

La Cour de cassation à l'épreuve du numérique et de l'intelligence artificielle

Il entre dans les missions de la Cour de se prononcer sur la conformité des jugements attaqués à la règle de droit. Il lui incombe ensuite de faire connaître ses décisions. Le digital, y compris dans sa composante d'IA, s'implante progressivement dans ces deux tâches : le traitement des pourvois et la diffusion de la jurisprudence.

Publié le 9 février

2021

Le numérique ou le digital est, pour une institution comme la Cour de cassation, un défi paradoxal.

La science qui s'y déploie n'est pas exacte et l'intelligence qui s'y exerce n'est pas artificielle. Cette

science et cette intelligence reposent sur le raisonnement de ses juges, sur les choix jurisprudentiels, le plus souvent opérés entre points de vue opposés. Elles s'attachent à la singularité de chaque procès.

Le poids des traditions et la rigueur de ses procédures pourraient offrir une résistance à l'évolution des méthodes de travail de la Cour. En traversant ses couloirs, ses salles d'audience et de délibéré, on peine à croire que les lieux sont équipés de postes et de connexions informatiques ainsi que, crise sanitaire oblige, de systèmes de visioconférence installés pour l'organisation de ses réunions.

Pourtant, la Cour de cassation a résolument pris le chemin de la modernité en explorant les potentialités du numérique et, depuis peu, en prêtant un intérêt et une attention soutenus et raisonnés aux possibilités offertes par l'intelligence artificielle (IA). C'est ce chemin que ces lignes proposent d'emprunter.

Le traitement de pourvois

De longue date, la Cour de cassation est utilisatrice du numérique pour son organisation et pour le traitement de ses pourvois. Elle a d'abord entrepris la dématérialisation de ses procédures puis celle des documents.

L'informatisation des tâches se poursuit, elle est à présent aidée par des techniques d'IA. Empreinte d'une détermination constante d'accompagner ces évolutions

de réflexions partagées avec les métiers sur le sens des tâches, la Cour met ces outils au service de l'évolution des méthodes de travail mais aussi des perspectives d'évolution des missions fondamentales de la Cour de cassation.

De l'informatique de gestion des procédures à la dématérialisation des documents

À l'origine, comme d'autres institutions et administrations, la Cour de cassation a eu recours à l'informatique essentiellement à des fins d'aide à la gestion des procédures au service des greffes, pour les fonctions de recherche des dossiers, de consultation des données, d'édition des courriers, de dactylographie, de statistiques. L'objectif de cette dématérialisation était, pour l'essentiel, l'amélioration des conditions de travail. Une approche documentaire avait néanmoins déjà permis de créer une première base de jurisprudence au service des magistrats.

Au cours des quinze dernières années, les méthodes de travail et le traitement des procédures ont été profondément transformés. L'architecture du système d'information et de communication a été refondée et, si le développement de l'informatique de gestion s'est poursuivi, la mise en place d'une véritable communication électronique avec les partenaires de la juridiction a permis l'institution d'un dossier numérique partagé au sein de la juridiction.

La finalité a été ici de permettre aux utilisateurs, magistrats, fonctionnaires de la juridiction, avocats, d'accéder et de consulter simultanément le dossier, de récupérer tout ou partie d'un texte, de télécharger et de classer des pièces ou encore d'archiver la procédure. La dématérialisation "dynamique" de l'organisation du travail a ainsi simplifié les tâches et réduit les délais de traitement, au bénéfice des utilisateurs et des justiciables.

Elle s'est poursuivie et parachevée par l'instauration à la Cour de la signature électronique des décisions dès 2013. À ce moment, la Cour s'est trouvée la première Cour suprême d'Europe à se doter d'un tel dispositif.

Une méthode pragmatique, progressive et concertée a présidé à la réalisation et au développement de ce projet de dématérialisation, s'appuyant sur un dialogue institutionnalisé et constant avec les utilisateurs internes et externes de la juridiction. En se familiarisant avec la dématérialisation, tous les acteurs des procédures conduites à la Cour ont apporté leur expérience et ont formulé leurs demandes d'amélioration et d'évolutions, réunis au sein de groupes associant les métiers de la Cour et/ou les avocats aux Conseils.

Par cette méthode, des interrogations sont apparues, exprimant la volonté des équipes informatiques de comprendre les contraintes et les finalités procédurales des règles du pourvoi et du rôle de la Cour de cassation et, réciproquement, la conscience des acteurs du procès de l'existence de marges de progression et d'amélioration apportées par les nouvelles technologies.

À titre d'exemple, les objectifs de la signature électronique des décisions méritent aujourd'hui d'être questionnés. Autorisée à la Cour depuis 2013, comme il a été vu, l'instauration de cette signature a été décidée principalement dans le but de parachever la dématérialisation du procès jusqu'à son terme. Or cette évolution importante ne prend sens que si elle s'inscrit dans la chaîne du procès, qui comprend sa phase d'exécution dématérialisée des décisions par les avocats et les huissiers de justice. Ainsi, il faut se souvenir que la dématérialisation sert les finalités de la procédure.

L'intelligence artificielle au service de la dématérialisation des procédures et du traitement des pourvois ?

Comme il a été dit, le digital s'est mis au service de l'administration de la justice et du traitement des pourvois. L'IA est désormais sollicitée pour simplifier des tâches chronophages jusqu'ici confiées à des magistrats et à des juristes. Elle pourrait demain faciliter l'identification et la sélection des pourvois présentant un intérêt ou une difficulté particulière.

L'automatisation de l'orientation vers les chambres

La Cour de cassation a développé et vient de mettre en production au sein de son service de documentation, des études et du rapport un moteur de pré-orientation des pourvois vers les chambres spécialisées. Un algorithme d'apprentissage automatique a été entraîné

sur une centaine de milliers de mémoires en demande déjà orientés ces dernières années. L'algorithme prend en entrée les moyens et, plus spécifiquement, le moyen du mémoire en demande des avocats appelé mémoire ampliatif, pour prédire vers quelle chambre pré-orienter le pourvoi. Les résultats obtenus sur les mémoires ampliatifs de l'année 2020, non présents dans le jeu d'entraînement, sont de 87% de bonne pré-orientation.

Le numérique et l'IA au service de la mise en œuvre de circuits différenciés ?

Trois circuits différenciés de traitement des pourvois viennent d'être institués à la Cour, à l'initiative de sa première présidente. L'objectif est d'ajuster les moyens employés pour résoudre le litige en fonction du degré de complexité de l'affaire et de réserver l'expression de la Cour de cassation, par des arrêts motivés en style direct, aux décisions présentant un apport normatif. Sont instaurés un circuit court, permettant de juger rapidement les recours dont la solution s'impose, un circuit approfondi à l'issue duquel sont jugées les affaires les plus complexes et un circuit intermédiaire. Une cellule d'orientation des dossiers vers l'un des trois circuits a été créée au sein de chaque chambre.

L'efficacité de cette orientation, distincte de l'orientation vers les chambres présentée plus haut, suppose la mise en place d'outils permettant de signaler le plus tôt possible les pourvois importants. À cet égard, la nécessité d'établir une connexion informatique permettant la transmission par les juridictions du fond des variables descriptives des affaires entrantes est apparue. C'est pourquoi l'intégration de la nomenclature

des affaires civiles (la NAC) dans le système d'information de la Cour de cassation, pour mieux connaître les natures d'affaires traitées et relier les contentieux de la Cour avec ceux des juridictions du fond vient d'être mise en production. Sur le modèle de la NAC, une nouvelle nomenclature des affaires orientées vers les chambres (NAO), hiérarchisée et antisismique, c'est-à-dire indépendante des changements éventuels dans les attributions des chambres, permettra d'améliorer la connaissance des affaires traitées par la Cour, d'accompagner leur orientation dans un circuit adapté et facilitera la désignation d'un rapporteur spécialisé.

La Cour s'est également engagée dans des programmes de recherche destinés à faciliter la détection des pourvois les plus importants de façon à les orienter dans les circuits de traitement approfondis. En partenariat avec l'ordre des avocats au Conseil d'État et à la Cour de cassation, elle explore les potentialités de l'intelligence artificielle aux termes d'une convention conclue avec des équipes de recherche d'HEC Paris et de l'École polytechnique. Un grand nombre de pièces de procédure et d'arrêts, préalablement pseudonymisés, a été mis à disposition des équipes, afin d'identifier les arguments et les questions juridiques, de rechercher des connexités et rapprochements et de tenter d'objectiver la notion de complexité et/ou d'importance d'une affaire. Une autre recherche, moins avancée, a été confiée à une équipe de l'Institut national de recherche en informatique et en automatique (l'INRIA), portant sur la détection,

conceptuellement très délicate, des divergences de jurisprudence au sein de la Cour et/ou avec les juridictions du fond.

Le processus de diffusion de la jurisprudence

La Cour achève parallèlement la dématérialisation de la plupart de ses publications, bulletins, lettres de chambre et rapport annuel. Elle est également devenue responsable de traitements de données personnelles.

Ainsi, aux termes des articles R. 433-3 et R. 433-4, le Service de documentation et d'études tient une base des décisions de la Cour de cassation et des décisions rendues par les autres juridictions présentant un intérêt particulier, ainsi qu'une base de données distincte rassemblant l'ensemble des arrêts rendus par les cours d'appel. Ces bases, dénommées JuriCAca (voir plus bas) pour les arrêts des cours d'appel et Jurinet pour ceux de la Cour de cassation et pour les décisions des autres juridictions de l'ordre judiciaire présentant un intérêt particulier, n'ont pas été créées dans l'optique de l'*open data* mais pour les besoins des magistrats dans l'accomplissement de leurs tâches juridictionnelles. Elles ont été autorisées par une délibération n° 2012-245 du 19 juillet 2012 de la CNIL. Actuellement, 872 501 décisions (dont 711 100 provenant de la Cour de cassation) sont stockées dans la base Jurinet et 2 167 610 décisions dans la base Jurica.

La loi n° 2016-1321 du 7 octobre 2016, puis la loi n° 2019-222 du 23 mars 2019 et le décret n° 2020-797 du 29 juin 2020 ont posé le principe de la mise à la disposition du public des décisions de l'ordre judiciaire à titre gratuit sous forme électronique, soit l'*open data* des décisions judiciaires. L'entrée en vigueur de l'*open data* est conditionnée à la parution d'arrêtés.

Afin de protéger leur vie privée et leur sécurité, la loi prévoit que le nom et les prénoms des personnes physiques mentionnées dans la décision, lorsqu'elles sont parties ou tiers, sont occultés préalablement à la mise à la disposition du public. Ce faisant, le législateur a institué un premier niveau d'occultation, une occultation dite socle, obligatoire et dépourvue d'appréciation, des nom et prénoms des personnes physiques, parties ou tiers, mentionnées dans la décision. Il a prévu un second niveau d'occultation, facultatif, dit d'"occultations complémentaires", confié au juge ayant rendu la décision, portant sur tout élément permettant d'identifier les parties, les tiers, les magistrats et les membres du greffe, lorsque la divulgation d'éléments d'identification est de nature à porter atteinte à la sécurité ou au respect de la vie privée de ces personnes ou de leur entourage.

Le décret du 29 juin 2020 a confié à la Cour de cassation la responsabilité de la mise à la disposition du public des décisions judiciaires. L'*open data* introduit un changement d'échelle considérable dans la mission traditionnelle de diffusion de la jurisprudence.

Actuellement, moins de 15 000 décisions (principalement de la Cour) sont diffusées chaque année sur Légifrance. L'objectif est de diffuser plus de 3 millions de décisions par an.

En lien avec l'administration centrale de la justice, la Cour s'emploie, dans la mesure des moyens mis à sa disposition, à développer les outils de cette diffusion numérique. Elle se préoccupe pareillement des usages et de la réutilisation qui seront faits de ces données.

Les outils de l'*open data* : l'IA en amont de la diffusion

Pour mettre en œuvre l'*open data*, la Cour s'est dotée d'un moteur performant de pseudonymisation recourant à une solution dite de *machine learning* utilisant une interface d'annotation en cours de refonte. Un moteur de recherche, qui sera installé sur son site internet, est également en cours de développement.

Le développement d'un moteur de pseudonymisation

La pseudonymisation est définie par le Règlement général sur la protection des données (RGPD)n°2016/679 du 27 avril 2016, comme le traitement de données à caractère personnel "de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires, pour autant que ces informations supplémentaires soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données à caractère personnel ne sont pas attribuées à une personne physique identifiée ou identifiable".

Actuellement, seule une partie des arrêts de la Cour de cassation et quelques arrêts de cours d'appel sont mis à disposition en ligne sur le site internet Légifrance, sur lequel sont disponibles les textes légaux, les textes réglementaires et les décisions de cours suprêmes. Pour être mis à disposition en ligne, les arrêts sont pseudonymisés avant cette mise en ligne. Les données identifiant les tiers et parties personnes physiques mentionnées dans les décisions de la Cour de cassation sont occultées par le Service de documentation, des études et du rapport de la Cour de cassation.

Grâce au concours d'Etalab, département de la direction interministérielle du numérique (DINUM) qui coordonne notamment la conception et la mise en œuvre de la stratégie de l'État dans le domaine de la donnée, la Cour a bénéficié d'un premier programme dit d'"Entrepreneurs d'Intérêt Général" (EIG 3), qui permet d'engager pour dix mois des spécialistes de l'IA dans des organismes publics pour relever des défis d'amélioration du service public. Elle a ainsi pu moderniser son moteur de pseudonymisation, au moyen d'une application remplaçant la solution déterministe utilisée jusque-là. Ce moteur développe des techniques d'apprentissage automatique afin d'identifier les données à pseudonymiser dans les décisions de justice avant de les rendre accessibles et réutilisables. Les décisions sont ensuite relues par des agents qui utilisent une interface d'annotation. Cette réalisation pourrait le cas échéant s'appliquer à d'autres types de données publiques.

Afin de faciliter la mise en œuvre de l'*open data*, et conformément à la délibération n° 2020-021 du 6 février 2020 portant avis sur un projet de décret relatif à la mise à disposition du public des décisions des juridictions judiciaires et administratives, la Cour a décidé d'accompagner les juridictions en leur proposant une doctrine d'harmonisation des occultations complémentaires tout en préservant l'objectif d'intelligibilité des décisions. Des groupes de travail ont été constitués à cet effet, installés d'abord à la Cour de cassation, puis auprès des cours d'appel, pour dégager des lignes directrices en considération, notamment, de la sensibilité des contentieux. Ces travaux sont achevés à la Cour et viennent de débiter pour les cours d'appel. L'objectif est de finaliser la réflexion au premier semestre 2021.

L'interface d'annotation

En 2020, la Cour a bénéficié d'un second programme EIG (EIG4), dont l'objectif est de développer un outil d'annotation ergonomique en *open source* pour accélérer la pseudonymisation des décisions de justice et leur mise en *open data* à plus grande échelle. À cette fin, le Service de documentation, des études et du rapport de la Cour accueille en son sein des développeurs et un designer qui, en étroite collaboration avec les magistrats, directeurs de greffe, *data scientists* et agents de la cellule d'anonymisation, s'efforcent d'optimiser et d'améliorer le travail d'annotation de ces agents par la réalisation d'une nouvelle interface d'annotation.

Le développement d'un nouveau moteur de recherche

Le nouveau site internet de la Cour est en cours de réalisation. Il devrait être disponible au cours du dernier trimestre de l'année 2021.

Ce site hébergera un moteur de recherche de la jurisprudence mise progressivement en open data, avant la fin de l'année 2021 pour les arrêts de la Cour de cassation, courant 2022 pour les arrêts des cours d'appel en matière civile, commerciale et sociale puis pour les juridictions du premier degré.

Le fonctionnement et la mise en production de ce moteur devront se faire en cohérence avec la création à intervenir du portail internet placé sous la responsabilité du garde des Sceaux, ministre de la Justice, prévu par le décret du 29 juin 2020 précité.

Les enjeux de l'*open data* : l'avènement de l'IA en aval de la diffusion

Les principaux enjeux de l'*open data*

Le premier des objectifs de la diffusion générale des décisions sous forme numérique est la transparence. Cette transparence participe de la publicité des décisions de justice.

Au-delà de cette finalité, l'*open data* doit favoriser la compréhension des contentieux. Il est primordial de transformer ces données devenues disponibles en connaissance par la réalisation d'outils méthodologiques, d'études et de recherche. Des dispositifs d'aide à la décision voire à la définition de

stratégies, de prévisibilité voire de prédiction des issues contentieuses sont ou seront proposés. Le secteur public comme le secteur privé, emmené par les principaux éditeurs juridiques et les acteurs de la *legal tech*, élaborent d'ores et déjà des solutions au moyen d'algorithmes entraînés et alimentés par des décisions de justice.

Les enjeux de l'*open data* ont maintes fois été décrits : évolution de l'office du juge, cristallisation de la jurisprudence, horizontalisation de celle-ci et émergence possible d'un système de précédent ou de "cas", mutation de la place et de l'autorité des juridictions supérieures, menaces de profilage et existence de biais dans le traitement des données....

Le tri informatique est de nature à créer de la puissance et de la gouvernance algorithmiques. La Cour de cassation ne peut se désintéresser des usages et des risques que pourraient engendrer une réutilisation incontrôlée ou dévoyée des données dont elle est chargée d'assurer la diffusion massive.

Un nécessaire contrôle et une régulation de la réutilisation des données

La Cour propose donc de tenir compte des recommandations du **rapport établi en janvier 2017 par le professeur Loïc Cadiet** à la demande du garde des Sceaux, de celles du Conseil de l'Europe et de la Commission européenne pour l'efficacité de la Justice (CEPEJ), ainsi que des travaux conduits sous l'égide de l'Union européenne. Tous s'accordent à penser qu'une régulation et un contrôle des algorithmes et de la réutilisation des données s'imposent, selon des

modalités à déterminer. Il est essentiel que le recours aux outils dits de justice prédictive ou prévisionnelle soit gouverné par des principes de transparence des algorithmes ainsi que par l'adoption de mécanismes de contrôle par la puissance publique. Une réflexion est engagée sur ce point, conjointement avec le Conseil d'État et la Chancellerie, associant la communauté des réutilisateurs des données, les avocats et l'université.

Abonnez-vous à nos lettres d'information

vous recevrez nos actualités par courriel

S'abonner

En renseignant votre adresse électronique, vous acceptez de recevoir nos actualités par courriel. Vous pouvez vous désinscrire à tout moment à l'aide des liens de désinscription ou en nous contactant.