prénom des professionnels (juges, greffiers, avocats) ; différentes dates civiles (naissance, décès et mariage) ; les adresses ; les personnes morales ; les cadastres ; les plaques d'immatriculation ; les comptes bancaires ; les numéros d'identification INSEE ; les numéros de téléphone/fax/mail. Ces différentes catégories sont repérées par le moteur de recherche, et nous avons la possibilité de décider si nous les anonymisons ou non. Actuellement, nous analysons le taux d'erreurs commises par le moteur de recherche automatique ; l'objectif est d'avoir, au fur et à mesure, des relectures humaines plus ciblées en fonction des taux d'erreurs obtenus ».

Concilier notamment la mise à disposition des décisions de justice avec la protection des données personnelles.

Jean-Michel Sommer a, en outre, au cours de ce colloque numérique, dressé la liste des différents enjeux de cet open data judiciaire :

- concilier la mise à disposition des décisions de justice avec la protection des données personnelles. « *Il est, en quelque sorte, nécessaire d'articuler deux ordres de législation : le droit nouveau de l'open data judiciaire et le droit plus ancien et traditionnel de la protection des données (la loi informatique et libertés de 1978 et ses textes d'application, le règlement général sur la protection des données – ou RGPD –, etc.)* » ;
- rendre les informations accessibles et créer des outils méthodologiques. « *Il s'agit de transformer des données en connaissances. Il faut réussir à transformer cette masse énorme de décisions qui seront mises en ligne – à terme, potentiellement 4 millions par an –, en connaissances. Nous souhaitons, pour notre part, créer grâce à ces nouvelles connaissances qu'on pourra acquérir des outils méthodologiques pour nos collègues magistrats mais aussi éventuellement pour les experts judiciaires* » ;
- analyser les pratiques professionnelles. « *Il faut que les juges, notamment, réfléchissent à ce qu'est la jurisprudence. Jusqu'à présent, la jurisprudence en tant que telle, ce n'est que les arrêts de la Cour de cassation ; les arrêts des cours d'appel ou les jugements de première instance n'en font pas partie. Cependant, les rapports de rapporteurs à la Cour de cassation ou les avis des avocats généraux à la Cour de cassation font référence assez régulièrement à des éléments issus d'arrêts rendus par les cours d'appel et disponibles sur la base de données JuriCa* » ;
- modifier possiblement l'office du juge. « *Cette modification pourrait être entraînée dans un avenir proche par cette diffusion généralisée des décisions de justice. Il pourrait y avoir une analyse par territoire ou juridiction des décisions rendues, et il n'est pas impossible que prochainement les parties et leurs avocats plaident une juris-*

prudence de telle juridiction locale rendue un ou deux mois auparavant ».

Jean-Michel Sommer a, par ailleurs, évoqué l'utilisation possible par des éditeurs privés et le monde de la LegalTech de ces décisions de justice publiées sur Internet, signalant qu'il existe « *un débat qui n'est pas encore clos et qui s'ouvre même à peine sur le niveau de contrôle que nous pourrions avoir concernant la réutilisation de ces données par des entreprises de la LegalTech. Le Conseil d'État s'est, pour sa part, prononcé en faveur d'un encadrement par une autorité administrative indépendante qui pourrait participer à une régulation de la réutilisation par des acteurs privés économiques des données issues de l'open data. La Cour de cassation n'a pas encore arrêté de position définitive sur ce sujet* ».

Maître Frédéric Bibal, avocat en droit du dommage corporel s'est, de son côté, interrogé quant à l'intérêt de l'open data des décisions de justice pour les experts du fait. Quoi qu'il en soit, M^e Frédéric Bibal a évoqué l'idée, en constatant le travail entrepris pour l'anonymisation des décisions de justice, que ces avancées techniques soient utilisées pour la création, dans un avenir plus ou moins proche, d'une base de données de rapports d'expertise anonymisés : « *Les experts, notamment, pourraient alors accéder au raisonnement auxquelles ils sont eux-mêmes confrontés. Il faut cependant avoir conscience des enjeux que cela représenterait étant donné les questions de secret et d'intimité qu'on peut trouver dans chaque expertise.* »

Open data des décisions de justice et droit du dommage corporel

M^e Frédéric Bibal a, par ailleurs, signalé que « *le droit du dommage corporel sert souvent de terrain d'essai lorsqu'il s'agit de traitement des données. En matière de dommage corporel, cette question du traitement des données n'est pas nouvelle, elle existe depuis que sont apparus des recueils de jurisprudence* ».

La volonté de ne pas être pris de court par les évolutions entraînées par le traitement des données a conduit l'Association nationale des avocats de victimes de dommages corporels (ANADAVI) – à laquelle appartient M^e

Frédéric Bibal – à réaliser un guide des bonnes pratiques. « *Ces bonnes pratiques dans le traitement des données en matière de dommage corporel ont été détaillées dans un article de Maître Aurélie Coviaux³ ; il s'agit d'une sorte d'exigence minimale pour que l'outil puisse être accepté et acceptable* », précise M^e Frédéric Bibal.

Ce guide a été établi en prenant pour modèles les cinq principes de la Commission européenne pour l'efficacité de la justice (CEPEJ) du Conseil de l'Europe pour une Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement : principe de respect des droits fondamentaux (notamment garanties d'accès au juge et d'un procès équitable) ; principe de non-discrimination (avec le danger de reproduction voire d'aggravation de discriminations par l'outil numérique) ; principe de qualité et sécurité (sources certifiées et données intangibles, environnement technologique sécurisé) ; principe de transparence, de neutralité et d'intégrité intellectuelle (accessibilité des méthodologies de traitement des données, autoriser les audits externes) ; principe de maîtrise par l'utilisateur (le professionnel de la justice doit à tout moment pouvoir revenir aux décisions et données judiciaires ayant été utilisées pour produire un résultat et continuer à avoir la possibilité de s'en écarter au vu des spécificités de l'affaire concrète).

M^e Frédéric Bibal a présenté ainsi les principes proposés par l'ANADAVI :

1. Assurer une recherche permettant de choisir, *a minima*, la juridiction considérée. « *Il existe une dualité juridictionnelle en France avec les ordres judiciaire et administratif ; une des conditions de pertinence d'un outil est de pouvoir sélectionner l'ordre juridictionnel, voire plus précisément la juridiction, parce qu'on y trouve des méthodologies différentes. Certains affirment que la nomenclature Dintilhac⁴ est aujourd'hui universelle, mais dans la pratique, on constate notamment que des juges administratifs peuvent travailler avec l'avis Lagier⁵ qui classe les préjudices quelque peu différemment.* »
2. Procéder de façon systématique à l'analyse des prétentions et moyens des parties pour ne retenir que les décisions pertinentes. « *En*

matière civile, nous nous trouvons dans un système dans lequel les juges ne peuvent pas allouer moins que ce qui est proposé par la partie défenderesse – l'assureur, par exemple – ni davantage que ce qui est demandé par la partie demanderesse – la victime, le plus souvent ; par conséquent, si on ne connaît pas les bornes de la décision – ce qui a été demandé et ce qui a été proposé –, les données n'ont strictement aucun sens. Un chiffrage anormalement bas peut être la conséquence d'une demande qu'on peut considérer comme insuffisante et un chiffrage anormalement haut peut être dû au fait qu'un défendeur qui n'a pas fait son métier correctement a fait une proposition qu'on peut qualifier de trop importante. Ne pas connaître l'ensemble des données d'un litige ne serait donc pas pertinent dans cette situation. »

3. Seules les sommes allouées avant imputation des éléments de minoration ne peuvent être prises en considération. « On pense ici en particulier aux questions de partage de responsabilité ou de perte de chance : une donnée n'a pas de sens si on ne restitue pas un chiffrage intégral et si on donne par exemple seulement la moitié d'un montant parce qu'on n'a pas pris en considération le fait qu'il y ait eu un partage de responsabilité. »
4. Doivent être exclues de toute analyse de masse les données chiffrées des postes de préjudices dont le mode d'évaluation ne découle pas d'une appréciation médico-légale. « Cela n'a pas de sens, par exemple, d'inclure des frais ou des pertes de salaire, parce que cela dépend de chaque cas particulier. »
5. Offrir une recherche multi-critères propre à assurer la personnalisation de la réparation. « En matière de dommage corporel, par exemple, il existe certes des éléments communs entre les différents domaines – victimes d'un attentat, d'une erreur médicale, d'un accident domestique ou d'un accident de la circulation –, mais il y a également toute une série d'informations caractéristiques – les circonstances de survenue du dommage, l'environnement familial, l'origine sociale, un lieu d'habitation plus ou moins isolé –, qui sont tout aussi essentielles pour l'appréciation de la réparation. »

Les questionnements concernant le décret DataJust

M^e Frédéric Bibal a ensuite évoqué le décret « DataJust » du 27 mars 2020⁶, en soulignant que « ce texte ne signifie pas, comme certains l'ont craint, la mise en circulation d'un algorithme, d'une base de données ou d'un référentiel pour l'évaluation des préjudices en matière de dommages corporels. C'est en réalité un texte, paru de manière inopportune un dimanche en plein confinement, autorisant une équipe de recherche à travailler sur un algorithme qui devrait permettre : d'évaluer les politiques publiques en matière de responsabilité civile ou administrative ; d'élaborer un référentiel indicatif d'indemnisation des préjudices corporels ; d'aider les parties à évaluer le montant de l'indemnisation à laquelle elles peuvent prétendre afin de favoriser un règlement amiable des litiges ; d'aider les juges amenés à statuer sur ce type de contentieux à acquérir des informations et de la documentation utiles. C'est un travail dont l'importance peut être considérable pour la réparation des préjudices nés du dommage corporel, avec d'un côté le risque d'une barémisation inacceptable, mais d'un autre côté, si des garde-fous sont instaurés, la possibilité d'avoir accès à des données pertinentes et de façon plus pratique que cela peut être le cas actuellement ».

À la suite de la publication de ce décret, plusieurs associations de victimes de dommages corporels ont rédigé une motion commune. Dans ce texte, elles ont affirmé :

- « leur attachement au principe de la réparation intégrale et de la prise en compte individualisée de la situation de chaque victime » ;
- « leur attachement à une diffusion loyale et exhaustive de la jurisprudence rendue par les tribunaux et à l'accessibilité gratuite à cette jurisprudence » ;
- « leur opposition absolue à tout barème d'indemnisation, y compris sous la forme édulcorée d'un référentiel, qui figera à terme l'évolution nécessaire des réparations accordées aux victimes et empêchera l'individualisation de ces réparations ».

M^e Frédéric Bibal a souligné « le danger d'un caractère autoprescripteur du barème : à partir d'un certain moment, le barème n'est plus l'illustration de décisions de justice ; au contraire, les décisions deviennent l'application du barème, et cela peut se produire assez rapidement. Il est nécessaire d'avoir à la fois une vision de ce qui se fait le plus souvent en matière de jugements, mais aussi de pouvoir découvrir ceux qui ont fait différemment afin de voir si cela peut nous faire progresser ».

Il a ajouté, craignant que l'utilisation d'algorithmes ne conduise à mettre en place un système de moyennes ou de médianes laissant de côté les cas les plus graves : « Il faudrait un système de fourchettes, qui ne doit surtout pas éliminer les décisions extrêmes, au contraire. Dans les bases de données, les décisions extrêmes doivent être signalées comme telles ; mais leur présence permettra aux utilisateurs se trouvant dans des situations extrêmes – avec un chiffrage particulièrement élevé ou particulièrement bas – de découvrir les raisons pour lesquelles telle décision a été rendue dans des cas extrêmes autres que le leur. »

Jacques Cohen, expert agréé par la Cour de cassation en immunologie biologique, a également fait part de sa crainte que cet outil « conduise à bloquer la situation. Cela peut entraîner la mise en place d'un système normatif qui gèle une jurisprudence, ne permette pas son évolution, et représente la "fin de l'histoire". Cela ressemble à un film de Federico Fellini, *Prova d'orchestra* : lors d'une répétition, les membres d'un orchestre se disputent, et finalement le chef d'orchestre est renvoyé et remplacé par un énorme métronome. Avec cet outil, au lieu de s'améliorer, l'orchestre ne joue pas très intelligemment, et cela ne résout pas davantage les conflits ».

Au moment de conclure les échanges de ce colloque numérique de la CNEMJ, M^e Frédéric Bibal a souhaité souligner, « pour garder espoir », que ces nouveaux outils « peuvent nous permettre de découvrir des pépites dans la jurisprudence, d'avoir de très bonnes surprises. Quand il m'arrive, par exemple grâce à un confrère, de découvrir certaines décisions de juges du fond de première instance et que je peux y déceler un vrai travail de recherche, des originalités, des efforts de rédaction, je constate que tout n'est pas bâclé dans la justice, qu'il y a de belles décisions, des

gens qui font leur travail parfaitement. Une des conséquences de ces outils d'open data, qui nous feront accéder à toutes les décisions de justice, pourrait être de nous permettre d'éliminer le pire et de mettre en valeur le meilleur ».

NOTES

1. Selon ce principe, la justice doit être compréhensible et prévisible.
2. Le Conseil d'État est, lui, responsable de la diffusion des décisions rendues par les juridictions administratives.
3. « Pour un guide des bonnes pratiques dans le traitement des données chiffrées de la réparation du préjudice corporel », par Maître Aurélie Coviaux, *Gazette du Palais*, 22 janvier 2019.
4. La nomenclature Dintilhac fixe des principes pour l'évaluation de la réparation résultant d'infractions ayant causé des dommages corporels à une victime. Elle fixe vingt postes pour les victimes directes et sept postes pour les victimes indirectes. Elle n'a pas de force obligatoire, elle est simplement indicative et un instrument pour les praticiens. Elle n'est pas non plus exhaustive, c'est-à-dire que le juge pourra décider d'indemniser un poste qui ne figure pas dans la nomenclature.
5. Dans l'« avis Lagier » (CE, Section, avis, 4 juin 2007, Lagier et Consorts Guignon, n°s 303422 304214), le Conseil d'État avait instauré sa propre classification des postes de préjudice. La différence fondamentale entre cette classification et celle suggérée par la nomenclature Dintilhac réside dans l'absence de distinction entre la période antérieure à la consolidation et celle postérieure à la consolidation.
6. Décret n° 2020-356 du 27 mars 2020 portant création d'un traitement automatisé de données à caractère personnel dénommé « DataJust ».

Calendrier prévisionnel du développement de l'open data des décisions de justice :

Jean-Michel Sommer a indiqué qu'un calendrier progressif, en trois temps, a été mis en place pour le développement de l'open data des décisions de justice :

- dans un premier temps, à échéance de l'automne 2021, toutes les décisions de la Cour de cassation seront mises à disposition sur un site Internet rénové de la haute juridiction. Un moteur de recherche présent sur ce site permettra d'accéder aux arrêts rendus par la Cour de cassation en matières civile et pénale (soit environ 13 à 15 000 arrêts par an) ;
- dans un deuxième temps, au premier trimestre 2022, doivent être mises en ligne, en open data, toutes les décisions civiles, sociales et commerciales des cours d'appel (soit environ 230 à 240 000 décisions par an) ; le moteur de recherche de la Cour de cassation permettra d'accéder à ces décisions ;
- dans un troisième temps, dont l'échéance n'est pas encore fixée, il est prévu de mettre en place un open data des décisions de toutes les autres juridictions de première instance, y compris spécialisées – les conseils de prud'hommes, les tribunaux de commerce, les tribunaux judiciaires ; cela représente entre 3,9 et 4 millions de décisions par an.

Jean-Michel Sommer a indiqué que la base de données serait alimentée à partir des deux bases qui sont administrées par la Cour de cassation : la base Jurinet des arrêts de la Cour de cassation et la base JuriCa des arrêts des cours d'appel. La Cour de cassation gère seule ces deux bases de données dans des conditions qui ont été autorisées par la CNIL en juillet 2012 (délibérations nos 2012-245 et 2012-246).

Estelle Jond-Necand a précisé que « dès qu'il y aura une modification dans les bases JuriCa et Jurinet, cela sera reporté dans la base de données en open data, après un passage par le processus d'anonymisation. Il n'y aura donc pas de discordances entre les bases de données ». Elle a, par ailleurs, signalé qu'une suppression des bases Jurinet et JuriCa n'est pour l'instant pas prévue : « Actuellement, nous créons seulement une nouvelle base de données en open data qui sera sécurisée ».



Création de la CNEMJ
Jacques HUREAU
1993



Colloques annuels à Paris
Réunions en province et en outre-mer
Entretiens présentiels et dématérialisés
Membre institutionnel de l'EEEE
Membre du CNP de médecine légale et d'expertises médicales



En ligne et accessible à tous :

- Actualités,
- Activités de formation,
- Références de nos experts,
- Livrets des colloques,
- Liens utiles aux experts

Nous rejoindre pour :

- Connaître les attentes des magistrats
- Maîtriser les procédures de l'expertise
- Discuter de cas pratiques d'expertise
- Communiquer avec les avocats
- Avoir les conseils amicaux des confrères

Nous joindre par mail à : CNEMJ@orange.fr