

RAPPORT

Préparer la Cour de cassation de demain
Cour de cassation et intelligence artificielle

ia



Avril 2025



SYNTHESE

Par lettre de mission en date du 29 mai 2024, le premier président et le procureur général ont institué un groupe de travail interne à la Cour de cassation chargé à la fois d'identifier les cas d'usage possibles de l'intelligence artificielle (IA) à la Cour et de conduire une réflexion sur les interrogations que ces usages pouvaient susciter, au regard de l'office du juge et du respect des droits humains.

1. La méthodologie du groupe : l'association étroite de juristes et de scientifiques

Le groupe de travail, composé de représentants de la première présidence, du parquet général, de chacune des chambres de la Cour et du Service de documentation des études et du rapport (SDER) ainsi que du greffe, a procédé, dans un premier temps, à un **recensement très large des besoins en IA**, en interrogeant l'ensemble des magistrats de la Cour ainsi que des personnels de greffe. Dans un second temps, il s'est attelé à une **analyse de chacun de ces cas d'usage**.

Pour mener à bien cette étude, le **groupe de travail a auditionné une vingtaine de personnes**, représentants du ministère de la justice, de l'Ecole nationale de la magistrature (ENM), des professions judiciaires, de la Commission européenne pour l'efficacité de la justice (CEPEJ) ainsi que des experts en matière d'IA dans le champ de la justice, scientifiques ou juristes, universitaires, éditeurs juridiques et issus des *legaltechs*. Il a aussi **échangé, avec les autres hautes cours nationales (Conseil constitutionnel et Conseil d'Etat) et avec la cour d'appel de Paris**, sur leur appréhension de l'IA et a dressé, avec l'aide du Service des relations internationales de la Cour, un panorama des usages de l'IA dans les **juridictions étrangères**, notamment dans les cours suprêmes.

Il s'est aussi appuyé très largement sur **l'expérience existante de la Cour de cassation en matière d'IA et sur l'expertise des ingénieurs du laboratoire d'innovation du SDER**, membres du groupe de travail, qui ont apporté leur éclairage technique sur les cas d'usage recensés. Il faut souligner, à cet égard, que la Cour de cassation occupe une position unique au sein de l'institution judiciaire car elle a pu constituer une équipe interne de *data science*, dédiée aux travaux d'IA. Dans le cadre du projet d'open data des décisions de justice, la Cour de cassation a conçu dès 2019 un logiciel d'anonymisation reposant sur de l'IA. En 2020, la Cour a mis en place un outil d'aide à l'orientation des mémoires ampliatifs vers les différentes chambres civiles de la Cour qui a considérablement amélioré l'organisation du service. La Cour a également développé un projet ambitieux d'aide à la recherche de divergences de jurisprudence, ayant donné des résultats intéressants en matière de recherche de similarité entre des décisions, de résumé automatisé de décisions et de classification des décisions par mots clés.

Pour ses travaux, le groupe a entendu appréhender largement la notion d'IA sans se limiter à l'examen des potentialités de la seule IA générative. Il s'est aussi référé à des

technologies plus anciennes telles que des systèmes experts ou l'apprentissage automatique ainsi qu'à la combinaison de différents types d'IA.

Tout en exprimant un intérêt pour les services d'IA existant sur le marché, à usage général ou pour les besoins des juristes, le groupe de travail s'est concentré sur l'examen de cas d'usage pouvant être développés par la Cour de cassation pour répondre à ses besoins propres. Il ressort en effet de l'expérience de la Cour que le développement interne d'outils offre des facilités d'agilité dans la gestion des projets, d'interconnexion des profils techniques et métier, et des coûts bien moins importants que le recours à des prestataires. Cela permet également d'assurer une meilleure maîtrise des données et, dès lors, de mieux garantir la fiabilité et la transparence des algorithmes. Cette approche simplifie également la maîtrise de l'hébergement des systèmes d'IA (SIA), sur des serveurs propres de la Cour de cassation ou un *cloud* sécurisé.

2 L'identification des critères d'évaluation des cas d'usage

Les travaux du groupe l'ont conduit à définir des critères d'évaluation des cas d'usage, de nature à favoriser la comparaison de leurs intérêts et limites.

Ces critères sont, en premier lieu, des critères éthiques, juridiques et fonctionnels.

Les critères éthiques identifiés relèvent de la **conciliation du SIA avec les droits fondamentaux**, et de questions relatives au **développement durable**, tenant plus particulièrement à la frugalité des algorithmes, c'est-à-dire à la proportionnalité de la puissance de calcul et donc de l'énergie déployée pour répondre à une problématique.

Les critères juridiques reflètent les contraintes du cadre juridique devant être respecté pour la mise en œuvre d'un SIA. De toute évidence, un soin particulier doit être apporté au **respect du règlement sur l'intelligence artificielle (RIA)**, avec la question, essentielle, de savoir si un système d'intelligence artificielle (SIA) destiné à être mis en œuvre relève ou non d'un SIA à haut risque. À l'étude, il apparaît que même si les SIA destinés à être utilisés par les autorités judiciaires sont en principe à haut risque, les exceptions sont nombreuses et relativement larges, de telle sorte qu'il est probable que la plupart des cas d'usage envisagés dans le domaine de la justice ne relèvera pas de cette catégorie.

De même, le **respect du règlement général sur la protection des données (RGPD)** doit faire l'objet d'une vigilance particulière, dès lors notamment que le SIA requiert le traitement de catégories particulières de données à caractère personnel au sens de l'article 9 du RGPD (données de santé notamment), ce qui sera très fréquent pour les données de justice.

Les critères fonctionnels correspondent à l'intérêt pour les métiers au sein de la Cour, et potentiellement, plus généralement, de l'institution judiciaire de mettre en place un système d'IA. Cet intérêt peut résulter de la perspective soit d'un **gain de qualité**, soit d'un **gain d'efficacité**, qu'il convient d'évaluer précisément.

En second lieu, parallèlement aux critères éthiques, juridiques et fonctionnels, le groupe de travail a identifié des critères techniques et économiques. Les critères techniques visent à évaluer, dans l'ensemble, la faisabilité du projet, sous deux angles. Le premier est celui

de la **disponibilité de la donnée**. En effet, la plupart des projets d'IA repose sur l'exploitation de jeux de données et il est nécessaire, d'une part, de s'assurer que cette donnée est bien disponible, le cas échéant avec l'accord d'une tierce partie, d'autre part, d'évaluer si la donnée est utilisable telle quelle ou si elle nécessite des travaux de transformation ou d'annotation. Le second est celui de la **faisabilité selon l'état de l'art** en matière d'IA.

Les critères économiques permettent d'évaluer le coût d'ensemble d'un projet, en distinguant les ressources humaines (**ressources en science de la donnée, ressources métier, autres ressources de développement**), le coût de la **puissance de calcul** pouvant être nécessaire pour l'entraînement de modèles algorithmiques et le **coût d'exploitation**.

Il est à noter que cette méthodologie d'évaluation pourrait servir de référence aux juridictions qui s'inscriraient dans une démarche similaire et souhaiteraient à leur tour développer et expérimenter des programmes d'IA pour leurs besoins spécifiques.

3 Le recensement des cas d'usage

Le groupe de travail a envisagé cinq catégories de cas d'usage répondant à des besoins distincts. Même si ces cas d'usage font l'objet de déclinaisons particulières pour la Cour de cassation, ils paraissent applicables à l'ensemble des juridictions.

En premier lieu, le groupe de travail a envisagé le recours à une IA pour structurer et enrichir les documents transmis, par exemple en repérant les références normatives ou jurisprudentielles citées dans des mémoires. Il a aussi identifié des cas d'usage spécifiques au greffe.

En second lieu, il a mis en exergue trois grandes familles de cas d'usage. Il s'agit d'abord des cas d'usage relatifs à l'**exploitation des écritures des parties**. Ces cas d'usage permettraient à la Cour d'améliorer son système d'orientation des pourvois, de mieux repérer les connexités matérielles ou intellectuelles entre les pourvois, de faciliter la première analyse du dossier (par exemple pour évaluer la complexité), ou encore de faciliter la compréhension du litige en cartographiant les éléments clés.

Il s'agit ensuite des cas d'usage relatifs à **la recherche et à l'exploitation des bases de données documentaires** qui sont très nombreux et répondent, pour une part, aux besoins propres de la Cour. Pour une autre part, ces cas d'usage s'inscrivent dans le développement de sa mission de diffusion de la jurisprudence auprès des juridictions du fond (dans la continuité des bases de données Jurica et Jurinet accessibles aux magistrats) et auprès du grand public (dans la continuité de la base de données Judilibre regroupant l'ensemble des décisions disponibles en *open data*). Il pourrait s'agir, notamment, d'améliorer la recherche dans ces bases de données par l'implémentation d'outils d'IA de génération augmentée de récupération (RAG) ou d'outils conversationnels. Des fonctionnalités de rapprochements de jurisprudence, et éventuellement d'aide à la recherche de divergences, ou encore d'analyse automatisée de la jurisprudence, peuvent également être envisagées.

Il s'agit enfin des cas d'usage relatifs à **l'aide à la rédaction**. Il peut s'agir de cas d'usage assez simples, d'uniformisation de la rédaction ou encore de respect des normes de saisie,

comme de cas d'usage plus complexes, tels qu'un outil reposant sur la recherche de rapprochements avec des affaires similaires précédemment jugées, ou encore un outil d'aide au traitement de litiges sériels. Il est évident que les outils d'aide à la rédaction posent d'autant plus de questions éthiques qu'ils sont susceptibles d'avoir un impact sur la décision.

A cet égard, il convient de noter qu'aucun besoin d'**aide à la décision** n'a été identifié au sein de la Cour de cassation. Cela tient, notamment, à la forte spécialisation des magistrats de la Cour, en ce qui concerne tant les matières qu'ils traitent que la technique de cassation. Tout autant, le principe même de la mise en œuvre d'outils d'aide à la décision poserait des questions éthiques et juridiques majeures à la Cour, juridiction de dernier recours sur l'application et l'interprétation du droit, probablement de manière plus aigüe encore que pour des usages dans des juridictions du fond.

Le groupe de travail a identifié, par ailleurs, d'autres cas d'usage, comme des applications de traduction et d'interprétariat, ou encore de transcription de la parole, qu'il n'a pas approfondis, considérant qu'ils n'étaient pas spécifiques à la Cour de cassation et relèveraient plus naturellement d'un développement par le ministère de la justice pour être utilisés par l'ensemble des juridictions et par l'administration.

Le groupe de travail a également souligné que nombre de besoins exprimés relevaient en réalité plus des applicatifs métiers de traitement de la procédure que d'applications d'IA.

4 Les pistes dégagées par le groupe de travail pour la poursuite des projets d'IA à la Cour de cassation

Chacun des cas d'usage recensés a été évalué de manière à la fois littérale et chiffrée, évaluation. Cette évaluation a permis au groupe de travail de dégager des pistes de développement de l'IA en mettant en relation, d'une part, les efforts requis pour mettre en œuvre des projets, d'autre part, le retour sur investissement attendu, tels qu'ils peuvent être évalués à ce jour.

Il apparaît, d'abord, que **certains cas d'usage relativement simples et ne posant pas de difficulté juridique ou éthique sont en mesure d'offrir des gains fonctionnels importants, sans nécessiter d'investissements lourds**. Il s'agit, par exemple, de cas d'usage transversaux de structuration et d'enrichissement des documents, qui pourraient bénéficier à de nombreux autres cas d'usage, qu'ils soient relatifs à l'exploitation des écritures des parties, à l'aide à la recherche et l'exploitation des bases de données documentaires ou encore à l'aide à la rédaction.

Ensuite, **certains cas d'usage, plus complexes quant aux critères d'analyse juridiques, éthiques ou techniques, nécessitent un investissement plus lourd qui apparaît pleinement justifié au regard des résultats attendus de manière établie par la littérature scientifique et l'état de l'art**. Il s'agit, notamment, en ce qui concerne l'exploitation des écritures des parties, des cas d'usage de recherche de connexité matérielle et intellectuelle entre mémoires ampliatifs, ainsi que des cas d'usage de détection des précédents et rapprochements de jurisprudence dans l'analyse de ces mémoires. Il s'agit également des cas d'usage relatifs à l'exploitation des bases de

données documentaires, dont la recherche d'éléments jurisprudentiels utiles aux travaux de l'observatoire des litiges judiciaires (questions de droit nouvelles, divergences entre décisions du fond, litiges sériels, questions d'intérêt sociétal).

Enfin, certains projets, en particulier relatifs à l'aide à la rédaction, présenteraient un fort intérêt, notamment en ce qui concerne le traitement des litiges sériels mais revêtent **une complexité particulière, liée notamment à des enjeux éthiques, juridiques et techniques majeurs.**

5 La mise en exergue de principes directeurs pour développer une IA judiciaire

Les réflexions menées au sein du groupe de travail ont permis de prendre la mesure des enjeux du développement de l'IA, au sein de la Cour, comme pour l'ensemble de l'institution judiciaire.

Bien que la réflexion du groupe de travail ait été centrée sur **les cas d'usage** pour les besoins propres de la Cour, plusieurs de ceux dont il a été préconisé le développement **intéressent l'ensemble des juridictions.** Ainsi, l'usage des SIA documentaires de recherche dans les bases de données interne et externe, comme les SIA d'analyse de données jurisprudentielles massives ont vocation à être accessibles aux magistrats des juridictions du fond. De même, les cas d'usage d'aide à la rédaction pourraient aboutir à la création de technologies utilisables ensuite pour l'élaboration d'outils adaptés aux besoins des juridictions du fond.

Le groupe de travail a été convaincu des potentialités qu'offre d'ores et déjà l'IA, potentialités qui devraient être décuplées à l'avenir tant les progrès sont rapides.

Le développement de l'IA constitue ainsi **une opportunité à saisir**, non seulement **pour gagner en efficacité**, à l'heure où l'institution est confrontée à une crise de moyens sans précédent, puisque l'IA est en capacité de réaliser en quelques secondes des tâches chronophages et sans réelle valeur ajoutée quand elles sont assurées par l'humain, mais encore **pour gagner en qualité.** Ce gain de qualité est particulièrement évident, par exemple, s'il s'agit de repérer, dans une grande masse de données, des questions nouvelles et/ou sérielles ou de détecter des divergences de jurisprudence. De telles recherches, qui ne sont pas réalisables raisonnablement sans l'assistance de l'IA, permettraient notamment aux juridictions de rationaliser et de mieux coordonner le traitement des contentieux émergents et d'enrichir le débat juridique et le dialogue des juges.

Cependant, le groupe de travail estime que **l'engagement dans la voie de l'IA est conditionné à la réalisation de prérequis** sur les plans de la technique, de l'éthique et de la gouvernance.

En premier lieu et sur le plan technique, le groupe de travail souligne l'interaction très forte qui existe entre les projets d'IA et **le bénéfice d'un système d'information solide, fiable, centré sur les besoins des utilisateurs et prenant en considération des objectifs de bonne structuration** et conservation des données utiles. Par ailleurs, la question de l'hébergement des services et de la puissance de calcul doit répondre à des impératifs de

souveraineté. La maîtrise directe de l'hébergement, soit sur les serveurs de la Cour de cassation, soit sur des *clouds* obéissant à des exigences de sécurité et de territorialité pour s'inscrire dans le droit applicable à l'Union européenne, permet de garantir l'indépendance des juridictions dans leur fonctionnement, la sécurité des données et la transparence de leur utilisation.

En deuxième lieu, le groupe de travail entend insister sur **la primauté des principes directeurs éthiques**.

En effet, le respect du RGPD et du RIA, pour essentiel qu'il soit, et bien qu'il contribue à protéger les droits fondamentaux, ne suffit pas à garantir un usage de l'IA vertueux, qui préserve la plénitude de l'office du juge et ne vienne pas amoindrir ou fragiliser les équilibres fondamentaux du procès équitable. Le groupe de travail insiste sur la nécessité, en complément des impératifs de conformité juridique, de prendre **ainsi en considération des principes éthiques** relevant de l'impact sur la fonction de juger, de l'impact sur les droits fondamentaux, comme de la transparence, de l'éthique et de la frugalité des algorithmes.

A cet égard, le groupe de travail considère que la **Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires de la CEPEJ**, adoptée en décembre 2018, **demeure un document de référence** ayant conservé toute son actualité malgré les évolutions techniques majeures intervenues depuis. Le groupe de travail a pris en compte les cinq principes dégagés par la charte **de respect des droits fondamentaux, de non-discrimination, de qualité et sécurité des algorithmes, de transparence et d'explicabilité ainsi que de maîtrise humaine des décisions**. Ce dernier principe est essentiel pour conserver intact l'office du juge. Sa mise en œuvre impose de toujours conserver une intervention humaine dans les différentes étapes du processus décisionnel, depuis l'analyse du dossier jusqu'à la rédaction de la décision, le recours à l'IA n'intervenant que pour apporter une aide ponctuelle et toujours à la demande et sous le contrôle du magistrat. Les critères éthiques, qui s'appuient sur les droits humains comme sur des considérations épistémologiques, déontologiques ou sociales, constituent, selon le groupe de travail, **des principes directeurs ou cardinaux, qui précèdent et priment les autres critères fonctionnels, juridiques, techniques et économiques**. Ils doivent dicter le choix des outils d'IA à retenir, mais aussi accompagner leur développement et guider les juges dans leur utilisation, ce qui pose la question de la gouvernance des IA judiciaires.

S'agissant, **en troisième lieu, de la gouvernance**, le groupe de travail a entendu donner quelques repères dont la portée dépasse la Cour.

D'abord, il estime que la méthodologie proposée pour évaluer les cas d'usage de l'IA a vocation non seulement à s'appliquer *ex ante*, pour éclairer les choix et déterminer les programmes d'IA à lancer, mais encore devrait être appliquée de manière continue tout au long du cycle de vie d'un projet, jusqu'à la décision, le cas échéant, de l'abandonner. Afin de mettre en œuvre cette **méthode d'évaluation « de bout en bout »**, le groupe de travail a proposé l'instauration **d'un comité de suivi du développement des SIA à la Cour**, en charge de cette évaluation, composé notamment de magistrats et de scientifiques du laboratoire d'innovation du SDER, ainsi que des autres services de la Cour.

Ensuite, l'effectivité de la mise en œuvre des usages éthiques de l'IA est suspendue à la nécessité pour les utilisateurs, magistrats, fonctionnaires et membres du greffe, d'une acculturation aux enjeux de l'IA. Cette acculturation passe nécessairement par une démarche de **formation des professionnels**. De même, elle suppose d'élaborer un **guide des bonnes pratiques, un référentiel interne ou une « charte éthique »**, à l'instar de celles mises en place par de nombreuses juridictions étrangères.

Enfin, il conviendrait que l'autorité de surveillance compétente pour les juridictions, qui sera prochainement désignée pour l'application du règlement IA, joue pleinement son rôle. En complément, la question de la création, à l'échelle de l'institution, d'un comité consultatif d'éthique pour les SIA judiciaires se posera, comme celle de la place que devra y tenir la Cour de cassation. A cet égard, le groupe de travail souligne l'importance de garantir l'indépendance de ce comité, dans la mesure où il serait notamment consulté sur l'impact sur les droits humains des SIA utilisés par les magistrats dans le cadre de leur activité juridictionnelle.

Pour conclure, ni technophile ni technophobe, le groupe de travail a identifié plusieurs grands axes de développement de l'IA au sein de la Cour, pour son usage propre ou au bénéfice de l'ensemble des magistrats, voire de la communauté des juristes, de manière réaliste et prudente, dans le souci d'utiliser l'IA au mieux de ses potentialités, tout en la laissant à sa juste place, afin de préserver la plénitude de l'office du juge dans un Etat de droit et de garantir la confiance et la maîtrise humaine.

A l'avenir, le rôle de la Cour dans le développement des SIA judiciaires, si les moyens lui en sont donnés, pourrait s'avérer majeur, compte tenu de l'expertise technique qu'elle a acquise et de ses missions propres, tenant à l'interprétation et l'harmonisation du droit comme à la diffusion de la jurisprudence de l'ensemble des juridictions.

PRÉPARER LA COUR DE CASSATION DE DEMAIN – COUR DE CASSATION ET INTELLIGENCE ARTIFICIELLE

SOMMAIRE

SYNTHESE	1
SOMMAIRE	10
PRÉSENTATION DU GROUPE DE TRAVAIL.....	14
LETTRE DE MISSION.....	14
COMPOSITION DU GROUPE DE TRAVAIL.....	15
Présidente du groupe de travail	15
Rapporteurs	15
Représentants de la première présidence	15
Représentants du parquet général près la Cour de cassation	15
Représentants des chambres de la Cour de cassation	15
Représentants du greffe.....	15
Laboratoire d'innovation du SDER.....	16
METHODOLOGIE RETENUE PAR LE GROUPE DE TRAVAIL	16
INTRODUCTION	18
LES EVOLUTIONS HISTORIQUES DE L'INTELLIGENCE ARTIFICIELLE	18
DEFINITION DE L'IA RETENUE PAR LE GROUPE DE TRAVAIL	20
CHOIX DES CAS D'USAGE ETUDIES : LE DEVELOPPEMENT D'OUTILS INTERNES POUR LES BESOINS DE LA COUR DE CASSATION.....	20
PRINCIPES DIRECTEURS DE LA MISE EN ŒUVRE DE SIA A LA COUR DE CASSATION	23
PARTIE 1 CONTEXTE DE LA MISE EN ŒUVRE DE L'IA POUR LES BESOINS DE L'ACTIVITÉ JUDICIAIRE	26
CHAPITRE 1 – PRESENTATION DE L'IA A LA COUR DE CASSATION	27
Section 1 – Mise en place en 2019 d'un outil de pseudonymisation des décisions de justice.....	28
Section 2 – Mise en place en 2020 d'un outil d'orientation automatique des mémoires ampliatifs : une IA frugale s'appuyant sur des données de qualité.....	29
Section 3 – Le projet « Divergences »	30
Section 4 – Une expertise reconnue aux niveaux national et international.....	31
CHAPITRE 2 – PRESENTATION D'OUTILS ET DE PROJETS JUDICIAIRES BASES SUR L'IA AU NIVEAU NATIONAL OU INTERNATIONAL.....	33
Section 1 – Au niveau national.....	33
Section 2 – Au niveau international	35
§ 1 – Gestion du flux juridictionnel.....	35
§ 2 – Retranscription et interprétariat	36
§ 3 – Accès à la justice et information du justiciable.....	37
§ 4 – Analyse de dossiers et enrichissement des données de procédure.....	37

§ 5 – Aide à la rédaction des décisions	37
§ 6– Rédaction de chartes éthiques et d'encadrement des usages de l'IA.....	38
PARTIE 2 PRESENTATION DES CRITERES D'EXAMEN	42
CHAPITRE 1 – PRESENTATION DES CRITERES D'EVALUATION FONCTIONNELS, ETHIQUES ET JURIDIQUES.....	43
Section 1 – Critères éthiques.....	43
§ 1 – Conciliation du système d'IA étudié avec les droits fondamentaux.....	43
§ 2 – Développement durable : frugalité des systèmes d'intelligence artificielle	45
Section 2 – Critères juridiques.....	47
§ 1 – Mise en conformité des SIA à la réglementation IA.....	47
§ 2 – Mise en conformité avec la réglementation sur les données personnelles	50
Section 3 – Critères fonctionnels	52
§ 1 – Gain de qualité.....	52
§ 2 – Gain d'efficacité	53
CHAPITRE 2 – PRESENTATION DES CRITERES D'EVALUATION TECHNIQUES ET ECONOMIQUES.....	54
Section 1 – Critères techniques	55
§ 1 – Disponibilité de la donnée.....	55
§ 2 – Preuve de faisabilité	57
Section 2 – Critères économiques	58
§ 1 – Temps de ressources en science de la donnée.....	58
§ 2 – Temps de ressources métiers	59
§ 3 – Autres ressources de développement	60
§ 4 – Puissance et temps de calcul pour l'entraînement des modèles	62
§ 5 – Coût d'exploitation	63
PARTIE 3 PRÉSENTATION ET EVALUATION DES CAS D'USAGE IDENTIFIÉS.....	65
CHAPITRE 1 – CAS D'USAGE TRANSVERSAUX : STRUCTURATION ET ENRICHISSEMENT DES DOCUMENTS.....	66
CHAPITRE 2- L'EXPLOITATION DES ECRITURES DES PARTIES	68
Section 1 – Traitement des pourvois par le SDER	68
§ 1 – Orientation des mémoires ampliatifs	68
§ 2 – Pré-signalé automatique des pourvois.....	70
Section 2 - Détection de connexité matérielle et intellectuelle	71
§ 1 – Connexité matérielle	71
§ 2 – Connexité intellectuelle	72
Section 3 – Première analyse du dossier par les chambres	73
§ 1 – Détection des critères de complexité	73
§ 2 – Estimation du temps de traitement	74
§ 3 – Détection des précédents et rapprochements de jurisprudence au sein des décisions rendues par la Cour de cassation	75
Section 4 - Cartographie du litige	76
CHAPITRE 3 – AIDE A LA RECHERCHE ET EXPLOITATION DES BASES DE DONNEES DOCUMENTAIRES.....	78
Section 1 – Exploitation de bases de données pour un usage interne	78
§ 1 – Création d'un moteur de recherche sémantique	80
§ 2 – RAG (Retrieval Augmented Generation)	81
§ 3 – Outil conversationnel	81
Section 2 Exploitation des bases de données pour un usage public.....	83
§ 1 – Création d'un moteur de recherche sémantique	84
§ 2 – RAG	84
§ 3 – Outil conversationnel	85

Section 3 Exploitation des travaux parlementaires	85
Section 4 – Veilles automatiques	87
§ 1 – Veille automatique européenne s'appuyant sur la jurisprudence de la Cour européenne des droits de l'Homme (CEDH) et la Cour de justice de l'Union Européenne (CJUE).....	87
§ 2 – Veille doctrinale avec synthèses et chaînage de la jurisprudence de la Cour	88
§ 3 – Veille médiatique.....	89
Section 5 – Identification des rapprochements ou divergences de jurisprudence	90
§ 1 – Rapprochements jurisprudentiels ou divergences en droit interne et européen	91
§ 2 – Identification de rapprochements ou divergences de jurisprudence entre les juridictions du fond	91
§ 3 – Identification de rapprochements ou de divergences entre les décisions des juridictions du fond et la jurisprudence de la Cour de cassation	92
§ 4 – Identification de rapprochements ou divergences de jurisprudence au sein de la Cour de cassation	93
Section 6 – Analyse de la jurisprudence du fond afin de déterminer des tendances et positions dominantes sur des questions précises	93
§ 1 – Tendances et positions dominantes au sein des juridictions du fond dans les contentieux chiffrés	93
§ 2 – Identification de questions nouvelles ou de litiges sériels – soutien aux missions de l'observatoire des litiges judiciaires.....	95
§ 3 – Outil général d'analyse de la jurisprudence du fond, notamment massive	97
CHAPITRE 4 - AIDE A LA REDACTION	99
Section 1 – Aide à la rédaction des rapports, avis et décisions	99
Section 2 – Outil d'aide à la rédaction pour les litiges sériels	100
Section 3 – Uniformisation et vérification de la rédaction (formulaires et guides de rédaction, normes de saisie, orthographe, syntaxe, qualité de la langue, complétude).....	101
§ 1 – Intégration des formulaires et guides de rédaction.....	101
§ 2 – Vérification automatique des normes de saisie	103
§ 3 - Harmonisation des styles rédactionnels	104
§ 4 – Vérification des oublis de moyens, de la cohérence des textes visés et des moyens	105
Section 4 – Génération automatique de résumés	105
§ 1 – Proposition de résumés automatiques pour la rédaction des travaux préparatoires et mise en valeur des faits, de la procédure, des moyens, de l'article 700 du code de procédure civile et des avis	106
§ 2 – Publication : proposition de résumé ou de sommaires pour publication au Bulletin ou à la Lettre de chambre, résumés de vulgarisation en vue d'une diffusion pour le grand public	107
Section 5 - Amélioration du titrage	108
CHAPITRE 5 – CAS D'USAGE DU GREFFE	110
Section 1 – Aide à la rédaction d'ordonnances de désistement	110
Section 2 – Aide à l'établissement par le greffe des certificats de non-pourvoi...	111
CONCLUSION.....	113
ANNEXES	117
ANNEXE 1 : LES LARGES MODELES DE LANGAGE : UNE REVOLUTION DANS L'INTELLIGENCE ARTIFICIELLE.....	117

I. COMPRENDRE LES MODELES DE LANGAGE	117
II. LES ACTEURS MAJEURS ET LEURS INNOVATIONS	117
III. IMPLICATIONS POUR LA PRATIQUE JURIDIQUE	118
IV. APPLICATION AUX CAS D'USAGE DE LA COUR	119
V. PANORAMA DES PRINCIPAUX MODELES DE LANGAGE	120
A. Les Modèles Commerciaux Propriétaires.....	120
B. Les Modèles Open Source de Nouvelle Génération	120
C. Les modèles spécialisés en langues européennes et en français.....	122
ANNEXE 2 : ETUDES DU SERVICE DES RELATIONS INTERNATIONALES.....	125
ANNEXE 3 : LISTE DES AUDITIONS	146
CONSEIL D'ÉTAT.....	146
CONSEIL CONSTITUTIONNEL	146
ÉCOLE NATIONALE DE LA MAGISTRATURE	146
JURIDICTIONS DE L'ORDRE JUDICIAIRE.....	146
LEGALTECH ET EDITEURS JURIDIQUES	147
MINISTERE DE LA JUSTICE	147
ORDRE DES AVOCATS AU CONSEIL D'ETAT ET A LA COUR DE CASSATION	148
PRESIDENCE DE LA REPUBLIQUE	148
THINK TANK.....	148
UNIVERSITAIRES	148
ANNEXE 4 : GLOSSAIRE	149

PRÉSENTATION DU GROUPE DE TRAVAIL

Lettre de mission

COUR DE CASSATION

Le procureur général

Le premier président

Paris, le 29 mai 2024

Le premier président de la Cour de cassation
Le procureur général près la Cour de cassation
à
Madame Sandrine Zientara, présidente de chambre,
directrice du service de documentation, des études et du rapport.

Objet : Mission « Préparer la Cour de cassation de demain - Cour de cassation et intelligence artificielle »

Avec le big data et avec les dernières avancées scientifiques en la matière, le développement de l'intelligence artificielle ne cesse de croître de jour en jour : de nombreuses applications qui peuvent dorénavant simplifier, voire automatiser les tâches du quotidien ont vu le jour. En outre, il est communément admis que l'intelligence artificielle a pour conséquence de transformer de nombreux secteurs d'activité.

La question du recours à l'intelligence artificielle dans la Justice a fait et fait encore l'objet de nombreux travaux et réflexions, aux niveaux national, européen et international.

La Cour de cassation doit poursuivre et approfondir sa réflexion pour identifier les usages qu'elle pourra faire de l'intelligence artificielle, notamment dans les domaines de la documentation, de la recherche et de l'aide à la décision, ainsi que pour l'orientation, l'identification et le traitement des affaires qui lui sont soumises, tant au niveau du SDER que dans les chambres. Cette réflexion doit faire leur part aux interrogations que l'usage de l'intelligence artificielle peut faire naître au regard de l'office du juge et des principes fondamentaux qui s'imposent à lui.

A cette fin, nous vous remercions de bien vouloir animer un groupe de travail, composé notamment de représentants des chambres de la Cour, du parquet général, d'ingénieurs du laboratoire d'innovation du SDER, du cabinet de la première présidence et du cabinet du procureur général. Dans la réalisation de cette mission, vous vous appuyerez sur les moyens du SDER, en particulier ceux du pôle 2. Vous veillerez à prendre connaissance des pratiques du Conseil d'Etat, des cours suprêmes étrangères et des juridictions européennes et à entendre toute personne, appartenant notamment au monde universitaire, susceptible d'enrichir la réflexion du groupe de travail.

Il vous appartiendra, notamment, d'indiquer quelles pourraient être les modalités de déploiement d'outils d'intelligence artificielle, les enjeux, les limites et les bonnes pratiques susceptibles d'être développées.

Ce groupe pourrait se réunir la première fois en juin 2024 et rendre ses conclusions en décembre 2024.


Rémy Heitz


Christophe Soulard

Composition du groupe de travail

Présidente du groupe de travail

- M^{me} Sandrine Zientara, présidente de chambre, directrice du service de documentation, des études et du rapport de la Cour de cassation (SDER).

Rapporteurs

- M. Edouard Rottier, conseiller référendaire, adjoint à la directrice du SDER, chef du pôle du numérique et de la diffusion de la jurisprudence.
- M. Matthieu Allain, auditeur, chef du bureau du droit du numérique et de la protection des données.

Avec l'assistance de Mme Emma Palombi, Mme Maïlys Prats, M. Marcus Woodcock, assistants de justice, et de Mme Inès Haber, M. Giacomo Modolo et M. Galaad Tuzio, stagiaires.

Représentants de la première présidence

- M. Edouard de Leiris, conseiller, chargé de mission du premier président.
- M^{me} Clémence Bourillon, auditrice, cheffe du service des relations internationales.

Représentants du parquet général près la Cour de cassation

- M^{me} Sonya Djemni-Wagner, avocate générale à la chambre criminelle, chargée de mission relations internationales et institutionnelles, recherche et universités et environnement numérique.
- M. Hugues Adida-Canac, avocat général, 2^{ème} chambre civile.
- M. Samuel Aparisi, avocat général référendaire, 1^{ère} chambre civile.

Représentants des chambres de la Cour de cassation

- M^{me} Charlotte de Cabarrus, conseillère référendaire, 1^{ère} chambre civile.
- M. Pascal Montfort, conseiller référendaire, 2^{ème} chambre civile.
- M. Axel-Nicolas Choquet, conseiller référendaire, 3^{ème} chambre civile.
- M^{me} Laure Comte, conseillère référendaire, chambre commerciale.
- M^{me} Marie-Pierre Lanoue, conseillère référendaire, chambre sociale.
- M. Olivier Violeau, conseiller référendaire, chambre criminelle.

Représentants du greffe

- M^{me} Caroline Savier, directrice de greffe de la Cour de cassation.

- Mme Fouzia Adeli, cheffe de service du greffe des chambres civiles.

Laboratoire d'innovation du SDER

- M. Amaury Fouret, *data scientist*.
- M. Paul Déchorgnat, *data engineer*.

Méthodologie retenue par le groupe de travail

La composition du groupe de travail traduit l'ambition d'une réflexion pluridisciplinaire et transversale sur les usages de l'intelligence artificielle à la Cour de cassation en s'appuyant sur l'expérience du SDER en matière d'intelligence artificielle et en réunissant des représentants de la première présidence, du parquet général, des chambres et du greffe de la Cour de cassation.

Le groupe de travail a visé à donner une portée particulièrement concrète à ses travaux en se fixant pour objectif principal de dresser un inventaire le plus exhaustif possible, selon l'état de l'art actuel, des cas d'usage pouvant présenter un intérêt pour la Cour de cassation. Ces cas d'usage ont été évalués suivant des critères fonctionnels, éthiques, juridiques, techniques et économiques. Cette évaluation permettra de mesurer l'intérêt de leur développement et de leur mise en œuvre.

Les membres du groupe de travail, pour éclairer leur réflexion, ont échangé dès le début de leurs travaux sur les caractéristiques techniques de l'intelligence artificielle, à partir de présentations effectuées par les *data scientists* du laboratoire d'innovation du SDER.

Les différents membres de la mission ont ensuite pu faire remonter au groupe de travail une liste de cas d'usage identifiés par les chambres ou services auxquels ils étaient rattachés. Ces cas d'usage ont été portés à la discussion des membres de la mission.

Le groupe de travail a également souhaité nourrir ses réflexions d'un large spectre d'auditions. Il a entendu, parfois en groupe restreint, une vingtaine de personnes, représentants du ministère de la justice, de l'Ecole nationale de la magistrature (ENM), des professions judiciaires, de la Commission européenne pour l'efficacité de la justice (CEPEJ) ainsi que des experts en matière d'IA dans le champ de la justice, scientifiques ou juristes, universitaires, éditeurs juridiques et issus des *legaltechs*. Il a aussi échangé, avec les autres hautes cours nationales (Conseil constitutionnel et Conseil d'Etat) et avec la cour d'appel de Paris.

Le groupe de travail a, en outre, pu bénéficier de l'expertise du service des relations internationales de la Cour, qui a réalisé une étude sur les usages de l'IA dans les juridictions étrangères, en particulier les cours suprêmes.

L'ensemble de ces auditions et travaux a notamment permis aux membres du groupe d'avoir une vision nationale et internationale de l'état de l'art et des possibilités offertes par les différentes technologies d'IA en matière de justice, tant par les acteurs publics que

les acteurs privés. Elles ont également permis d'affiner la liste des cas d'usage et les critères retenus pour leur examen.

INTRODUCTION

Les évolutions historiques de l'intelligence artificielle

L'intelligence artificielle (IA), en tant que domaine scientifique, regroupe l'ensemble des sciences, théories et techniques ayant pour but la conception et le développement de machines capables de reproduire les capacités cognitives d'un être humain.

Le terme « intelligence artificielle » a été introduit en 1956 lors de la conférence de Dartmouth, considérée comme l'acte fondateur de ce domaine. Les premières recherches se sont concentrées sur l'IA symbolique, une approche qui vise à reproduire le raisonnement humain en manipulant des symboles et en appliquant des règles logiques prédéfinies. Cette méthode a notamment donné lieu à la création de systèmes experts capables de résoudre des problèmes spécifiques en suivant des ensembles de règles établies par des experts métier notamment via des arbres de décision. Les limites de cette approche sont cependant apparues rapidement, dès les années 1960, les systèmes d'intelligence artificielle (SIA) n'étant pas en capacité de s'adapter à des situations imprévues ou de faire face à des tâches nécessitant une phase d'apprentissage à partir de données.

La révolution numérique des années 1990, la puissance de calcul et les masses de données qui s'y associent ont favorisé l'émergence d'une nouvelle approche fondée sur l'apprentissage automatique et posent les prémices des SIA modernes. Cette approche, qui n'est plus fondée sur la définition de règles par l'humain mais sur l'apprentissage automatique par la machine à partir d'un ensemble de données, s'est particulièrement développée dans les années 2010 avec la disponibilité plus large de données en masse et de la puissance de calcul nécessaire.

Ces deux approches, par règles et par apprentissage, ne sont pas exclusives, et des systèmes dits hybrides offrent notamment aux systèmes experts un second souffle au contact de l'apprentissage automatique. Il s'agit, en combinant les technologies, de proposer des applications performantes exploitant le meilleur des deux mondes.

Dans le paysage contemporain de l'IA, on distingue ainsi plusieurs catégories de systèmes aux caractéristiques et applications distinctes.

L'IA générative, qui relève de l'apprentissage automatique et dont le déploiement massif au cours des années 2020 constitue la dernière révolution de l'usage de l'IA, se démarque par sa capacité à créer de nouveaux contenus, qu'il s'agisse de textes ou d'images. Cette technologie s'appuie sur des modèles sophistiqués, tels que les grands modèles de langue (LLM) entraînés sur d'immenses corpus de données. Ces systèmes fonctionnent selon une approche statistique : ils ne comprennent pas intrinsèquement le sens des informations traitées, mais déterminent les séquences les plus probables en réponse aux requêtes des utilisateurs, appelées « prompts ». A la différence des systèmes experts, le

risque d'inexactitudes, de réponses erronées, ou « hallucinations » dans le langage courant, est donc nécessairement accru.

L'IA classificatoire ou prédictive représente un autre pan majeur du domaine. S'appuyant également sur l'apprentissage automatique, ces systèmes se spécialisent dans la catégorisation d'entités, de faits ou d'événements selon des typologies prédéfinies. Leur capacité à « prédire » repose sur l'analyse statistique des données historiques pour déterminer les probabilités d'occurrence d'événements futurs, ouvrant ainsi de nombreuses perspectives d'application dans des domaines variés.

Le développement et le déploiement de ces systèmes d'IA nécessitent des ressources computationnelles significatives. Les modèles les plus sophistiqués requièrent des infrastructures spécialisées, équipées de processeurs adaptés tels que les GPU (unités de traitement graphique), capables de traiter en parallèle d'importants volumes de calculs. Néanmoins, il existe aujourd'hui une grande variété de modèles plus légers ou « frugaux », accessibles sur des infrastructures plus modestes, démocratisant ainsi l'usage de ces technologies.

Il convient de noter que les systèmes d'IA par apprentissage automatique se distinguent fondamentalement des algorithmes déterministes traditionnels de l'IA symbolique. Alors que ces derniers produisent invariablement le même résultat pour une entrée donnée, suivant un cheminement prédéfini et parfaitement reproductible, les systèmes d'IA modernes intègrent une dimension probabiliste et adaptative qui leur confère à la fois leur puissance et leur spécificité. Cette caractéristique, si elle ouvre de nouvelles possibilités, implique également une attention particulière dans leur mise en œuvre et leur utilisation.

Il est essentiel de distinguer deux phases fondamentales dans le cycle de vie des systèmes d'intelligence artificielle par apprentissage : la phase d'apprentissage et la phase d'inférence. La phase d'apprentissage peut elle-même se décomposer en deux approches distinctes, à savoir l'entraînement initial (ou « *from scratch* »), extrêmement gourmand en ressources, et l'affinement (ou « *fine-tuning* »), qui consiste à adapter un modèle pré-entraîné à des données spécifiques. Si l'entraînement initial des grands modèles requiert des infrastructures massives mobilisant des milliers de GPU sur plusieurs semaines ou mois, l'affinement peut être réalisé avec des ressources plus modestes, accessibles à des institutions comme la Cour de cassation disposant de GPU dédiés. Cette approche par affinement permet d'adapter précisément des modèles existants aux spécificités du langage juridique et aux particularités des données judiciaires françaises, sans nécessiter les investissements colossaux de l'entraînement initial. La phase d'inférence, qui correspond à l'utilisation du modèle une fois affiné pour produire des réponses ou des prédictions, nécessite encore moins de ressources. Cette distinction est déterminante dans l'appropriation de ces technologies par les institutions judiciaires : en privilégiant l'affinement de modèles pré-entraînés, il devient possible de développer des solutions d'IA adaptées aux besoins spécifiques de la justice, tout en maintenant une maîtrise des coûts et des ressources computationnelles.

Définition de l'IA retenue par le groupe de travail

Le groupe de travail retient pour ses travaux la notion d'intelligence artificielle dans son acception la plus large, englobant l'ensemble des technologies exposées ci-dessus. Cette approche holistique vise à appréhender de manière exhaustive les enjeux, les opportunités et les défis que présente l'intégration de ces systèmes dans nos environnements professionnels et sociétaux, sans se limiter aux enjeux spécifiques des évolutions les plus récentes de l'IA, telles que les perspectives ouvertes par l'usage de l'IA générative.

Dès les premiers échanges, il est toutefois apparu au groupe de travail la nécessité de bien distinguer ce qui relève de l'IA et ce qui relève de projets informatiques classiques ou d'outils bureautiques, dont certaines fonctionnalités peuvent également reposer sur des règles déterministes simples. Ainsi, le groupe de travail a constaté que plusieurs besoins remontés initialement auprès de ses membres relevaient non pas d'usage de l'IA, mais plutôt d'améliorations des outils informatiques et bureautiques mis à disposition.

La définition de l'IA retenue par le groupe de travail est à rapprocher de celle adoptée par le règlement européen sur l'intelligence artificielle du 12 juillet 2024 (règlement IA)¹.

L'article 3 de ce règlement définit le système d'intelligence artificielle (SIA) comme « *un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels* ».

Les contours de la notion de SIA au sens du règlement IA sont précisés par le considérant 12 du règlement et par des lignes directrices de la Commission européenne sur la définition des systèmes d'IA². Ils visent à couvrir un spectre large d'utilisations de l'IA, tout en excluant des systèmes fondés sur une programmation informatique basique ne présentant aucune autonomie dans leur fonctionnement.

Choix des cas d'usage étudiés : le développement d'outils internes pour les besoins de la Cour de cassation

Le groupe de travail a pu observer avec grand intérêt, au travers notamment de démonstrations et d'auditions, les développements récents des outils d'IA existant sur le

¹ Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle

² [Lignes directrices de la Commission sur la définition des systèmes d'IA établie par le règlement IA, approuvées par la Commission le 6 février 2025 mais non encore formellement adoptées au 28 février 2025](#)

marché, qu'il s'agisse d'outils d'IA généraliste ou d'outils offerts aux juristes par les acteurs privés de l'édition juridique et de la *legaltech*.

Il est apparu toutefois au groupe de travail qu'il ne lui appartenait pas de se prononcer sur les mérites comparés de ces différents services. D'une part, une telle comparaison supposerait de mettre en œuvre un processus de test approfondi de ces produits par des juristes de la Cour et sur une durée de quelques jours, ce que la mission n'a pas été en mesure de faire. D'autre part, ces produits se développent et évoluent très rapidement et sont, en ce qui concerne les services offerts aux juristes, le plus souvent adossés à des fonds documentaires et éditoriaux propres qu'ils viennent valoriser, ou sont fondés sur une approche originale visant à répondre aux besoins des juristes et en particulier des avocats³.

Le groupe de travail note toutefois que la Cour de cassation, en tant que juridiction suprême de l'ordre judiciaire, devrait en principe pouvoir souscrire à tous les abonnements des éditeurs afin de bénéficier d'outils d'IA utiles à l'exploitation des fonds doctrinaux de référence et d'avoir un accès à une documentation exhaustive, qui ne soit pas inférieure à celle dont disposent certains cabinets d'avocats.

Par ailleurs, le groupe de travail relève que l'utilisation des outils offerts sur le marché doit s'analyser différemment, selon que cette utilisation implique ou non la transmission d'informations couvertes par le secret professionnel ou la protection des données personnelles. Dans cette dernière hypothèse, une analyse approfondie serait nécessaire afin de s'assurer que le traitement de ces données réponde aux exigences de sécurité et de conformité juridique, ce qui pourrait nécessiter une mise en œuvre de ces services sur des serveurs contrôlés par l'autorité judiciaire. En l'absence d'une telle analyse, il est nécessaire d'exclure toute transmission d'informations, quelles qu'elles soient, provenant de pièces de procédure.

Il apparaît également que ces services, s'ils peuvent répondre à de larges spectres de besoins, ne sauraient couvrir tous ceux qui sont spécifiques à l'institution judiciaire. Dans ce contexte, le groupe de travail a concentré ses travaux sur les SIA qui pourraient être développés par la Cour de cassation, seule ou avec des partenaires, afin de répondre à des besoins spécifiques à la Cour de cassation ou communs aux partenaires.

La Cour de cassation a l'avantage, à cet égard, de disposer d'une équipe interne de *data science* et d'une expérience réussie en matière de développement interne de projets d'IA.

Elle bénéficie en outre d'une direction des systèmes d'information, qui est un atout majeur, seul à même de permettre d'intégrer de façon coordonnée et harmonieuse des SIA dans le système d'information juridictionnel de la Cour.

³ Me Kami Haeri, président de la commission « *Impact de l'intelligence artificielle sur les professions du Droit* » du club des juristes, a ainsi pu témoigner lors de son audition de l'engagement croissant des cabinets d'avocats avec les outils d'IA pour effectuer principalement des tâches mécaniques telles que le référencement l'analyse forensique ou la traduction.

Il ressort notamment de cette expérience que le développement interne d'outils offre des facilités d'agilité dans la gestion des projets, d'intercompréhension des profils techniques et métier, et des coûts bien moins importants que le recours à des prestataires. Cette approche simplifie également la maîtrise de l'hébergement des SIA, sur des serveurs propres de la Cour de cassation ou un cloud sécurisé.

Les besoins spécifiques de la Cour de cassation ressortissent à l'ensemble de ses missions, qu'elles relèvent de sa place institutionnelle au sommet des juridictions de l'ordre judiciaire, de l'activité juridictionnelle des chambres et du parquet général ou encore de son rôle de diffusion et de valorisation de la jurisprudence, notamment dans le cadre du projet d'*open data* des décisions de justice.

Le groupe de travail s'est concentré sur les outils de traitement automatisé du langage naturel (TAL) qui correspondent à ses besoins spécifiques. Le groupe de travail a ainsi exclu de son champ d'étude des SIA pouvant présenter un grand intérêt pour l'ensemble des juridictions et dont la mise à disposition paraîtrait plutôt relever du ministère de la justice, tels que des outils de transcription (outils dits « *speech-to-text* ») ou des outils de traduction et d'interprétariat. Le groupe de travail relève, à cet égard, que le ministère de la justice a d'ores et déjà identifié de tels besoins et entend y répondre.

Sur cette base, le groupe de travail a identifié trois grandes familles de cas d'usage, portant sur l'exploitation des écritures des parties, l'aide à la recherche et à la documentation ainsi que l'aide à la rédaction. Parallèlement à ces familles de cas d'usage, le groupe de travail a identifié des cas d'usage transversaux pouvant bénéficier à de nombreux projets, et un cas d'usage spécifique aux missions du greffe (voir *infra*, partie 3, présentation et évaluation des cas d'usage identifiés).

Le groupe de travail a exclu d'analyser de manière approfondie des cas d'usage qui relèveraient de l'aide à la décision, en l'absence de besoin identifié par les magistrats des chambres. Il est apparu que les magistrats de la Cour de cassation, experts dans les matières qu'ils traitent et dans la technique de cassation, ne sont pas en recherche d'aide à la prise de décision.

Principes directeurs de la mise en œuvre de SIA à la Cour de cassation

Le groupe de travail a identifié des critères, éthiques, juridiques, fonctionnels, techniques et économiques, qui sont présentés dans la deuxième partie du rapport et à l'aune desquels il a examiné chacun des cas d'usage retenus aux termes de l'expression des besoins.

Certains de ces critères non seulement ont été utilisés par le groupe comme outils d'évaluation des cas d'usage, mais encore constituent des prérequis indispensables au développement des SIA.

Ainsi, le choix de développer des cas d'usage suppose d'identifier clairement des gains fonctionnels, qu'ils relèvent d'un gain de qualité ou d'un gain d'efficacité dans les missions de la Cour de cassation. L'intérêt de l'utilisation de l'IA pour atteindre ces gains fonctionnels doit en outre être apprécié à l'aune des autres moyens éventuellement susceptibles d'être mobilisés pour atteindre les mêmes objectifs.

En outre, il importe de souligner la nécessité, pour les membres de la Cour de cassation, de bénéficier d'outils numériques adaptés à leurs besoins, alors que l'application Nomos et le logiciel d'éditique qui lui est associé (Corel Wordperfect) sont anciens et qu'un projet de refonte (Nomos 2) est actuellement en cours de mise en œuvre par la Cour de cassation.

Le groupe de travail constate la nécessité d'une bonne articulation entre les outils informatiques, bureautiques et les systèmes d'intelligence artificielle, dans la mesure où une informatique et une gestion des données de qualité sont essentielles au succès de projets d'intelligence artificielle, qui ont généralement vocation à se nourrir des bases de données existantes et à s'intégrer aux applications informatiques métier (ou, si une intégration est impossible ou difficile à mettre en œuvre, à être implantées de manière complémentaire dans les systèmes informatiques). Il apparaît ainsi essentiel que les projets informatiques ou les choix applicatifs effectués prennent en considération les opportunités ou limitations qu'ils pourraient apporter à des projets d'intelligence artificielle.

De même, le développement et la mise en œuvre de cas d'usage de l'IA doit répondre à des principes directeurs juridiques. En particulier, il conviendra de s'assurer que les cas d'usage peuvent répondre aux exigences du règlement IA et du RGPD⁴ et d'évaluer la charge induite par la nécessité de conformité à ces instruments. Toutefois, les SIA judiciaires dont relèveraient les cas d'usage retenus par le groupe de travail demeurent autorisés et leur éventuelle classification au sein des SIA à haut risque impliquerait alors

⁴ [Règlement \(UE\) 2016/679 du Parlement européen et du Conseil du 27 avril 2016](#) relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données)

simplement certaines exigences notamment de contrôle, de fiabilité, de transparence ou de sécurité (voir infra, partie 2, chapitre 1, section 2).

Le respect de ce cadre juridique, pour essentiel qu'il soit, et bien qu'il contribue à protéger les droits fondamentaux, ne suffit donc pas à garantir un usage de l'IA vertueux, qui préserve la plénitude de l'office du juge et ne vienne pas amoindrir ou fragiliser les équilibres fondamentaux du procès équitable. Le groupe de travail insiste sur la nécessité, en complément des impératifs de conformité juridique, de prendre ainsi en considération des principes éthiques relevant de l'impact sur la fonction de juger, de l'impact sur les droits fondamentaux, comme de la transparence, de l'éthique et de la frugalité des algorithmes.

A cet égard, le groupe de travail considère que la Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires, adoptée en décembre 2018, demeure un document de référence ayant conservé toute son actualité malgré les évolutions techniques majeures intervenues depuis.

D'une manière générale, le groupe de travail estime que le développement de l'usage de l'IA doit s'accompagner d'une réflexion critique constante sur ses limites et ses risques, spécialement au sein d'une cour suprême. En effet, dans l'exercice de son office normatif, la Cour de cassation ne peut aborder le raisonnement juridique comme un pur syllogisme, un calcul objectif et prévisible. Elle est souvent amenée, dans un esprit d'ouverture, à développer des approches interdisciplinaires, conséquentialistes et contextuelles pour prendre en compte les évolutions notamment sociales, sociétales, économiques, environnementales ou technologiques. Le risque d'appauvrissement du raisonnement juridique, comme celui de voir se figer la jurisprudence, les solutions produites par l'IA reposant, par hypothèse, sur l'analyse des décisions déjà rendues, doit être précisément mesuré. De même, des analyses épistémologiques et éthiques sur l'identification des tâches qui, par nature, ne devraient et ne pourraient pas être modélisables par une IA, comme par exemple la mise en balance des valeurs, mériteraient d'être engagées.

Le groupe de travail relève aussi que la question, technique, de l'hébergement des services et de la puissance de calcul, répond directement à des impératifs juridiques et éthiques et notamment de souveraineté. La maîtrise directe de l'hébergement, soit sur les serveurs de la Cour de cassation, sur des *clouds* répondant à des exigences de sécurité et de territorialité pour s'inscrire dans le droit applicable à l'Union européenne, permet de garantir l'indépendance des juridictions dans leur fonctionnement, la sécurité des données et la transparence de leur utilisation.

Par ailleurs, les objectifs poursuivis par ces prérequis et l'effectivité de leur mise en œuvre est suspendue à la nécessité pour les utilisateurs, magistrats, fonctionnaires et membres du greffe d'une acculturation aux enjeux de l'IA éthiques, techniques, juridiques, épistémologiques et anthropologiques. Cette acculturation passe nécessairement par une démarche de formation des professionnels, dont ils sont au demeurant demandeurs.

De même, elle suppose d'élaborer un guide des bonnes pratiques, un référentiel interne ou une « charte éthique », à l'instar de celles mises en place par de nombreuses

juridictions étrangères⁵, notamment pour sensibiliser les utilisateurs au respect des règles de sécurité et de protection des données personnelles, limiter les biais pour conserver la maîtrise humaine de l'outil, ou encore se conformer efficacement aux exigences de vérifications ou de transparence. Le groupe de travail préconise donc l'adoption d'un tel document pour l'usage de l'IA au sein de la Cour.

⁵ Voir sur ce point Chapitre 2 – Section 2 – Paragraphe 6 : Rédaction de chartes éthiques et d'encadrement des usages de l'intelligence artificielle

PARTIE 1

CONTEXTE DE LA MISE EN ŒUVRE DE L'IA POUR LES BESOINS DE L'ACTIVITÉ JUDICIAIRE

Chapitre 1 – Présentation de l'IA à la Cour de cassation

L'IA représente une opportunité et une nécessité pour l'institution judiciaire, confrontée qualitativement à une complexité croissante des contentieux et quantitativement à une masse de données inédite résultant notamment de la mise en œuvre de l'*open data*⁶.

Le volume des décisions actuellement diffusées par la Cour de cassation sur son site internet Judilibre dépasse à ce jour le million de décisions de justice rendues par la Cour de cassation, l'ensemble des cours d'appel et une partie des tribunaux judiciaires.

Cette masse de données considérable ne constitue cependant que l'annonce d'un changement radical d'échelle à l'horizon de décembre 2025, date à laquelle l'ensemble des décisions publiques transmises par les tribunaux judiciaires en matière civile, sociale et commerciale, les conseils de prud'hommes et les tribunaux de commerce alimenteront la base de données Judilibre à raison de près d'un million de décisions par an selon les projections actuelles. Suivra ensuite le déploiement de l'*open data* des décisions définitives rendues par les juridictions de jugement de première instance et d'appel en matière pénale.

Cet horizon immédiat achève de projeter la Cour de cassation, comme l'ensemble des acteurs judiciaires, dans la sphère des données de masse auxquelles il apparaît objectivement impossible de faire face humainement.

Le recours à l'IA comme moyen de collecte, de traitement et de publication de ces données a donc constitué une nécessité pour la Cour. Elle constitue également une opportunité dans l'exercice de sa fonction « traditionnelle » de régulation et d'harmonisation de la jurisprudence sur le territoire.

Pour ce faire, la Cour de cassation peut s'appuyer sur les ressources internes et le savoir-faire technique qu'elle a pu développer depuis 2019 au sein de son laboratoire d'innovation, service composé actuellement de 10 profils techniques (développeurs, *devops*, *data scientists*...) et rattaché au SDER.

Trois algorithmes ont ainsi été développés ces dernières années en interne, parfois en partenariat avec des institutions publiques, afin de faire face à des besoins identifiés par la Cour de cassation.

⁶ Le décret n° 2020-797 du 29 juin 2020 prévoit en effet, pour la mise à disposition du public des décisions des juridictions judiciaires que « *la Cour de cassation est responsable de la mise à la disposition du public, sous forme électronique, des décisions de justice rendues par les juridictions judiciaires, dans les conditions définies à l'article L. 111-13 ainsi qu'au présent chapitre et à l'article R. 433-3. Les décisions sont mises à la disposition du public dans un délai de six mois à compter de leur mise à disposition au greffe de la juridiction.* »

Sur ces trois modèles d'IA développés par la Cour, deux sont actuellement en production, tandis que les recherches exploratoires sur le dernier projet se poursuivent.

Section 1 – Mise en place en 2019 d'un outil de pseudonymisation des décisions de justice

Il incombe à la Cour de cassation, lors du processus de publication des décisions judiciaires en *open data*, d'assurer techniquement la pseudonymisation des décisions. Concrètement, cela implique d'occulter dans le texte de la décision certains éléments de nature à permettre la réidentification des personnes concernées, tels que des noms et prénoms, adresses, plaques d'immatriculation, numéro de cadastre, coordonnées électroniques, etc.

Pour faire face à ce défi technique, la Cour de cassation a investi significativement dans l'expertise en *data science* en pourvoyant au recrutement de deux *data scientists*. En complément de cette expertise humaine, la Cour de cassation a aussi acquis des équipements informatiques disposant d'une puissance de traitement suffisante grâce notamment à l'inclusion de processeurs graphiques GPU. Ces travaux ont pu être réalisés grâce à la participation de la Cour de cassation à la 3^{ème} promotion du programme « Entrepreneur(e)s d'intérêt général » (EIG3) animé par Etalab⁷. Ils ont donné lieu à une publication scientifique par les *data scientists* du projet⁸.

La Cour de cassation a ainsi mis au point un outil d'IA de traitement du langage naturel reposant sur deux algorithmes utilisant des procédés d'apprentissage automatique (*machine learning*). Le premier algorithme de reconnaissance des entités nommées (*named entity recognition* ou NER) permet d'identifier les éléments à occulter dans la décision. Pour la construction et la mise à jour de cet algorithme, la Cour de cassation s'est appuyée sur une équipe d'annotateurs, chargée de relire et d'annoter une partie des décisions publiées⁹. Le second modèle d'IA générative a, lui, permis de modéliser le langage des décisions de justice françaises en s'appuyant pour cela sur une base de près de deux millions d'arrêts de la Cour de cassation et de différentes cours d'appel afin d'augmenter les performances du premier algorithme. En 2020 et 2021, une nouvelle participation de la Cour de cassation au programme « Entrepreneurs d'Intérêt Général » (EIG4) a permis de développer une nouvelle interface d'annotation très ergonomique (logiciel LABEL) permettant une meilleure maîtrise du processus complet de pseudonymisation et une amélioration des conditions d'annotation des décisions¹⁰.

⁷ [Voir la présentation du projet sur le site numerique.gouv.fr](#)

⁸ [May I Check Again? — A simple but efficient way to generate and use contextual dictionaries for Named Entity Recognition. Application to French Legal Texts](#)

⁹ Le rôle de cette équipe d'annotateurs, actuellement composée de 20 personnes, a fait l'objet d'une étude sociologique de Mme Camille Girard-Chanudet exposée dans sa thèse : « *La justice algorithmique en chantier. Sociologie du travail et des infrastructures de l'Intelligence Artificielle* ».

¹⁰ [Voir la présentation du projet sur le site numerique.gouv.fr](#)

L'algorithme de pseudonymisation, très performant, a fait l'objet d'améliorations continues. Un modèle plus robuste, dont l'exploitation repose sur un processeur graphique, a été mis en service au cours de l'année 2024. Cet algorithme, comme son prédécesseur, est régulièrement réentraîné pour en améliorer les performances.

Ce nouveau modèle a été enrichi d'un second algorithme d'IA dit de fiabilité, qui permet d'identifier de manière automatique certaines décisions pour lesquelles une relecture humaine sera jugée nécessaire en raison d'un doute sur le résultat de la pseudonymisation automatique.

Section 2 – Mise en place en 2020 d'un outil d'orientation automatique des mémoires ampliatifs : une IA frugale s'appuyant sur des données de qualité

La création en 2004 d'un nouveau système informatique (Nomos) a permis à la Cour de cassation de se montrer pionnière dans le processus de dématérialisation du traitement des affaires. La Cour reçoit, en matière civile, des mémoires ampliatifs transmis par les cabinets d'avocats aux Conseils sous forme de fichiers PDF. Une grande partie de l'activité juridictionnelle repose sur une gestion électronique des documents (GED), accessible depuis un bureau virtuel.

Jusqu'en 2020, pour les matières civile, sociale et commerciale, ces mémoires ampliatifs étaient analysés par des juristes et magistrats du SDER, puis orientés vers les chambres compétentes à l'aide d'une liste de codes matières. Le SDER est en effet composé de bureaux dits « miroirs » des chambres de la Cour, chacun placé sous la direction de magistrats auditeurs, et spécialisés dans les contentieux de leur chambre de rattachement.

Un roulement des bureaux était ainsi organisé afin d'assurer une permanence d'orientation des mémoires ampliatifs qui étaient transmis à la Cour sans pré-tri. À raison de 30 à 50 codes matières différents par chambre, ce travail d'analyse nécessitait par conséquent, pour l'orientateur du SDER, de maîtriser près de 200 codes différents couvrant l'intégralité du spectre contentieux des chambres civiles, commerciale et sociale de la Cour de cassation selon des règles de priorité parfois complexes. C'est dans ce contexte qu'un projet de création d'un algorithme de reconnaissance et de tri des mémoires ampliatifs civils en vue d'assurer leur orientation vers les chambres compétentes, a émergé en 2020.

Cet outil d'analyse et de traitement du langage a été développé en interne par un *data scientist* en collaboration avec les magistrats du SDER sur une période de deux mois, puis intégré au système Nomos au terme d'un an de travaux. La disproportion du temps de conception de l'outil rapporté à celui de son intégration dans un système préexistant nécessite de penser l'infrastructure du système informatique, la structuration des données qu'il renferme et l'ensemble des applicatifs métiers, sous l'angle nouveau de leur capacité à accueillir des modèles d'IA développés en interne.

L'algorithme d'orientation s'appuie, pour son entraînement, d'une part sur le caractère structuré des mémoires ampliatifs qui permet d'identifier des zones d'intérêt dans le texte (zonage), d'autre part sur le stock de mémoires ampliatifs déjà orientées (affaires terminées).

Ce projet a en effet pu s'appuyer sur des données d'excellente qualité et structurées selon des normes précises de rédaction, en l'espèce les mémoires ampliatifs. Cette précision rédactionnelle a également été enrichie de plusieurs dizaines de milliers d'analyses humaines réalisées par les orientateurs sur une période de 15 ans à la date du projet. Ces conditions ont formé un contexte idéal pour l'ajustement (*fine tuning*) et l'apprentissage automatique (*machine learning*) du modèle d'IA.

Une fois en production, l'algorithme a pu également être amélioré de manière régulière en confrontant ses résultats à l'analyse des agents et magistrats orientateurs selon un processus d'apprentissage supervisé.

Avec un taux de succès de plus de 90 %, cet outil a considérablement simplifié et allégé la tâche du SDER en matière d'orientation.

Désormais, chaque bureau miroir de chambre est destinataire d'une liste de mémoires ampliatifs pré-orientés et dont l'analyse reposera essentiellement sur un nombre réduit des codes de la nomenclature des affaires orientées. Le modèle d'IA a ainsi permis d'accélérer grandement le processus d'orientation vers les chambres de la Cour et de gagner en précision dans l'orientation des mémoires ampliatifs vers les sections compétentes au sein des chambres.

Cette expérience est riche d'enseignements. La capacité de la Cour de cassation à conduire à bien ce projet résulte en effet d'une combinaison de facteurs métiers et techniques.

Pour le portage du projet, ces facteurs relèvent ainsi de la capacité du métier à identifier au sein des services les tâches chronophages mettant en œuvre de vastes jeux de données et à cadrer précisément le besoin. Pour la mise en œuvre technique, la possibilité de s'appuyer sur d'importants jeux de données structurées et annotées par des experts métier est apparue décisive.

De même, le temps d'intégration du SIA à un environnement préexistant est un critère important et potentiellement limitant. Il doit en conséquence être chiffré en amont et précisément.

Section 3 – Le projet « Divergences »

Le projet « Divergences » a été conçu dans la dynamique qui s'est attachée à la réussite du projet pseudonymisation. La Cour de cassation a, en effet, remporté en juillet 2019 un appel à projets animé par Etalab qui lui a permis de bénéficier de l'intervention d'une

équipe de chercheurs de l'Institut national de recherche en sciences et technologies du numérique (INRIA) au soutien de ses propres équipes¹¹.

Le projet « Divergences » ambitionnait d'aider à détecter au sein des arrêts de la Cour cassation les éventuelles divergences d'interprétation existant entre chambres ou sections. Cet objectif a nécessité l'exploration de différentes architectures d'IA et méthodologies d'analyse des arrêts et a conduit incidemment au développement de deux algorithmes d'IA permettant, pour l'un, le titrage automatique des arrêts avec un taux de réussite de 75 %, et pour l'autre, le rapprochement automatique de décisions de la Cour de cassation en s'appuyant sur les chaînes de titrage générées avec un taux de réussite de 35 % sur une base de plus de 79 000 arrêts publiés et plus de 850 000 arrêts diffusés. Les travaux effectués avec l'INRIA ont donné lieu à la publication d'un article scientifique¹².

Le code source de l'algorithme a été transmis au laboratoire d'innovation à l'expiration de la mission de l'INRIA et repris en interne par les *data scientists* et magistrats du SDER qui continuent de l'améliorer afin de permettre notamment la rédaction automatique de sommaires selon un système d'IA (SIA) générative.

Sur ce point, les résultats récents ont témoigné de bonnes performances du modèle pour la génération de sommaires à partir de la motivation des arrêts rendus par la Cour de cassation et le rapprochement de ces mêmes arrêts, publiés et diffusés, entre eux. En l'état, le modèle développé représente donc une opportunité d'enrichissement du fond existant de près de 800 000 décisions diffusées par la Cour. Le modèle affiche cependant de meilleures performances sur les arrêts récents et intégrant une motivation enrichie.

Si le projet « Divergences » en lui-même reste ainsi au stade de développement, ses résultats intermédiaires, eux, apparaissent immédiatement mobilisables pour les besoins des magistrats et fonctionnaires de la Cour dans leurs activités de recherche.

L'appréhension des arrêts d'appel puis des décisions de première instance par l'algorithme constitue la suite logique de ce projet qui permettrait ainsi à la Cour de cassation d'affiner sa connaissance de la jurisprudence du fond.

Section 4 – Une expertise reconnue aux niveaux national et international

L'expertise humaine de la Cour de cassation est reconnue au niveau national comme international, ce qui lui permet d'intégrer des projets innovants en partenariat avec des institutions publiques partenaires telles que le *Sorbonne Center for Artificial Intelligence*

¹¹ Voir la présentation du projet sur le site modernisation.gouv.fr

¹² [Complex Labelling and Similarity Prediction in Legal Texts: Automatic Analysis of France's Court of Cassation Rulings](#)

(SCAI), dont le projet PostGenAI@Paris ouvre des perspectives particulièrement prometteuses.

Ce projet, qui rassemble de nombreux partenaires institutionnels et industriels, constitue en effet l'un des neuf cluster IA nationaux visant à créer un pôle d'excellence international dédié à l'IA post-générative. Il s'appuie pour ce faire sur une dotation de 35 millions d'euros sur cinq ans, annoncée par le président de la République en mai 2024.

Des membres du SDER interviennent régulièrement à l'invitation d'acteurs internationaux et en particulier européens afin de présenter les projets d'IA de la Cour de cassation et de participer à des projets de coopération sur ce thème¹³.

¹³ Par exemple : [atelier du Conseil de l'Europe sur le résumé automatique de décisions de justice les 15 et 16 mai 2023 à la Cour suprême de Chypre](#) ; [Hackathon du Conseil de l'Europe des 18-19 novembre 2023 à Paris](#) ; [conférence cyberjustice organisée par l'IERDJ en partenariat avec le Conseil de l'Europe et le laboratoire de cyberjustice de Montréal le 24 novembre 2023 au Conseil de l'Europe](#) ; séminaire entre la Cour suprême de Singapour et la Cour de cassation du 31 janvier 2024 sur « [les nouvelles technologies et l'intelligence artificielle : sujets de justice et objets de justice](#) » ; [16^e session de formation des magistrats francophones de l'AA-HJF le 6 novembre 2024 à Porto-Novo \(Bénin\)](#) ; [Hackathon du Conseil de l'Europe « Digital Future of Justice 2.0 » les 16-17 novembre 2024 à Bologne \(Italie\)](#) ; webinaire de l'Unesco du 21 novembre 2024 : « [AI in the courts: Exploring practical applications and transformative use cases in the judicial system](#) ».

Chapitre 2 – Présentation d'outils et de projets judiciaires basés sur l'IA au niveau national ou international

Section 1 – Au niveau national

Il y a lieu de constater que l'intégration de SIA au service des institutions et administrations françaises est aujourd'hui un sujet de société.

La conception d'une IA publique ou IA générale à finalités de valorisation des données publiques mais également d'accroissement des performances des administrations est un enjeu dont l'importance a été soulignée par le président de la République lors du sommet pour l'action sur l'IA qui s'est tenu les 10 et 11 février 2025.

Cette prise de conscience demeure cependant récente à l'échelle des réflexions qui ont pu être conduites sur l'IA.

L'ambition du rapport Villani, en mars 2018, de « *Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne* »¹⁴ n'a en effet pas connu de traduction immédiate au sein de l'institution judiciaire ni suscité de réelles réflexions sur la place de ces technologies au sein du processus judiciaire.

Il a fallu attendre l'année 2022 pour que les potentialités de l'IA frappent les consciences des acteurs du monde judiciaire au travers du phénomène mondial ChatGPT, révélé en novembre 2022 et dont le lancement suit de peu la publication d'une étude d'ampleur publiée en août 2022 par le Conseil d'Etat¹⁵.

Le rapport remis au président de la République en mars 2024 par la Commission de l'intelligence artificielle complète cette réflexion et formule 25 propositions afin de faire de la France un acteur majeur de la révolution technologique qu'offre l'IA.

Au sein des cours suprêmes et parmi les acteurs juridiques institutionnels du droit, cette ambition prend forme au travers de différents axes d'exploration de l'intelligence artificielle qui supposent en premier lieu de **définir des stratégies internes**, une feuille de route de développement des SIA.

¹⁴ Rapport de la mission parlementaire confiée à Monsieur Cédric Villani le 8 septembre 201 et remis au président de la République le 28 mars 2018.

¹⁵ Cette étude « *Intelligence artificielle et action publique : construire la confiance, servir la performance* », réalisée à la demande du premier ministre, promeut la construction d'outils algorithmiques comme facteur de progrès de la performance publique sous réserve de trois conditions essentielles : la vigilance dans le déploiement de ces outils, la capacité de la France à se doter des ressources et de la gouvernance adaptée à cette ambition et la capacité à créer la confiance publique.

Le Conseil constitutionnel a ainsi adopté un schéma stratégique d'IA le 28 novembre 2024, visant à aborder de manière vigilante l'utilisation potentielle de solutions d'IA. De même, le Conseil d'État a formé un groupe de travail qui a mis en évidence les risques et enjeux éthiques associés à l'utilisation de l'IA, mais également les opportunités d'amélioration qualitative du travail des juges.

Le ministère de la justice, quant à lui, travaille à établir un cadre pour accompagner le développement d'outils algorithmiques dans les juridictions. Cette initiative vise à sensibiliser les agents aux contraintes juridiques de l'utilisation des IA « grand public » dans l'exercice de leurs fonctions, démontrant une préoccupation pour l'encadrement éthique et légal de ces technologies.

Ces feuilles de route ont permis de **définir plusieurs cas d'usage** à développer, parmi lesquels l'analyse du contentieux électoral et l'amélioration de la gestion administrative au Conseil constitutionnel, l'aide à la recherche juridique et l'amélioration de l'accessibilité de la justice administrative aux requérants grâce à des Chatbots pour le Conseil d'État. Ce dernier explore également les potentialités de l'IA en tant qu'outil d'aide à la rédaction pour des tâches précises et supervisées par l'humain ou encore de préanalyse, au stade de l'instruction, de certaines demandes non-contentieuses.

Au sein du ministère de la justice, les récents travaux ont permis de définir des familles de cas d'usage¹⁶ et quatre types d'applications qui pourraient être développés à court terme. Le *speech-to-text*, l'interprétariat, la génération de résumé de documents et l'aide à la recherche font ainsi partie des cas d'usages envisagés par le secrétariat général du ministère de la justice, comme a pu le confirmer Mme Audrey Farrugia, cheffe du service de l'expertise et de la modernisation, lors de son audition.

Au sein des juridictions du fond, la cour d'appel de Paris a récemment lancé une expérimentation, en partenariat avec la DINUM, des outils développés par les services de l'État, tels que l'API Albert, afin d'optimiser le traitement des contentieux de masse et faciliter le travail des magistrats et greffes.

Ces transformations doivent s'appuyer **sur la formation et la sensibilisation des professionnels du droit** et en premier lieu des magistrats, notamment lors de leur formation initiale et continue¹⁷. Ce constat est partagé par le Conseil constitutionnel qui a mis en place une démarche interne de sensibilisation et de formation, mais également par le ministère de la justice et le Conseil d'État, au travers de chartes éthiques en cours d'élaboration.

¹⁶ [Développer l'intelligence artificielle \(IA\) pour la Justice : entre innovation et sécurisation](#), 6 février 2025. Quatre familles de cas d'usage ont été identifiées par le ministère de la Justice : la retranscription automatique ou « *speech to text* », l'interprétariat, le résumé de document ou synthèse documentaire et l'aide à la recherche juridique.

¹⁷ L'École nationale de la magistrature a ainsi enrichi sa formation et propose désormais un [cycle annuel du numérique](#) et une formation spécifique sur les usages de l'IA à partir de 2025.

L'introduction des SIA nécessite également **un suivi attentif de l'évolution de l'état de l'art et des conséquences de ces technologies émergentes**. L'École nationale de la magistrature, en partenariat avec l'université de Grenoble¹⁸, a ainsi pu développer une réflexion sur l'influence de la preuve algorithmique dans le processus décisionnel du magistrat et prévoit de dédier une équipe de recherche à l'analyse de l'évolution des SIA.

Section 2 – Au niveau international

Sur le plan international, plusieurs Etats ont d'ores et déjà recours à l'IA dans le domaine juridictionnel, et de nombreuses initiatives sont en cours de développement. Ces projets relèvent de plusieurs usages, allant de l'aide à l'analyse des dossiers à la rédaction même des décisions. Il apparaît cependant que ces derniers cas d'usage restent minoritaires et que les systèmes déployés s'appliquent davantage aux tâches périphériques à la rédaction de la décision.

Les exemples suivants¹⁹, bien que non exhaustifs, permettent d'illustrer diverses initiatives à travers le monde.

§ 1 – Gestion du flux juridictionnel

La gestion des flux entrants, que ce soit au travers de fonctions de triage automatique, à l'instar de ce qui a pu être mis en place à la Cour de cassation, ou d'une meilleure exploitation des données entrantes, constitue un des champs les plus exploités de l'IA pour l'activité judiciaire.

En Italie, le projet *Aut Dedere, Aut Judicare* exploite de manière automatisée les données issues des productions judiciaires, notamment les mandats d'arrêt, ainsi que les demandes d'extraditions et de transferts, afin de fournir des statistiques en matière de coopération judiciaire pénale.²⁰ A l'instar des juridictions italiennes, le ministère de la justice finlandais a développé *Robot Process Automation*, un outil permettant de relier les paiements relatifs aux sanctions pénales aux jugements correspondants, facilitant leur gestion et le suivi des défauts de paiement.²¹

Des systèmes d'IA sont également mis en place dans plusieurs pays pour filtrer et prioriser les affaires. Au Brésil, le système *VICTOR*, utilisé depuis 2017 par la Cour suprême, identifie les affaires présentant une « répercussion générale » pour mieux filtrer les recours extraordinaires. Un système supplémentaire *ATHOS* permet d'analyser les

¹⁸ Voir notamment l'article de Géraldine Vial, [Prise en main d'un outil d'intelligence artificielle par des auditeurs de justice : l'office du juge sous l'influence des algorithmes – Géraldine Vial – D. 2022. 1928](#)

¹⁹ Voir annexe 2 études du service des relations internationales de la Cour de cassation

²⁰ Centre de recherches et de diffusion juridiques du Conseil d'Etat, Service de recherches juridiques, Cellule de droit comparé, « *Note de droit comparé sur l'utilisation de l'intelligence artificielle au sein des juridictions* », 7 mai 2024

²¹ *Ibid.*

contentieux récurrents afin d'identifier la jurisprudence applicable ou examiner la recevabilité des recours.²²

L'identification et l'aide au traitement de contentieux sériels a également été appréhendée par la cour d'appel de Stuttgart, qui s'est pour cela appuyée sur le logiciel *OLGA* (*OberLandesGerichtAssistent : assistant du tribunal régional supérieur*) pour le traitement des affaires liées au *Dieselgate*, ce afin d'identifier et grouper des affaires semblables. Cet outil combine traitement du langage naturel et apprentissage automatique pour numériser les jugements et mémoires de première instance, facilitant ainsi l'affectation des dossiers en appel selon des critères thématiques.²³ Il s'agit cependant d'une initiative locale et non nationale.

L'implication de l'IA au stade de la répartition et de l'orientation des dossiers peut également avoir pour objet d'assurer le respect des droits fondamentaux et de garantir notamment l'impartialité du tribunal. Ainsi, en Pologne, un système d'attribution aléatoire des dossiers a été mis en place afin de prévenir toute ingérence dans la répartition des contentieux auprès des juges, notamment afin d'éloigner des risques de conflit d'intérêt ou de tentative de corruption²⁴.

§ 2 – Retranscription et interprétariat

Du fait des capacités de l'IA à appréhender de nombreux langages, qui correspond à une des caractéristiques d'entraînement de la plupart des LLM (multilinguistiques), la traduction et l'interprétariat constituent également des cas d'usage évidents de l'IA pour l'activité juridictionnelle. La transcription figure également en bonne place eu égard aux nombreux modèles privés et *open source* appliqués à ce cas d'usage.

Ainsi, en Estonie, le système *Salme* enregistre les audiences et intègre un logiciel de reconnaissance vocale permettant d'automatiser la rédaction des procès-verbaux. Des outils similaires sont en cours de développement en Croatie et en Estonie pour automatiser la prise de notes durant les audiences.

La traduction constitue un autre champ d'application majeur. En Chine, la cour primaire de Hunchun a mis en place le modèle *i-法院 (I-Court)*, qui permet la transcription et la traduction en temps réel des interventions des parties ainsi que des documents soumis par les avocats. Ce système est particulièrement utile dans cette ville frontalière de la Russie et de la Corée du Nord, où les affaires multilingues sont fréquentes. En Inde, le système *SUVAS (Supreme Court Vidhik Anuvaad Software)* traduit automatiquement les documents judiciaires entre l'anglais et les langues locales, facilitant ainsi l'accès à la

²² Service des relations internationales de la Cour de cassation, « *Utilisation de l'intelligence artificielle au sein de la Justice Brésilienne* », 2024

²³ Centre de recherches et de diffusion juridiques du Conseil d'Etat, Service de recherches juridiques, Cellule de droit comparé, « Note de droit comparé sur l'utilisation de l'intelligence artificielle au sein des juridictions », 7 mai 2024 préc.

²⁴ Commission européenne pour l'efficacité de la justice (CEPEJ), *Centre de ressources sur la cyberjustice et l'IA*, 2024

justice dans un pays où l'anglais n'est véritablement parlé que par une minorité de la population.²⁵

§ 3 – Accès à la justice et information du justiciable

En Suède, l'Agence nationale des tribunaux (*Domstolsverket*) a lancé un projet de *chatbot* destiné à répondre aux questions des justiciables et à permettre le suivi de l'état d'avancement de leurs dossiers sans solliciter directement les greffes des tribunaux²⁶.

La Suède développe également un outil de traduction automatique des arrêts de la Cour suprême afin de faciliter leur accès dans d'autres langues, notamment l'anglais. Celui-ci vise à faciliter l'accès aux décisions judiciaires suédoises à l'international, tout en participant aux efforts de convergence et d'harmonisation du droit européen.²⁷

§ 4 – Analyse de dossiers et enrichissement des données de procédure

L'IA est également mise au service des juges afin d'améliorer la lisibilité de la procédure et de faciliter l'analyse des dossiers.

Il apparaît ainsi qu'en Allemagne, des juridictions de Rhénanie-du-Nord-Westphalie et de Basse-Saxe utilisent des systèmes capables de générer une chronologie des litiges et de lier automatiquement les textes aux dispositions juridiques pertinentes. La Cour suprême autrichienne a également mis en place un outil permettant d'afficher toutes les citations contenues au sein d'un texte et d'établir des liens vers leurs références respectives.²⁸

Plusieurs juridictions ont développé des systèmes semblables en Chine. La Haute Cour de Pékin a conçu *Ruì fǎguān* (*Juge Ruì*), un programme capable de fournir précisément la jurisprudence et les textes juridiques applicables à un litige. La Haute Cour de Shanghai a, quant à elle, mis en place le système 206, conçu pour détecter les vices de procédure, analyser les preuves fournies à chaque étape d'un procès et signaler d'éventuelles incohérences.

§ 5 – Aide à la rédaction des décisions

Certains Etats expérimentent l'utilisation de l'IA pour la rédaction – partielle ou totale – des décisions judiciaires.

²⁵ Service des relations internationales de la Cour de cassation, « L'utilisation de l'Intelligence Artificielle et des nouvelles technologies par le système judiciaire en Inde », 2024. En Annexe

²⁶ Commission européenne pour l'efficacité de la justice (CEPEJ), *Centre de ressources sur la cyberjustice et l'IA*, 2024

²⁷ *Ibid.*

²⁸ Centre de recherches et de diffusion juridiques du Conseil d'Etat, Service de recherches juridiques, Cellule de droit comparé, « Note de droit comparé sur l'utilisation de l'intelligence artificielle au sein des juridictions », 7 mai 2024

En Basse-Saxe, le système *FT Cam* est utilisé depuis plusieurs années pour pré-rédiger certaines décisions. Celui-ci aide notamment à la rédaction d'ordonnances de première instance en matière familiale et au calcul automatique des montants de pensions alimentaires. Le système *FRAUKE* (*FRAnkfurts Urteil Konfigurator Elektronisch*) permet de traiter plus rapidement les contentieux de masse liés aux plaintes de passagers aériens ayant trait aux retards ou annulations de vols. Celui-ci permet de générer des propositions de rédaction de jugements, ainsi que des références jurisprudentielles, afin de traiter plus efficacement des litiges substantiellement similaires.

En Estonie, l'IA est utilisée pour générer des décisions judiciaires dans les procédures accélérées en commandement de payer, bien que cette utilisation soit limitée aux affaires portant sur des montants inférieurs à 8 000 euros. Plusieurs *chatbots* ont également été développés en ce sens à Singapour afin d'accompagner les justiciables dans certaines procédures. *Divorce AIDE*, par exemple, aide à estimer la répartition des biens lors d'un divorce. Un autre *chatbot*, fondé sur l'IA générative, est actuellement testé au tribunal des petites créances afin de fournir des conseils juridiques simplifiés aux citoyens.²⁹

Enfin, en Argentine, le parquet de Buenos Aires utilise *Prometea*, un système d'IA entraîné sur 2 400 jugements et 1 400 avis juridiques, permettant d'accélérer le traitement des affaires en proposant des décisions automatisées dans des litiges sociaux, notamment en matière d'attribution de logements ou d'aides publiques. Cet outil a été adopté en Colombie sous le nom de *PretorIA*, où il serait capable d'analyser en cinq secondes 2 700 décisions de *tutela* soumises à la Cour constitutionnelle, avec un taux de succès de 95% sur 13 critères d'examen. Outre la génération de décisions, certains outils visent à uniformiser et améliorer la qualité des documents juridiques. *Prometea* joue également un rôle de correcteur automatique, permettant d'éliminer 99 %³⁰ des erreurs typographiques dans les décisions judiciaires.³¹

§ 6– Rédaction de chartes éthiques et d'encadrement des usages de l'IA

En complément des différents SIA déployés à titre expérimental ou mis en production, de nombreux Etats ont entendu encadrer l'usage de l'IA dans le domaine juridictionnel en mettant en place des chartes éthiques ou d'encadrement de ses usages.

Si les pays de *Common Law* ont globalement opté pour une approche de régulation souple et non contraignante de l'IA via des chartes éthiques, certains pays de l'Union

²⁹ Service des relations internationales de la Cour de cassation, « L'utilisation de l'intelligence artificielle dans le système judiciaire en Argentine », 2024

³⁰ Ce taux de succès repose sur l'analyse du directeur du laboratoire de l'innovation de la faculté de droit de Buenos Aires. [Entrevista a Juan Gustavo Corvalán: "Prometea es el primer sistema de inteligencia artificial diseñado y desarrollado por el Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires" - The Technolawgist](#) Aucune publication scientifique attestant de ces performances n'a cependant été identifiée.

³¹ Service des relations internationales de la Cour de cassation, « L'Utilisation de l'Intelligence Artificielle et des nouvelles technologies par le système judiciaire en Colombie », 2024

Européenne ont en revanche entendu se doter de textes contraignants à valeur législative pour encadrer l'usage de l'IA dans la justice.

Si ces chartes éthiques prennent des formes variées et émanent de multiples institutions, l'esprit qui les anime et les principes déclinés sont relativement similaires et correspondent aux grands principes de prudence et de maîtrise de la technologie.

L'*AI Guidance for Judicial Office Holders* énumère ainsi 7 principes clés repris en tout ou partie par la Charte de la Cour suprême du Delaware sur l'utilisation par le personnel de l'IA générative³² :

1. *Comprendre l'IA et ses applications*
2. *Respecter la confidentialité et la vie privée*
3. *Garantir la responsabilité et l'exactitude*
4. *Être conscient des biais*
5. *Assurer la sécurité*
6. *Assumer ses responsabilités*
7. *Être conscient que les utilisateurs des cours et tribunaux peuvent utiliser des outils d'IA*

Les chartes adoptées aux États-Unis, Royaume-Uni et Australie³³ mettent ainsi l'accent sur la responsabilité des utilisateurs et rappellent ces derniers à leurs obligations de confidentialité, de prudence et de vérification des contenus générés. Elles prévoient également certaines restrictions à l'usage de l'IA, comme l'interdiction de déléguer la prise de décision à l'IA ou son utilisation pour certains documents juridiques spécifiques.

³² [La Charte de la Cour suprême du Delaware sur l'utilisation par le personnel de l'IA générative](#) prévoit cinq principes clés applicables aux fonctionnaires pour l'utilisation de l'IA :

1. L'utilisateur autorisé reste responsable. Toute utilisation des résultats du GenAI est en fin de compte la responsabilité de l'utilisateur autorisé. Les utilisateurs autorisés sont responsables de l'exactitude de tous les produits de travail et doivent faire preuve de prudence lorsqu'ils s'appuient sur les résultats de GenAI.
2. Utilisation informée. Les utilisateurs autorisés ne devraient pas utiliser le GenAI approuvé sans avoir une connaissance et une compréhension pratiques des outils. Les utilisateurs autorisés devraient être formés aux capacités techniques et aux limites des GenAI approuvés avant de les utiliser.
3. Prise de décision. Les utilisateurs autorisés ne peuvent pas déléguer leur fonction de prise de décision à l'IA approuvée.
4. Conformité avec les lois et les politiques du pouvoir judiciaire. L'utilisation des GenAI doit être conforme à toutes les lois applicables et aux politiques du pouvoir judiciaire.
5. IA générative non approuvée. Les utilisateurs autorisés ne peuvent pas introduire d'informations non publiques dans des GenAI non approuvés. Les GenAI non approuvés ne peuvent pas être utilisés sur les ressources technologiques de l'État.

³³ Voir notamment le [Guide de bonnes pratiques sur l'IA générative de la Cour suprême de Nouvelle-Galles du Sud applicable aux procédures à compter du 3 février 2025](#)

Au sein de l'Union Européenne, l'Italie, l'Espagne et l'Allemagne se sont en revanche dotés de supports législatifs nationaux pour encadrer l'usage de l'IA dans la justice.

En Italie, un projet de loi non définitif a ainsi été diffusé au printemps 2024³⁴ par le gouvernement. Il ambitionne ainsi de restreindre l'utilisation de l'IA à des fins de support, telles que l'organisation et la simplification du travail judiciaire ou encore la recherche jurisprudentielle et doctrinale. Le projet de loi précise par ailleurs que la décision sur l'interprétation de la loi, l'appréciation des faits et des preuves, ainsi que l'adoption de toute mesure restent exclusivement du ressort du magistrat.

En Espagne, la *Carta de Derechos Digitales* (Charte des droits numériques), élaboré par le ministère des Affaires Économiques et de la Transformation Numérique d'Espagne et approuvée par le gouvernement espagnol le 14 juillet 2021, édicte un ensemble de droits pour protéger les intérêts et les libertés des citoyens dans l'environnement numérique.

Elle s'articule avec le décret-loi royal du 19 décembre 2023 portant approbation de mesures urgentes pour l'exécution du plan de redressement, de transformation et de résilience, dont le préambule prévoit que l'usage de l'IA dans le « *domaine de l'administration de la justice* » aura un « *caractère instrumental de support et de soutien de l'activité juridictionnelle, dans le plein respect des garanties procédurales et constitutionnelles* », sous le contrôle et au service de l'humain.

Le décret-loi poursuit en identifiant aux articles 56 et 57, au sein du service public de la justice, trois catégories d'activités ou d'actions pour lesquelles est prévu un usage plus ou moins poussé des outils d'IA'. Sont ainsi distinguées les actions automatisées³⁵, les actions dites « *proactives* »³⁶ et les actions assistées³⁷.

³⁴ [Communiqué de presse du Conseil des ministres n° 78, 23 avril 2024](#)

³⁵ Ainsi les actions automatisées pourront concerner « *les actes simples de procédure ou de décisions simples qui ne nécessitent pas d'interprétation juridique. Il s'agit, entre autres, de : a) La numérotation ou la pagination des dossiers ; b) La transmission des dossiers aux archives lorsque les conditions procédurales sont réunies ; c) La production de copies et de certificats ; d) La création de livres ; e) La vérification des représentants* » ; f) *La déclaration du caractère définitif d'une décision (...)* ». »

³⁶ Aux termes de l'alinéa 3 de l'article 56 les actions proactives sont « *les actions automatisées, auto-initiées par les systèmes d'information sans intervention humaine, qui exploitent, dans un but spécifique, des informations incorporées dans un dossier ou dans une procédure d'une administration publique, pour générer des avertissements ou des effets directs (...), dans le même dossier ou dans d'autres dossiers, de la même administration publique ou d'une autre administration publique, en se conformant, dans tous les cas, à la loi* ».

³⁷ Les actions assistées : *Conformément aux dispositions de l'article 57, « une action assistée est une action pour laquelle le système d'information de l'administration de la justice génère un projet total ou partiel de document complexe fondé sur des données, qui peut être produit par des algorithmes, et qui peut constituer la base ou le support d'une décision judiciaire ou procédurale. 2. En aucun cas, le projet de document ainsi généré ne constitue en lui-même une décision judiciaire ou procédurale, sans validation par l'autorité compétente.*

En Allemagne, la Loi fondamentale prévoit que seules les personnes physiques peuvent exercer le pouvoir judiciaire. L'IA ne peut être utilisée que pour assister les juges, sans jamais se substituer à eux.

PARTIE 2

PRESENTATION DES CRITERES D'EXAMEN

Afin d'évaluer la pertinence des cas d'usage identifiés par la mission IA, le groupe de travail a retenu deux séries de critères d'évaluation fondées sur des approches fonctionnelles, éthiques et juridiques, d'une part, et technique et économique, d'autre part.

Chapitre 1 – Présentation des critères d'évaluation fonctionnels, éthiques et juridiques

Section 1 – Critères éthiques

§ 1 – Conciliation du système d'IA étudié avec les droits fondamentaux

L'intégration de systèmes d'IA au sein de la Cour de cassation ne peut se faire sans prise en compte de leur impact sur les droits fondamentaux.

L'intensité de cet impact devra s'apprécier en fonction de son influence sur l'acte de juger lui-même et de sa proximité avec l'acte décisionnel susceptible de produire des effets sur les utilisateurs du service public, les justiciables. Par exemple, une assistance à la recherche documentaire ne saurait ainsi poser les mêmes questions éthiques qu'une assistance à la rédaction et *a fortiori* qu'une aide à la décision.

Cette logique intègre notamment l'économie des dispositions du règlement IA et procède également de la Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires³⁸, adoptée par la CEPEJ, en décembre 2018. Le groupe de travail relève que cette charte demeure un instrument de référence d'actualité, malgré les évolutions techniques intervenues depuis son adoption.

La Charte énumère ainsi cinq principes repris dans la Convention-cadre du Conseil de l'Europe :

- principe de respect des droits fondamentaux : assurer une conception et une mise en œuvre des outils et des services d'IA qui soient compatibles avec les droits fondamentaux, notamment la Convention de sauvegarde des droits de l'Homme et des libertés fondamentales³⁹ ;

³⁸ [Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires - Commission européenne pour l'efficacité de la justice \(CEPEJ\)](#)

³⁹ Il s'agit ici d'étudier l'interférence avec des droits fondamentaux spécifiquement protégés par des instruments juridiques tels que la convention européenne de sauvegarde des droits de l'homme, la charte des droits fondamentaux de l'Union européenne ou encore la Constitution. L'Institut des études et de la recherche sur le droit et la justice (IERDJ) identifie ainsi dans son [étude d'impact du numérique sur la Justice, volume 1, septembre 2024](#), des risques manifestement d'atteinte aux droits fondamentaux et notamment à l'article 6 de la Convention de sauvegarde des droits de l'Homme et des libertés fondamentales mais également des usages potentiels de l'IA permettant de contribuer à renforcer l'accès à la justice et à la connaissance ou encore d'accentuer la transparence des motivations des décisions de justice.

- principe de non-discrimination : prévenir spécifiquement la création ou le renforcement de discriminations entre individus ou groupes d'individus ;

- principe de qualité et sécurité : en ce qui concerne le traitement des décisions juridictionnelles et des données judiciaires, utiliser des sources certifiées et des données intangibles avec des modèles conçus d'une manière multidisciplinaire, dans un environnement technologique sécurisé ;

- principe de transparence, neutralité et intégrité intellectuelle : rendre accessibles et compréhensibles les méthodologies de traitement des données, autorisant les audits externes⁴⁰ ;

- principe de maîtrise par l'utilisateur : bannir une approche prescriptive et permettre à l'usager d'être un acteur éclairé et maître de ses choix.

Ces principes sont au cœur de la notion même de transparence ou d'explicabilité des résultats algorithmiques⁴¹, définis à l'article 8 de la convention cadre du 5 septembre 2024 et explicités dans son rapport explicatif⁴².

Cette même convention souligne également la nécessité de tenir compte du risque de biais volontaire ou involontaire des SIA pouvant être à l'origine de discriminations en raison de la nature des données utilisées, notamment, ou encore au regard de la plus grande confiance que les humains pourraient accorder « *aux machines et aux artefacts technologiques ou dans des situations où ils sélectionnent des informations qui confirment leurs propres opinions* »⁴³. Ce danger inhérent au biais de confirmation a par ailleurs été souligné par le professeur Etienne Vergès lors de son audition par la mission IA.

La maîtrise humaine de la décision est aussi apparue au groupe de travail comme un principe essentiel pour préserver intact l'office du juge. Sa mise en œuvre impose notamment de toujours conserver une intervention humaine dans les différentes étapes du

⁴⁰ Pour une approche comparée des principes de transparence dans les différentes chartes voir notamment les ouvrages de Yannick Meneceur : « *L'intelligence artificielle en procès* », Bruylant, 2019 et « *IA générative et professionnels du droit* », Lexis Nexis, 2024.

⁴¹ Sur la notion de transparence, voir notamment l'article de M. Clément Henin et Daniel Le Métayer, doctorant et directeur de recherche au sein de l'équipe Inria Privatics : « [Fournir des explications du fonctionnement des algorithmes compréhensibles par des profanes](#) », 6 janvier 2021

Par ailleurs, selon Yannick Meneceur, le principe de transparence ne saurait se restreindre à la communication du seul code informatique. Sont aussi essentielles la documentation de conception, une explication des arbitrages réalisés par les concepteurs dans le langage commun et la traçabilité de l'exécution du code informatique. Ainsi le principe de transparence renverrait de manière générale à l'étude d'impact sur les droits de l'Homme proposée par le commissaire aux droits de l'homme du Conseil de l'Europe qui a pour objet d' « *évaluer les éventuels effets des système d'IA sur les droits de l'homme en tenant compte de la nature, du contexte, de la portée et de l'objectif du système* » (« *Décoder l'intelligence artificielle : 10 mesures pour protéger les droits de l'Homme, mai 2019* », p. 205-206)

⁴² [Rapport explicatif de la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit](#)

⁴³ Commentaire de l'article 10 au sein du rapport explicatif de la Convention-Cadre précité

processus décisionnel⁴⁴, depuis l'analyse du dossier jusqu'à la rédaction de la décision, le recours à l'IA n'intervenant que pour apporter une aide ponctuelle et toujours à la demande et sous contrôle du juge.

Les lignes directrices en cours d'élaboration par l'UNESCO pour l'utilisation dans les systèmes d'IA dans les cours et tribunaux constituent également une source de référence pour un déploiement éthique de l'IA au sein des systèmes de justice⁴⁵.

Par ailleurs, le Conseil de l'Europe a développé une approche structurée pour réaliser des évaluations des risques et de l'impact des systèmes d'IA⁴⁶. Cette méthodologie HUDERIA est spécialement conçue pour protéger et promouvoir les droits humains, la démocratie et l'État de droit.

Il importe de prendre en compte l'ensemble de ces facteurs de risque et d'évaluer les mesures à prendre pour atténuer les risques et impacts négatifs. La confrontation des cas d'usages retenus à ces critères éthiques pourrait être déclinée en s'appuyant sur les questions suivantes : le SIA envisagé est-il susceptible d'affecter les droits fondamentaux des justiciables ? La transparence et l'explicabilité des algorithmes sont-elles assurées ? Des mécanismes de contrôle et d'évaluation doivent-ils être mis en place pour garantir le respect des droits fondamentaux ?

Note	Signification
0	Le SIA présente une interaction très forte avec les droits fondamentaux
1	Le SIA présente une interaction significative avec les droits fondamentaux
2	Le SIA présente une interaction faible avec les droits fondamentaux
3	Le SIA ne présente aucune interaction avec les droits fondamentaux

§ 2 – Développement durable : frugalité des systèmes d'intelligence artificielle

L'appréciation de l'impact écologique du SIA à développer pour répondre au cas d'usage examiné doit tenir compte de l'impact environnemental de la conception, du

⁴⁴ Voir IA générative et décision du juge, E. Legrand et M. Popa-Fabre, impact du numérique sur la justice, vo 2, oct. 2024, IERDJ

⁴⁵ [Document pour consultation publique: Lignes directrices de l'UNESCO pour l'utilisation des systèmes d'IA dans les cours et tribunaux - UNESCO Bibliothèque Numérique](#)

⁴⁶ Voir l'outil HUDERIA d'analyse de l'impact des systèmes d'IA sur les droits fondamentaux : [HUDERIA : Nouvel outil pour évaluer l'impact des systèmes d'IA sur les droits humains - Portal](#)

développement et de l'utilisation de ce SIA, en termes de consommation d'énergie et d'utilisation des ressources⁴⁷.

L'analyse de ce critère impliquera donc d'examiner notamment l'efficacité énergétique des algorithmes utilisés, la possibilité ou non de s'appuyer sur une IA frugale, la quantité de données nécessaires à l'entraînement des modèles algorithmiques, ainsi que le volume et la fréquence des données traitées en production⁴⁸. Il importera donc de mettre en balance les bénéfices attendus de l'IA et les coûts environnementaux induits⁴⁹.

La possible utilisation d'IA frugale pour atteindre les objectifs définis par les cas d'usage constituera ainsi un indicateur positif de soutenabilité environnementale du projet⁵⁰. Certains cas d'usage nécessiteront toutefois de s'appuyer sur des modèles à l'état de l'art et plus énergivores, ce qui induira un coût énergétique et environnemental plus important.

Ce critère peut donc être décliné, à l'état de l'art, en répondant aux questions suivantes : le cas d'usage identifié nécessite-t-il de recourir à un LLM ou une technologie d'IA énergivore ? Quelle sera la quantité de données utilisées pour l'entraînement ? Sur quelle quantité de données le modèle a-t-il vocation à s'appuyer en production ? A quel volume de requêtes/d'utilisateurs le modèle devra-t-il faire face ?

Note	Signification
0	Le SIA repose sur des algorithmes extrêmement gourmands en ressources et/ou une quantité de données d'entraînement excessive
1	Le SIA utilise des algorithmes relativement gourmands en ressources et/ou une quantité de données d'entraînement importante
2	Le SIA peut reposer sur des modèles relativement frugaux et une quantité de données raisonnable
3	Le SIA fonctionne de manière optimale en s'appuyant sur une IA frugale et requiert une quantité raisonnable de données d'entraînement.

⁴⁷ Des réflexions ont déjà été engagées par les pouvoirs publics sur ce sujet au travers notamment du [référentiel général pour l'IA frugale publié en juin 2024](#) par l'AFNOR et le ministère de la transition écologique

⁴⁸ Voir The Shift Project, *Synthèse de l'audition du 12 octobre 2017 par la Mission Villani – IA, « Intelligence Artificielle, Numérique et Environnement »* 5 novembre 2017 (Le Shift auditionné par la Mission Villani sur l'Intelligence artificielle – The Shift Project).

⁴⁹ V. Étude du Conseil d'État, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, adoptée en assemblée générale plénière le 31 mars 2022, p. 127.

Section 2 – Critères juridiques

§ 1 – Mise en conformité des SIA à la réglementation IA

Ce critère tient compte du poids des obligations qui pèseront sur la Cour de cassation en tant que fournisseur d'IA⁵¹.

Le système de classification par risque du règlement (UE) 2024/1689 du 13 juin 2024 module en effet la rigueur et l'étendue de ces obligations en fonction de la catégorie à laquelle appartient le SIA étudié⁵². Le règlement IA distingue en effet trois grandes catégories de SIA : les IA prohibées, les IA à haut risque et les IA à risque limité⁵³. A défaut d'intégrer l'une de ces trois catégories, les SIA développés intègrent une catégorie résiduelle pouvant se soumettre volontairement aux obligations prévues par le règlement.

L'économie générale du règlement IA peut ainsi être résumée schématiquement de la manière suivante :

⁵¹ Lors de son audition du 19 décembre 2024, M. Samir Merabet a indiqué que le règlement sur l'intelligence artificielle permet d'avoir une bonne visibilité sur l'état du droit actuel et à venir en matière de développement de systèmes d'IA. Il souligne toutefois que le texte du RIA est très général et n'est pas pensé pour l'implémentation de SIA dans le domaine de la justice.

⁵² V. Y. Meneceur, *Intelligence artificielle – Garantir les droits humains, la démocratie et l'Etat de droit à l'ère de l'intelligence artificielle – Les ambitions croisées du règlement sur l'« IA » de l'UE et de la convention-cadre du Conseil de l'Europe*, La Semaine juridique – Edition Générale, n°6, 10 février 2025, doct. 201.

⁵³ Les SIA, à risques limités, soumis à une obligation de transparence sont énumérés par l'article 50 du règlement IA. Cette obligation de transparence a principalement pour objet d'informer le public utilisateur direct ou indirect du SIA du recours à l'IA dans la génération de contenus ou le traitement de données.

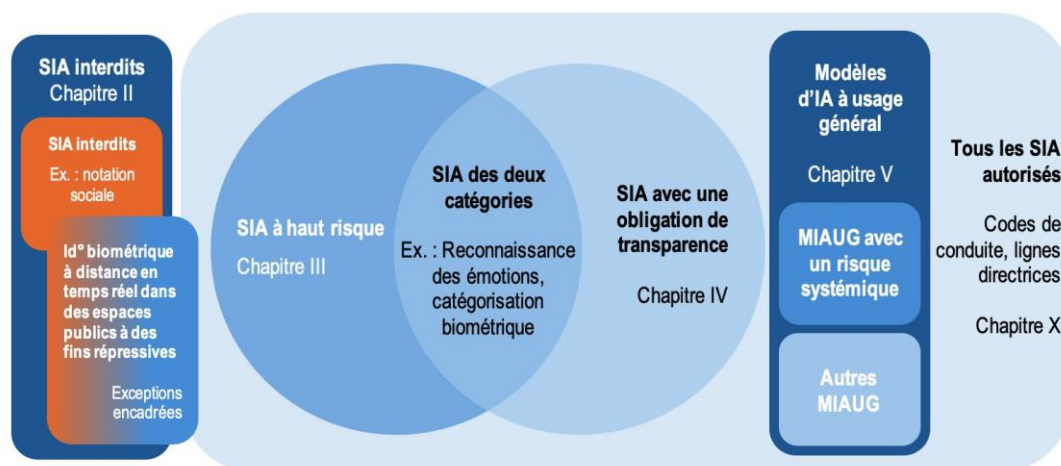


Fig.1 – Catégorisation des systèmes d'IA dans le règlement européen (RIA)

Source : Adapté et étendu à partir d'une publication sous licence *creative commons*
(A. Tiulkanov, EU AI Act: Getting the Basics Straight, LinkedIn, 2024)

54

Par principe, les SIA ayant vocation à être utilisés pour l'administration de la justice appartiennent à la catégorie des IA à haut risque selon les termes du considérant 61 du règlement IA⁵⁵ et de son article 6 § 2 (par renvoi à l'annexe III).

Aux termes de l'annexe III, §8 - a, sont réputés SIA à haut risque, « les systèmes d'IA destinés à être utilisés par les autorités judiciaires ou en leur nom, pour les aider à rechercher et à interpréter les faits ou la loi, et à appliquer la loi à un ensemble concret de faits, ou à être utilisés de manière similaire lors du règlement extrajudiciaire d'un litige. »

Toutefois, et par exception, l'article 6 § 3 précise que « par dérogation au paragraphe 2, un système d'IA visé à l'annexe III n'est pas considéré comme étant à haut risque lorsqu'il ne présente pas de risque important de préjudice pour la santé, la sécurité ou les droits fondamentaux des personnes physiques, y compris en n'ayant pas d'incidence significative sur le résultat de la prise de décision. »

⁵⁴ Schéma extrait de l'étude de Yannick Meneceur magistrat, maître de conférences associé à l'université de Strasbourg et responsable d'études et de recherches associé à l'IERDJ – « Garantir les droits humains, la démocratie et l'État de droit à l'ère de l'intelligence artificielle - Les ambitions croisées du règlement sur l'« IA » de l'UE et de la convention-cadre du Conseil de l'Europe »- La Semaine Juridique Edition Générale n° 6, 10 février 2025

⁵⁵ Considérant 61 : « Certains systèmes d'IA destinés à être utilisés pour l'administration de la justice et les processus démocratiques devraient être classés comme étant à haut risque, compte tenu de leur incidence potentiellement significative sur la démocratie, l'état de droit, les libertés individuelles ainsi que le droit à un recours effectif et à accéder à un tribunal impartial. **En particulier, pour faire face aux risques de biais, d'erreurs et d'opacité, il convient de classer comme étant à haut risque les systèmes d'IA destinés à être utilisés par une autorité judiciaire ou pour le compte de celle-ci pour aider les autorités judiciaires à rechercher et à interpréter les faits et la loi, et à appliquer la loi à un ensemble concret de faits.** »

Tel est le cas lorsque « *Le système d'IA est destiné à accomplir une tâche procédurale étroite* » ou est « *destiné à améliorer le résultat d'une activité humaine préalablement réalisée* », « *à détecter les constantes en matière de prise de décision ou les écarts par rapport aux constantes habituelles antérieures et n'est pas destiné à se substituer à l'évaluation humaine préalablement réalisée, ni à influencer celle-ci, sans examen humain approprié* » ou encore à « *exécuter une tâche préparatoire en vue d'une évaluation pertinente aux fins des cas d'utilisation visés à l'annexe III.* »

A titre d'exemple, le considérant 61 du préambule donne une illustration concrète des dérogations visées à l'article 6 § 3 en précisant que « *la classification des systèmes d'IA comme étant à haut risque ne devrait cependant pas s'étendre aux systèmes d'IA destinés à être utilisés pour des activités administratives purement accessoires qui n'ont aucune incidence sur l'administration réelle de la justice dans des cas individuels, telles que l'anonymisation ou la pseudonymisation de décisions judiciaires, de documents ou de données, la communication entre membres du personnel ou les tâches administratives.* »

Ainsi, les SIA ne présentant pas de risque d'atteinte aux droits fondamentaux des personnes physiques, sans incidence significative sur le résultat de la prise de décision ou destinés par exemple à l'accomplissement de tâches procédurales étroites pourront être soumis à des obligations de transparence et de documentation restreintes⁵⁶.

A l'inverse, l'appartenance des SIA envisagés à la catégorie des IA à haut risque supposera que la Cour de cassation, en qualité de fournisseur d'IA, justifie du respect d'obligations de conformité et de gestion des risques, d'information et de transparence à l'égard des usagers ou encore d'enregistrement auprès des autorités nationales compétentes⁵⁷.

Il semblerait donc, suivant une première analyse, que des outils de jurimétrie ou d'aide à la rédaction par de l'IA générative par exemple relèveraient des SIA à haut risque. A l'inverse, d'autres types d'outils, comme la transcription de voix vers du texte, la traduction et l'interprétation, la pseudonymisation de textes, le résumé de texte, hors pièces de procédure, ou l'orientation automatisée de courriers électroniques ne paraissent pas relever des dispositions les plus contraignantes du règlement IA. Les moteurs de recherche juridique paraissent aussi à exclure du champ des prévisions, dès lors qu'ils se limitent à présenter la liste de divers textes correspondant à des critères abstraits. Ces analyses provisoires devront être revues en tout état de cause à la lumière des lignes

⁵⁶ Voir Commission européenne pour l'efficacité de la justice (CEPEJ), *Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement*, adoptée lors de la 31e réunion plénière de la CEPEJ (Strasbourg, 3-4 décembre 2018), Conseil de l'Europe, février 2019, p. 26-27.

⁵⁷ V. Étude du Conseil d'État à la demande du Premier ministre, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, adoptée en assemblée générale plénière le 31 mars 2022, p. 200.

directrices qui seront publiées par la commission européenne au plus tard le 2 février 2026⁵⁸.

Cette analyse conduit d’ores-et-déjà à constater que le périmètre des IA à haut risque pourrait s’avérer assez limité compte tenu des exceptions prévues par le règlement.

Selon que les systèmes d’IA entrent ou non dans la catégorie des IA à haut risque, le poids des obligations à respecter et du formalisme auquel devra se soumettre la Cour n’est pas neutre et pèsera sur les ressources humaines mobilisées pour le développement, la mise en production comme le suivi des SIA⁵⁹.

Ce critère peut donc être décliné en s’appuyant sur les questions suivantes : le SIA envisagé est-il susceptible d’être qualifié d’IA à haut risque ? D’IA à usage général ? A risque minime ?

Note	Signification
0	Le SIA entre dans la catégorie des exceptions prévues aux IA prohibées
1	Le SIA entre dans la catégorie des IA à haut risque et pourrait avoir une influence majeure sur le processus décisionnel
2	Le SIA entre dans la catégorie des IA à haut risque
3	Le SIA entre dans la catégorie des IA à risque minime hors champs du règlement IA

§ 2 – Mise en conformité avec la réglementation sur les données personnelles

En ce qui concerne le RGPD, le groupe de travail rappelle qu’il est applicable à tout traitement de données à caractère personnel. Les données issues d’une procédure judiciaire sont en principe des données personnelles, sauf à avoir été anonymisées au sens du RGPD, ce qui suppose qu’elles ne se rapportent plus à une personne identifiée ou identifiable. A cet égard, il est important de souligner que le simple fait d’occulter les éléments les plus directement identifiants (noms, prénoms, adresses...) d’un document

⁵⁸ Article 6§5 du règlement IA, après consultation du Comité européen de l’intelligence artificielle (ci-après dénommé «Comité IA»), et au plus tard le 2 février 2026, la Commission fournit des lignes directrices précisant la mise en œuvre pratique du présent article, conformément à l’article 96, assorties d’une liste exhaustive d’exemples pratiques de cas d’utilisation de systèmes d’IA qui sont à haut risque et de cas d’utilisation qui ne le sont pas.

⁵⁹ V. Étude du Conseil d’État, préc., p. 186.

ne constitue pas une anonymisation de ce document mais une simple pseudonymisation qui n'a pas pour effet de rendre inapplicable le RGPD. C'est la raison pour laquelle, comme évoqué supra, un soin particulier doit être apporté à ne transmettre aucune information contenant des données personnelles à un tiers sans s'assurer de la conformité du traitement de données par ce tiers, ce qui peut supposer notamment l'établissement d'une convention de sous-traitance au sens du RGPD et un haut niveau de sécurité des systèmes d'information.

De manière générale, l'utilisation de données personnelles, à l'exception des données contenues dans les décisions mises en *open data* et dont l'utilisation est autorisée par les dispositions de la loi n° 78-17 du 6 janvier 1978 (loi informatique et libertés), aura un impact fort sur la complexité de conformité associée à un projet d'IA, en particulier en ce qui concerne l'hébergement des données.

Ce critère évalue l'impact de la mise en place des SIA étudiés sur les obligations relatives à la protection des données personnelles qui incomberont à la Cour de cassation, en tant que responsable de traitement et potentiellement fournisseur d'IA. Il ne s'agit pas ici d'aborder la question de la conformité de ces SIA à la réglementation applicable en matière de protection des données, qui constitue un prérequis, mais de définir l'importance des mécanismes à mettre en place en interne pour assurer cette conformité⁶⁰.

Le RGPD et la loi informatique et libertés modifiée imposent des exigences spécifiques en matière de collecte, de traitement, de conservation et de sécurité des données personnelles dont l'essence a déjà pu être évoquée précédemment. L'utilisation de systèmes d'IA, en particulier ceux s'appuyant sur des volumes importants de données, peut dès lors soulever des questions complexes en matière de protection des données⁶¹.

Ainsi, les SIA traitant des données personnelles de manière limitée, avec des mesures de minimisation et de pseudonymisation efficaces, et sans impact significatif sur les droits et libertés des personnes concernées, pourraient être soumis à des obligations allégées. Il s'agit notamment des projets s'appuyant sur des données ouvertes et peu sensibles (travaux parlementaires, textes législatifs)⁶².

A l'inverse, les SIA traitant des données sensibles potentiellement intègres sont soumis à des processus de contrôle et de suivi rigoureux et nécessiteront parfois la réalisation d'analyses d'impact sur la protection des données (AIPD)⁶³.

Le niveau d'effort requis pour assurer la conformité avec la réglementation en matière de protection des données personnelles aura un impact direct sur les ressources humaines,

⁶⁰ Commission européenne pour l'efficacité de la justice (CEPEJ), *Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement*, adoptée lors de la 31e réunion plénière de la CEPEJ (Strasbourg, 3-4 décembre 2018), Conseil de l'Europe, février 2019, p. 60.

⁶¹ V. Étude du Conseil d'État à la demande du Premier ministre, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, adoptée en assemblée générale plénière le 31 mars 2022, p. 169

financières et techniques à mobiliser pour le développement, le déploiement et la maintenance des SIA.

Ce critère peut donc être décliné en s'appuyant sur les questions suivantes : le SIA traite-t-il des données personnelles ? Si oui, quelles sont les catégories de données concernées et leur niveau de sensibilité ? Quelles sont les mesures à mettre en place pour garantir le respect des principes du RGPD (minimisation, limitation des finalités, transparence, sécurité...) ? Une AIPD est-elle nécessaire ?

Note	Signification
0	Il existe un risque majeur de non-conformité au RGPD
1	Le SIA traite des données personnelles sensibles non disponibles en <i>open data</i>
2	Le SIA traite des données personnelles sensibles contenues dans les décisions de justice publiées en <i>open data</i>
3	Le SIA ne traite pas de données personnelles sensibles

Section 3 – Critères fonctionnels

§ 1 – Gain de qualité

Le gain de qualité attendu d'un système d'IA doit être évalué à l'aune du travail préexistant de l'humain assisté des outils de recherche actuellement déployés au sein de la Cour de cassation⁶⁴.

La prise en compte de la pertinence d'un cas d'usage au travers de ce prisme qualitatif supposera donc de mettre en relation la qualité actuelle des travaux effectués par les magistrats, juristes et fonctionnaires de la Cour de cassation avec les résultats susceptibles d'être générés par l'IA.

La capacité de l'IA à accomplir une tâche qui n'était pas exécutable humainement, en raison du volume des données en présence par exemple, constitue nécessairement une amélioration qualitative. Il en va de même de la capacité de la Cour de cassation à détecter des contentieux et à organiser des audiences thématiques.

⁶⁴ Lors de son audition le 14 novembre 2024, Me Kami Haeri a souligné que l'intelligence artificielle doit être abordée comme un outil de performance par les structures qui l'incorporent.

A l'inverse, la génération de contenus sériels standardisés ou de qualité moyenne permettant au magistrat, juriste ou membre du greffe de disposer d'une première version de travail ne peut être regardée en soi comme améliorant la qualité des productions en présence mais plutôt l'efficacité d'un processus préexistant⁶⁵.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : la mise en place du SIA correspond-elle à une activité humaine préexistante ? Si oui, les résultats attendus sont-ils plus pertinents que ceux fournis par les outils actuels ? Plus exhaustifs ?

Note	Signification
0	L'utilisation du SIA n'aura pas d'impact sur la qualité des travaux
1	L'utilisation du SIA améliorera légèrement la qualité des travaux
2	L'utilisation du SIA améliorera significativement la qualité des travaux
3	L'utilisation du SIA améliorera fortement la qualité des travaux

§ 2 – Gain d'efficacité

Le gain d'efficacité apporté par un système d'IA au sein de la Cour de cassation doit être évalué en considérant le temps et les ressources actuellement nécessaires aux magistrats, juristes et fonctionnaires pour accomplir leurs tâches, en utilisant les outils à leur disposition⁶⁶.

L'analyse de ce gain d'efficacité impliquera donc de comparer les processus de travail existants, incluant les délais et les ressources mobilisées, avec les améliorations potentielles offertes par l'IA en termes de rapidité d'exécution, d'automatisation des tâches répétitives ou de réduction des coûts⁶⁷.

La capacité de l'IA à traiter un volume important de dossiers ou à effectuer des recherches documentaires complexes en un temps considérablement réduit constitue une

⁶⁵ V. Rapport d'information de Mmes Marie-Pierre de La Gontrie et M. Christophe-André Frassa, fait au nom de la commission des lois, *L'intelligence artificielle et les professions du droit*, Sénat, 18 décembre 2024, p. 36.

⁶⁶ *Ibid.*

⁶⁷ Lors de l'audition des membres de la Cour d'appel de Paris concernant leurs travaux sur l'intelligence artificielle du 17 décembre 2024, M. Nicolas Dupont-Frugier a indiqué que le temps de rédaction actuel des décisions de justice par les magistrats est une difficulté quotidienne qui paraissait insolvable avant l'avènement des systèmes d'IA.

amélioration d'efficacité manifeste. La qualité de la donnée générée sera en revanche probablement plus faible compte tenu de la complexité des tâches à effectuer.

À l'inverse, la simple automatisation de tâches administratives ou la génération de documents standardisés qui se substituent à des systèmes préexistants de trames, bien qu'utiles, ne représentent pas nécessairement un gain d'efficacité significatif si elles n'entraînent pas une réduction substantielle du temps de travail des personnels ou une optimisation des ressources.

Ce critère peut donc être décliné en s'appuyant sur les questions suivantes : l'IA permet-elle de réaliser plus rapidement les tâches existantes ? Permet-elle de traiter un volume de dossiers plus important avec les mêmes ressources ? Automatise-t-elle des tâches répétitives, libérant ainsi du temps pour des activités à plus forte valeur ajoutée ?⁶⁸

Note	Signification
0	Le SIA ne génère aucun gain de temps
1	Le SIA génère un gain de temps faible
2	Le SIA génère un gain de temps significatif
3	Le SIA génère un gain de temps très important

Chapitre 2 – Présentation des critères d'évaluation techniques et économiques

L'analyse des critères techniques retenus par le groupe de travail s'appuie en partie sur l'expérience acquise par la Cour de cassation pour la mise en production des systèmes d'IA conçus par le laboratoire innovation du SDER à savoir la préorientation automatique et la pseudonymisation des décisions de justice.

Elle distingue les critères purement techniques en lien avec la faisabilité scientifique du projet, des critères économiques relatifs au coût d'entraînement, de déploiement, de mise en production et d'entretien des SIA.

Le développement et la mise en œuvre de projets d'IA, en particulier fondé sur de l'apprentissage automatique, reposent sur la mobilisation des ressources suivantes :

⁶⁸ V. Gouvernement, Rapport de la commission de l'intelligence artificielle, *IA : notre ambition pour la France*, mars 2024, p. 81

- Des experts métiers en mesure d'orienter la conception et la réalisation du produit ;
- Des experts techniques en mesure d'en assurer la mise en œuvre, en particulier des *data scientists* ;
- Des données à exploiter et le cas échéant des annotateurs pour préparer la donnée afin d'entraîner un algorithme d'apprentissage automatique ;
- De la puissance de calcul, nécessaire aux phases d'entraînement des algorithmes d'apprentissage automatique, ainsi qu'aux phases dites d'inférence pour leur exploitation.

La Cour de cassation dispose d'ores et déjà d'une équipe composée de deux *data scientists* et de puissance de calcul, installée sur ses serveurs et dédiée aux applications d'intelligence artificielle.

Toutefois, les équipes internes telles qu'elles sont actuellement constituées pourraient nécessiter des renforts pour la mise en œuvre de nouveaux projets.

De même, de la puissance de calcul supplémentaire serait nécessaire, en particulier pour la mise en œuvre de projets qui reposeraient sur de larges modèles de langage (voir Annexe 1 : présentation des larges modèles de langage).

La puissance de calcul nécessaire au développement et au déploiement des projets peut être obtenue en souscrivant des services d'informatique en nuage (*cloud computing*), avec des contraintes supplémentaires, notamment de coût, si des données protégées doivent y être exploitées. Elle peut également être déployée sur les serveurs de la Cour de cassation, ce qui paraît, en l'état, la solution la plus économique permettant de répondre à l'ensemble des obligations de conformité pour le traitement de données protégées. La puissance de calcul nécessaire à des phases d'entraînement peut également être recherchée dans le cadre de partenariats avec des laboratoires de recherche.

A cet égard, il convient de préciser que la Cour de cassation est partenaire du *Sorbonne Cluster for Artificial Intelligence* et que ce partenariat peut ouvrir des possibilités d'accès à une forte puissance de calcul ainsi qu'apporter des ressources en *data science*. D'autres partenariats avec des universités publiques comme avec des acteurs institutionnels, voire avec d'autres professions du droit, pourraient être également envisagés.

Section 1 – Critères techniques

§ 1 – Disponibilité de la donnée

La question de la disponibilité de la donnée apparaît primordiale. Les outils d'IA reposant sur de l'apprentissage automatique (*machine learning*) nécessitent de grandes quantités de données, idéalement de bonne qualité⁶⁹.

Pour illustration, l'algorithme de pseudonymisation des décisions de justice a été entraîné sur 200 000 décisions de justice pour qu'il assimile le langage juridique français utilisé dans les décisions de justice. Il a ensuite été entraîné sur une dizaine de milliers de décisions annotées, c'est-à-dire dans lesquelles avaient été identifiées des entités d'intérêt (personnes physiques, personnes morales, adresses, etc.)

Il est donc important d'avoir accès techniquement (c'est-à-dire d'être en mesure d'accéder aux données) et légalement (avoir le droit d'utiliser ces données) à de larges quantités de données⁷⁰.

La donnée n'est toutefois pas forcément disponible sous une forme exploitable. Il faut parfois l'extraire (des mémoires ampliatifs au format PDF dans le cadre de la préorientation par exemple), la filtrer, la transformer, l'agréger.

Enfin, il est souvent nécessaire d'annoter la donnée pour qu'elle puisse être utilisée pour une tâche spécifique. Cette annotation peut être réalisée par IA (en combinant un ensemble de règles issues des processus métier). Elle peut également nécessiter un travail d'annotation humain dédié souvent très long ou encore s'appuyer sur des données déjà enrichies lorsque l'IA a notamment vocation à remplacer ou assister l'humain dans une tâche (le traitement des mémoires ampliatifs et leur attribution aux différentes chambres est ainsi effectué par le SDER depuis près de 20 ans).

Le point central associé au critère de disponibilité de la donnée est donc de savoir avec quelle facilité il est possible d'accéder à une donnée de qualité utilisable pour entraîner un algorithme de *machine learning*.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : la donnée existe-t-elle ? Faut-il des autorisations pour y accéder et l'utiliser ? Faut-il la retravailler/la filtrer/l'agréger ? Faut-il l'annoter ? Si oui, a-t-on besoin d'annotateurs experts ?

Note	Signification
0	La donnée ne serait disponible qu'avec l'accord d'une tierce partie.

⁶⁹ Lors de l'audition des membres de la CEPEJ du 18 décembre 2024, M. Matthieu Quiniou a insisté sur l'importance de disposer d'une base de données de qualité et bien structurée afin de travailler sur des systèmes d'intelligence artificielle analytiques ou génératifs.

⁷⁰ Lors de l'audition des membres du parquet général de la Cour d'appel de Paris concernant leurs travaux sur l'intelligence artificielle du 17 décembre 2024, M. Nicolas Dupont-Frugier a donné l'exemple de la qualité insuffisante de la numérisation des procès-verbaux qui empêche la production de synthèses, ce pour illustrer les conséquences potentielles d'une mauvaise disponibilité de la donnée.

1	La donnée est disponible et nécessite d'importants travaux de transformation/d'annotation.
2	La donnée est disponible et nécessite des travaux de transformation/d'annotation.
3	La donnée est utilisable telle quelle.

§ 2 – Preuve de faisabilité

Pour estimer la faisabilité d'un projet d'apprentissage automatique, il est important de connaître l'état des recherches engagées sur ce même sujet ou des sujets similaires par d'autres institutions, administrations, laboratoires de recherche ou acteurs privés⁷¹.

Le milieu de l'IA et de l'apprentissage automatique est en effet un milieu relativement ouvert, sur lequel différents acteurs communiquent régulièrement. De nombreux modèles pré-entraînés ou non sont rendus disponibles chaque jour, améliorant les performances de l'état de l'art sur des tâches spécifiques avec des jeux de données de référence.

En conséquence, la production scientifique et industrielle autour d'une problématique apparaît comme un bon indicateur de faisabilité d'un projet d'apprentissage automatique. En d'autres termes, si peu de chercheurs ou d'entreprises se sont penchés sur un sujet et ont proposé des solutions, cela nécessitera plus de ressources et de temps de développement⁷².

Ainsi, l'algorithme de pseudonymisation des décisions de justice est un algorithme de reconnaissance d'entités nommées fondé actuellement sur des modèles de langages entraînés par l'INRIA et affinés par la Cour. L'algorithme de préorientation est lui fondé sur un algorithme de classification assez classique (TF-IDF⁷³). Ces deux tâches relèvent de problématiques très connues de la recherche en intelligence artificielle (citées respectivement 1 600 et 6 000 fois sur *Google Scholar* en janvier 2025).

Il paraît donc important d'estimer la faisabilité d'un projet d'apprentissage automatique en identifiant s'il existe des productions scientifiques en attestant. Une autre preuve de faisabilité serait aussi la mise en place de dispositifs similaires dans des institutions publiques ou privées.

⁷¹ Lors de son audition du 11 décembre 2024, M. Thomas Lyon-Caen a indiqué que la qualité des systèmes d'intelligence artificielle est souvent corrélée à la taille du marché dans lequel ils s'inscrivent.

⁷² Lors de son audition du 8 janvier 2025, M. Arno Amabile a précisé que le développement de systèmes d'intelligence artificielle spécialisés doit passer par la mise à disposition des innovateurs de modèles sur lesquels ils pourront itérer, créer leur propre produit, ce à un coût moindre.

⁷³ Le TF-IDF (de l'anglais term frequency-inverse document frequency) est une méthode de pondération souvent utilisée en recherche d'information et en particulier dans la fouille de textes.

De plus, une large production scientifique et industrielle sur une problématique donnée assure plus fortement de trouver des briques déjà existantes et dont le code informatique et les modèles sont accessibles publiquement, ce qui accélérerait le développement d'un projet. C'est aussi la garantie de l'existence d'une communauté d'acteurs qui peuvent être des ressources pour le projet.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : existe-t-il des solutions connues ? Dans le milieu public, privé, dans des documents de recherche ? Est-ce une problématique identifiée ? Existe-t-il des solutions ouvertes ?

Note	Signification
0	Le problème ou une variante n'est pas résolu dans la littérature scientifique.
1	La littérature scientifique montre la difficulté de réaliser ce genre de projet.
2	Il existe des solutions ouvertes et de nombreux articles prouvant la faisabilité du projet.
3	Il s'agit d'une problématique techniquement simple.

Section 2 – Critères économiques

§ 1 – Temps de ressources en science de la donnée

Le temps de développement désigne ici le temps de travail nécessaire aux *data scientists* dans le cadre d'un projet d'apprentissage automatique/IA. Plus le projet sera complexe, plus il risque d'impliquer un temps de développement long.

Le travail des *data scientists* peut ainsi être segmenté de la manière suivante :

1. Cadrage du projet
2. Acquisition de la donnée
3. Transformation de la donnée
4. Analyse de solutions existantes (« *benchmark* »)
5. Adaptation des solutions (ré-entraînement, adaptation des modèles) ou développement d'un modèle *ad hoc* (« *from scratch* »)
6. Évaluation des solutions trouvées
7. Mise en place d'une interface ou d'un processus automatique de traitement

8. Suivi de l'outil et amélioration/adaptation si nécessaire

Ces étapes ne sont cependant pas strictement séquentielles, elles peuvent être parallélisables ou nécessiter un processus itératif.

Comme explicité au travers des deux premiers critères, les étapes de transformation de la donnée et d'adaptation de solutions existantes peuvent être chronophages. Les étapes de veille et d'auto-formation sont aussi non négligeables dans le processus.

A titre d'exemple, la solution de pseudonymisation des décisions de justice a nécessité 10 mois de travail de deux *data scientists*, de l'analyse des modèles adéquats, à la mise en production du moteur en passant par l'exploitation des données annotées et non annotées. Aujourd'hui, un ré-entraînement du modèle peut nécessiter jusqu'à un mois de travail.

La question du temps de développement est un des principaux facteurs de limitation des projets.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : la donnée nécessite-t-elle beaucoup de transformations ? Est-il possible d'utiliser des briques déjà existantes ? Faut-il ré-entraîner le modèle ? Si oui, peut-on se contenter de l'affiner ? Une formation des équipes internes sera-t-elle nécessaire ?

Note	Signification
0	Le projet nécessiterait de faire appel à un nombre conséquent de <i>data scientists</i> pendant une très longue période de temps.
1	Le projet peut être réalisé par deux <i>data scientists</i> en deux ans.
2	Le projet peut être réalisé par deux <i>data scientists</i> en moins d'un an.
3	Le projet peut être réalisé par deux <i>data scientists</i> en moins de six mois.

§ 2 – Temps de ressources métiers

Comme indiqué à l'occasion de l'examen du critère précédent, le temps de développement technique d'un projet d'IA inclut des étapes de cadrage du projet, de transformation des données et d'évaluation des solutions trouvées. Si les *data scientists* sont impliqués dans ces étapes, ils répondent à des problématiques métier dont les magistrats, greffiers et autres personnels de la Cour ont seuls la maîtrise. Les experts « métiers » possèdent en effet un savoir juridique et pratique dont ne disposent pas les experts techniques.

L'implication du métier dans le projet, à presque chaque étape de celui-ci, est donc primordiale pour ne pas s'éloigner des problématiques qu'essaie de résoudre la solution mise en place. C'est pourquoi le critère de disponibilité de la donnée évoqué précédemment tient notamment compte de la possibilité d'avoir recours à une annotation manuelle de celle-ci par les experts « métiers » qui, étant les destinataires finaux des outils développés, sont par ailleurs les mieux à même d'évaluer la pertinence et performance de la solution proposée.

Dans le cadre de la préorientation automatique, ce sont les experts « métiers » qui ont indiqué aux *data scientists* quelles parties du mémoire ampliatif étaient pertinentes pour la préorientation, évitant de prendre en compte tout le document et allégeant ainsi l'architecture du modèle. Pour l'algorithme de pseudonymisation, ce sont des agents de la cellule d'anonymisation qui ont annoté les décisions de justice.

L'absence d'implication suffisante du métier est régulièrement citée lors de conférences spécialisées en IA comme une cause d'échec des projets d'apprentissage automatique ou d'IA. Un tel projet ne peut se faire sans l'implication du métier qui doit l'irriguer en permanence et qui peut aussi être impliqué sur des tâches chronophages d'annotation ou d'évaluation des solutions.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : de quel niveau d'implication a-t-on besoin de la part du métier ? A-t-on besoin simplement d'implication ponctuelle ? A-t-on besoin d'une longue phase d'annotation/d'évaluation des solutions proposées ?

Note	Signification
0	Le projet nécessiterait une charge sur le personnel incompatible avec le fonctionnement de la Cour.
1	Le projet nécessite une équipe de personnels de la Cour dédiée pendant une grande partie de leur temps.
2	Le projet nécessite un ou deux personnels de la Cour à temps plein ou pendant une grande partie de leur temps.
3	Le projet nécessite l'intervention ponctuelle de personnels de la Cour.

§ 3 – Autres ressources de développement

La Cour de cassation possède actuellement de nombreuses ressources internes pour assurer le développement des projets d'IA.

Outre le temps et les compétences des *data scientists* et experts « métiers », il est cependant nécessaire de prendre en compte la nécessité de recourir à d'autres ressources techniques.

Certaines de ses ressources sont également disponibles à la Cour. Le laboratoire d'innovation compte en effet des développeurs *front/back* (parties visibles et non visibles des systèmes informatiques) ou ingénieurs *Devops* (en charge de mettre à disposition les ressources matérielles (*hardware*) aux applications (*software*)). Toutefois, lorsqu'un outil propose une interface à des utilisateurs, il est important de se reposer sur des designers UI (*User Interface* : interface utilisateurs) et/ou UX (*User Experience* : expérience utilisateurs) qui permettront de donner un outil ergonomique, utilisable et répondant bien aux besoins des utilisateurs finaux.

Lors du développement du logiciel de relecture des décisions pseudonymisées LABEL, un soutien *UX/UI designer* a été sollicité afin d'adapter l'interface aux besoins des agents annotateurs. Par ailleurs, la pseudonymisation des décisions de justice est entièrement soutenue par le travail des ingénieurs *Devops* du laboratoire innovation.

La Cour de cassation bénéficie en outre de sa propre direction des systèmes d'information, ce qui constitue une opportunité, dès lors que l'intégration de SIA touche au système d'information de la Cour lui-même, ce qui implique une mobilisation de cette direction (spécification des besoins, développement des interfaces, intégration des traitements de l'IA aux applications, infrastructure hébergeant les modèles, homologation de sécurité, etc.) Une coordination des différentes équipes est par conséquent une condition essentielle au succès de l'implantation de SIA à la Cour.

Il y a lieu également de prendre en compte les particularités liées à la sécurisation du réseau informatique de la Cour et du ministère de la justice. Eu égard à la sensibilité des données traitées à l'occasion de certains cas d'usage, la question de l'hébergement de l'infrastructure en tout ou partie sur un *cloud SecNumCloud*, indépendamment du coût associé à un tel hébergement, pourra nécessiter le recours à des partenaires extérieurs à la Cour.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : a-t-on besoin de faire intervenir des ressources externes ? Développeurs ? Ingénieurs *Devops* ? *Designer UX* ? *Designer UI* ? Ingénieurs réseaux du ministère ? Consultant *clouds* ?

Note	Signification
0	Le projet nécessite l'intervention de très nombreuses ressources externes à la Cour
1	Le projet nécessite l'intervention de ressources internes et externes à la Cour.
2	Le projet nécessite une intervention de ressources internes à la Cour

3	Le projet ne nécessite pas ou peu de ressources techniques (hors data science)
---	--

§ 4 – Puissance et temps de calcul pour l'entraînement des modèles

Les modèles d'IA les plus volumineux, reposant sur des LLM à plusieurs centaines de milliards de paramètres, sont aujourd'hui très consommateurs de ressources, notamment de calcul. Ces ressources sont un facteur limitant des projets d'IA.

Entraîner ou ré-entraîner un modèle d'IA nécessite de très nombreux calculs, en général assez simples et parallélisables, mais qui s'appuient tout de même des architectures reposant sur des *GPU* (*Graphical Processing Unit* : cartes graphiques).

L'algorithme de pseudonymisation fonctionne actuellement en exploitation sur une carte graphique A40. Il prend ainsi 4 secondes pour chaque décision. A titre de comparaison, sur un ordinateur personnel doté d'un *CPU* (*Central Processing Unit*), autrement dit unité centrale de traitement, l'opération nécessite 50 minutes.

Certains modèles de langage sont si volumineux qu'ils ne peuvent être chargés sur des architectures comme celle utilisée en ce moment à la Cour, encore moins entraînés ou ré-entraînés. Si certaines tâches peuvent être traitées avec des modèles de taille moyenne à petite, d'autres nécessitent, pour être à l'état de l'art, des modèles plus grands et donc difficilement traitables dans l'état actuel des ressources de la Cour.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : quelles sont les tailles des modèles communément utilisés pour répondre à cette tâche ? Combien de temps est nécessaire pour ré-entraîner le modèle ? Peut-on le faire avec les ressources actuelles ?

Note	Signification
0	Le projet nécessite une puissance de calcul inaccessible.
1	Le projet nécessite une puissance de calcul accessible par un investissement important dans des ressources matérielles ou l'accès l'extérieur à une forte puissance de calcul (partenariat, cloud).
2	Le projet nécessite une puissance de calcul importante mais pouvant être déployée sur l'environnement actuel de la Cour de cassation.
3	Le projet ne nécessite ni temps de calcul important, ni ressources additionnelles.

§ 5 – Coût d'exploitation

Le coût d'exploitation associé à un système d'IA correspond aux frais de maintenance et d'hébergement des modèles développés.

Comme vu au point précédent, les modèles d'IA nécessitent des architectures spécifiques pour être entraînés. Ils nécessitent aussi une architecture *ad hoc* lors de la phase d'exploitation (phase d'inférence), c'est-à-dire après la mise en production du modèle. Cette architecture dépend de plusieurs facteurs : la taille du modèle, la fréquence d'utilisation, la taille des données à traiter, les exigences de temps de réponse...⁷⁴.

Ainsi, la conception d'un agent conversationnel accessible par l'ensemble des personnels de la Cour nécessite qu'il soit en mesure de calculer très rapidement des réponses et éventuellement de supporter une charge importante d'utilisateurs en même temps.

Au contraire, si l'on doit calculer les rapprochements (matériels et/ou intellectuels) entre des pourvois incidents arrivés pendant la journée et les pourvois déjà déposés, ce calcul peut être fait une fois, par exemple la nuit, ce qui permet de se détacher partiellement des contraintes de temps en journée. Le modèle d'IA et sa potentielle interface nécessitent aussi d'être maintenus en conditions opérationnelles et surveillés. Cela nécessite du temps de travail des personnels techniques.

Enfin, l'algorithme doit être surveillé et ses performances mesurées. Si l'algorithme est entraîné sur des données à un moment donné, il n'est pas rare que ces données évoluent avec le temps. Pour ne pas affecter les performances de l'algorithme, il est important de ré-entraîner régulièrement le modèle.

L'algorithme de pseudonymisation des décisions de justice est ainsi ré-entraîné régulièrement pour prendre en compte les évolutions des sujets traités dans les décisions et les nouvelles sources de données comme les tribunaux de commerce. En l'occurrence, l'évolution de la rédaction des décisions de la Cour ainsi que l'intégration de nouvelles sources de données affectent la performance de l'algorithme. Une campagne de ré-entraînement prend un mois d'annotation et de calculs.

La question du coût et du temps de maintien en conditions opérationnelles n'est pas à négliger car elle peut impliquer des besoins humains importants. À cet égard, lors de leur audition, les représentants de la CEPEJ ont mis en exergue la nécessité pour les institutions publiques, notamment la Cour de cassation, de disposer de *clouds* souverains pour l'hébergement de données et la mise à disposition de puissance de calcul à distance.

⁷⁴ Lors de son audition du 8 janvier 2025, M. Arno Amabile a attiré l'attention sur le coût important associé aux SIA génératifs qui s'appuient notamment sur les derniers modèles et sur la nécessité d'une réflexion préalable sur la faisabilité des projets au regard de ces coûts.

Ce critère peut donc être décliné en s'appuyant sur une série de questions qui détermineront la « note » à associer : combien coûte l'hébergement de la solution choisie ? Quelles sont les exigences en termes de disponibilité ? Combien de temps faut-il passer à maintenir l'outil ? A quelle fréquence faut-il ré-entraîner le modèle ?

Note	Signification
0	L'exploitation implique une charge très lourde en matériel et/ou personnel.
1	L'exploitation implique la prise en compte d'une charge lourde en matériel et/ou personnel.
2	L'exploitation implique la prise en compte d'une charge supplémentaire limitée en matériel et/ou personnel.
3	L'exploitation n'implique pas de charge matérielle et/ou personnelle supplémentaire.

PARTIE 3

PRÉSENTATION ET EVALUATION DES CAS D'USAGE IDENTIFIÉS

Le groupe de travail a identifié, au cours de ses travaux, trois grandes familles de cas d'usage visant à l'exploitation ou l'analyse des écritures des parties, la recherche dans les bases de données internes ou publiques de jurisprudence et de doctrine et l'aide à la rédaction.

Ces familles reflètent le cycle de vie d'une affaire devant la Cour de cassation depuis le dépôt du mémoire ampliatif jusqu'à la rédaction de la décision mettant fin à l'instance, en passant par les travaux de recherches et de rédaction des travaux préparatoires qui permettront d'éclairer les débats devant les chambres.

Il est cependant apparu au cours des travaux que certains cas d'usage identifiés ne s'apparentaient pas spécifiquement à une de ces familles mais constituaient un socle exploitable pour l'ensemble des projets abordés ou du moins un paramètre facilitant leur développement. Ces cas d'usage dits « transversaux » seront donc abordés à titre liminaire.

Par ailleurs, les cas d'usage propres à l'activité du greffe, qui présentent des spécificités, seront évoqués séparément.

Chapitre 1 – Cas d'usage transversaux : structuration et enrichissement des documents

Il s'agit, selon les cas, de définir la structure des documents en identifiant leurs différentes parties, et de repérer les références aux textes législatifs ou réglementaires, à la jurisprudence (judiciaire ou administrative, nationale ou internationale), à des articles ou ouvrages de doctrine, et idéalement de les relier (par lien hypertexte) à la source à laquelle il est fait référence ou d'intégrer directement le contenu de cette source lorsqu'elle est disponible.

Ce cas d'usage peut s'appliquer à tout document (rapport, avis, décision, mémoire) et pourrait être notamment un préalable à une meilleure exploitation des écritures, de la jurisprudence contenue en base de données ainsi qu'à une aide à la rédaction des rapports et avis, et subséquemment des décisions. Des applications pourraient être notamment mises en service pour améliorer le moteur de recherche Judilibre, qui donne un accès gratuit à l'ensemble de la jurisprudence mise à disposition en *open data*.

A titre d'illustration, le travail d'orientation des pourvois vers les chambres et leurs sections réalisé par le SDER, comme celui d'orientation vers les différents circuits procéduraux effectué au sein des chambres par les conseillers orientateurs, s'appuie sur les écritures des parties et nécessite de prendre connaissance des textes normatifs cités au soutien des moyens des parties.

Cette action impose souvent d'interroger le site Légifrance, ce qui, pris isolément, ne représente pas une opération particulièrement complexe ni chronophage. Toutefois, la charge de travail induite, et mécanique, doit être appréciée à l'aune des 13 500 affaires dont est saisie annuellement la Cour de cassation à raison, *a minima*, en matière civile, d'un mémoire ampliatif et d'un mémoire en défense, sauf rares cas, par affaire, et le plus souvent de plusieurs références par mémoire.

➤ *Evaluation du cas d'usage :*

L'enrichissement des documents utilisés et générés par la Cour de cassation représenterait un gain de qualité substantiel en ce qu'il permettrait d'en faciliter l'exploitation et la lecture. La recherche de références jurisprudentielles ou textuelles par les membres de la Cour de cassation constitue par ailleurs un exercice indispensable et chronophage, comme cela a pu être rappelé précédemment.

En outre, ce travail d'identification des textes est de nature à simplifier la lecture des travaux préparatoires, rapports et avis, et à en valoriser le contenu auprès des utilisateurs internes ou externes à la Cour. Il est de nature à générer un gain de temps pour les magistrats des chambres qui doit s'apprécier à l'aune du nombre de textes cités au sein de ces travaux.

Pour autant, en ce qui concerne la référence aux textes normatifs, il convient de noter que l'outil renverrait vers la version en vigueur qui ne sera pas nécessairement la version applicable au litige. Il appartiendra à l'utilisateur de naviguer vers les différentes versions applicables dans le temps.

Sur le plan éthique, l'enrichissement étant réalisé à partir d'un contenu fini et en utilisant des références de jurisprudence ou de textes publiés, ce cas d'usage n'est pas de nature à affecter les droits fondamentaux des justiciables et le cas d'usage reposerait sur une IA frugale et un lot délimité de données d'entraînement. Compte tenu cependant de la présence de nombreuses données personnelles au sein des documents à enrichir (rapports et avis ou mémoires des parties), une attention particulière devra être accordée au respect du RGPD dans le cadre de ce traitement.

Il s'agit enfin d'une problématique technique simple pouvant être mise en œuvre à moyen terme (un an) par deux *data scientists* en s'appuyant sur les ressources de la Cour. La complexité de ce projet dépendra cependant de l'étendue du périmètre des enrichissements et de la variété des bases de données à interroger ou à intégrer.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	3
	Coût d'exploitation	3

Chapitre 2- L'exploitation des écritures des parties

L'orientation et le traitement des pourvois entrant à la Cour de cassation repose sur un travail intellectuel de première analyse du mémoire ampliatif et d'identification des questions de droit posées.

Il se décompose ainsi en plusieurs phases d'identification des matières juridiques concernées via une grille de « codes », de détection du caractère récurrent ou au contraire nouveau de la question de droit, ainsi que d'une première analyse de la jurisprudence en présence.

L'objectif des différentes phases de triage est aussi d'identifier matériellement la chambre compétente, de définir la complexité, à première vue du dossier, afin de répartir la charge de travail des conseillers rapporteurs, de proposer une formation de jugement adaptée aux difficultés posées par le dossier et enfin de privilégier une organisation thématique des audiences. Au sein de cette phase, l'exploitation des écritures des parties revêt une place importante.

L'ensemble de ces objectifs constitue un terrain favorable à l'usage de l'IA et donne lieu à des cas d'usage développés et évalués ci-dessous.

Pour la mise en œuvre de ces critères d'évaluation, le groupe de travail relève que les cas d'usage de cette section nécessitent de s'appuyer sur les écritures intègres des parties, ce qui crée nécessairement un facteur de risque au regard de la réglementation sur la protection des données personnelles.

Par ailleurs, la plupart des cas d'usage envisagés n'est, a priori, pas susceptible d'affecter le processus décisionnel, ce qui constitue un facteur facilitant le respect de la réglementation sur l'IA.

Section 1 – Traitement des pourvois par le SDER

§ 1 – Orientation des mémoires ampliatifs

En matière civile, les pourvois font l'objet d'une orientation vers les chambres de la Cour de cassation par le SDER dont les équipes de juristes, sous la direction de magistrats, auditeurs, analysent les mémoires ampliatifs et renseignent les codes matières correspondants.

Cette opération repose sur une nomenclature des affaires orientées (NAO) qui compte, à l'instar de la nomenclature des affaires civiles (NAC) pour les juridictions du fond, plus de 900 codes différents se rattachant à la nature des affaires. Chaque code est rattaché à une section de chambre.

Cette orientation humaine est assistée par un algorithme de préorientation qui a déjà été présenté dans le cadre de ce rapport. Elle représente malgré tout un temps de travail non-

négligeable estimée entre 25 et 40 % des équivalents temps plein travaillé des 12 juristes présents au sein des bureaux miroirs des chambres civiles, commerciale et sociale.

Ce travail est contrôlé et enrichi au sein des chambres par les conseillers orientateurs.

En matière pénale, la question d'un choix d'orientation vers une chambre ne se pose pas. Les dossiers sont orientés vers les différentes sections au terme d'une analyse humaine reposant sur la nature des affaires.

L'amélioration du dispositif existant, afin de permettre l'orientation automatique des pourvois, sans contrôle humain ou avec un contrôle humain allégé, vers les différentes sections des chambres, pourrait constituer une opportunité.

➤ *Evaluation du cas d'usage :*

L'orientation réalisée actuellement par l'algorithme de préorientation est fiable à 90 %. Cette analyse est contrôlée par les membres des bureaux de chambre du SDER qui corrigent les erreurs marginales commises par le modèle et affectent aux pourvois, à cette occasion, un code « Nomenclature des Affaires Orientées » (NAO).

Compte tenu du temps de travail précédemment évoqué des juristes des bureaux miroirs des chambres civiles, commerciale et sociale que représente l'orientation humaine des pourvois, l'automatisation totale de ce processus représenterait un gain substantiel de productivité.

Si ce cas d'usage repose sur une technologie identifiée et frugale, il apparait cependant qu'il se heurte à un obstacle technique en l'absence de données annotées en nombre suffisant. La granularité de la NAO impose en effet de disposer de nouveaux jeux de données, ce qui nécessiterait de mobiliser de manière importante des personnels de la Cour pour entraîner le modèle en temps réel.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	1
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	3
	Ressources métier	0
	Autres ressources (Designer, Devops)	3

	Puissance et temps de calcul (entraînement)	3
	Coût d'exploitation	3

§ 2 – Pré-signallement automatique des pourvois

Le pré-signallement correspond à une opération réalisée par le SDER lors de la phase d'orientation des mémoires ampliatifs vers les chambres.

Il s'appuie sur les consignes adressées par les chambres aux bureaux miroirs du SDER et a pour objectif d'identifier les affaires répondant à certaines questions de droit qu'elles posent. Les signalements sont adressés chaque mois. Cette opération est réalisée humainement.

➤ *Evaluation du cas d'usage :*

Le pré-signallement des pourvois reposent sur une analyse humaine fine difficilement reproductible par un SIA.

Par ailleurs, la variété de ces critères et leur évolution dans le temps supposerait de recourir à un grand modèle de langage, coûteux et énergivore, ainsi que de mobiliser des ressources techniques et métiers conséquentes.

La mission IA identifie par conséquent un risque majeur de dégradation de la qualité des signalements s'ils devaient être réalisés par un SIA.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	1
Critères Techniques	Disponibilité de la donnée	1
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

Section 2 - Détection de connexité matérielle et intellectuelle

La recherche de connexité matérielle (identité de parties ou de décision attaquée) ou intellectuelle (identité de la question de droit posée) au sein des mémoires ampliatifs pendants devant la Cour de cassation est essentielle à la bonne administration de la justice. Elle permet, en cas de connexité matérielle, la jonction d'affaire pour les instruire ensemble, favorise, notamment pour les connexités intellectuelles, la tenue d'audiences thématiques et permet de prévenir d'éventuelles divergences de jurisprudence.

L'état des applicatifs internes de la Cour de cassation et la qualité de données transmises ne permettent toutefois pas une détection optimale de ces connexités qui relèvent par ailleurs d'un travail essentiellement humain.

L'exploitation par IA des mémoires ampliatifs, éventuellement couplées à une exploitation des décisions attaquées, permettrait ainsi d'affiner la connaissance de la Cour des pourvois en cours d'examen. La détection des connexités est particulièrement importante en ce qui concerne les litiges sériels.

Il n'est pas identifié d'impact négatif sur les droits fondamentaux lié à une assistance de ces recherches de connexité par un SIA. En outre, le gain de qualité apparaît majeur.

§ 1 – Connexité matérielle

➤ *Evaluation du cas d'usage :*

La détection des connexités matérielles existant entre les pourvois pendants devant la Cour constitue un cas d'usage simple en ce qu'il s'appuie sur des métadonnées de qualité et des règles aisément définissables (identité des deux parties, identité de décision attaquée...).

Il constitue ainsi un projet achevable à court terme, sous réserve de son intégration au système d'information de la Cour de cassation.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	3
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	3

	Ressources métier	3
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	3
	Coût d'exploitation	3

§ 2 – Connexité intellectuelle

➤ *Evaluation du cas d'usage :*

La recherche de connexité intellectuelle revêt une importance essentielle pour la Cour de cassation en ce qu'elle contribue à une meilleure vision des affaires pendantes et à une meilleure instruction des affaires et organisation des audiences. Cette opération repose aujourd'hui essentiellement sur l'expertise des conseillers en chambres ou des orientateurs du SDER en l'absence d'outils internes de recherche suffisamment performants.

Il s'agit cependant d'un projet complexe qui nécessiterait un investissement métier significatif mais pourrait s'appuyer sur des technologies relativement frugales, ce qui permettrait de réduire le coût d'exploitation d'un tel SIA.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2
	Ressources métier	1
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

Section 3 – Première analyse du dossier par les chambres

Une première analyse du dossier est réalisée par les magistrats des chambres afin de recenser notamment les moyens et branches des mémoires. Cette opération pourrait relever d'un applicatif métier n'ayant pas recours à l'intelligence artificielle grâce à l'utilisation de données structurées. D'autres aspects de cette première analyse pourraient néanmoins bénéficier d'un recours à l'IA.

§ 1 – Détection des critères de complexité

Sept critères de complexité d'un dossier peuvent être au stade d'une première analyse du dossier dans la chambre qui en est attributaire :

- Question de droit nouvelle pour la Cour ;
- Application de sources de droit variées (droit de l'Union européenne, droit international) ;
- Mise en balance complexe de droits fondamentaux (application de règles de droit national ou de droit européen) ;
- Existence d'une jurisprudence divergente ;
- Pourvoi posant une ou des questions susceptibles de relever de plusieurs chambres de la Cour ;
- Question de droit complexe mettant en exergue une articulation nécessaire entre le juge judiciaire et le juge administratif ;
- Pourvoi comportant plusieurs questions juridiques distinctes ne relevant pas toutes nécessairement de la section du conseiller.

Ces critères sont appréciés le cas échéant à l'aide des indications pouvant être données par les avocats aux Conseils et par les magistrats de la chambre. Une assistance de technologies d'IA pour l'identification de ces critères, pris ensemble ou chacun isolément, peut être envisagée. Certains de ces items sont à rapprocher de la détection de questions de droit nouvelles ou des contentieux émergents envisagés *infra* dans la section portant sur l'aide à la recherche et l'exploitation des bases de données.

➤ *Evaluation du cas d'usage :*

Les critères examinés par les chambres pour évaluer la complexité des dossiers qui leur sont soumis recourent en réalité de nombreux cas d'usage qui seront envisagés *infra*.

Il en va ainsi de l'identification des questions nouvelles, des rapprochements de jurisprudence interne ou internationale ou encore de la reconnaissance des références de textes ou de jurisprudence citées. Si l'examen de ces critères dans les mémoires ampliatifs présente des spécificités, il est proche de la recherche de ces éléments dans les décisions du fond.

La difficulté du cas d'usage d'évaluation de la complexité des dossiers reflète ainsi celle qui s'attache à l'ensemble des sous-cas qui le composent. La poursuite d'une évaluation automatique de la complexité des pourvois paraît donc être envisagée comme un objectif

complémentaire pouvant être atteint à l'issue d'autres projets. Elle pourrait ainsi se concevoir à long terme suivant l'avancée d'autres cas d'usage.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	1
Critères Techniques	Disponibilité de la donnée	1
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	2

§ 2 – Estimation du temps de traitement

Une estimation du temps prévisible que représentera le traitement des affaires distribuées entre les magistrats est utile.

➤ *Evaluation du cas d'usage :*

Il ressort de l'analyse de ce cas d'usage que la grande variété des critères retenus pour évaluer le temps nécessaire au traitement des dossiers est difficilement objectivable et pondérable par un SIA, et nécessiterait en toute hypothèse un entraînement significatif par référence au temps effectivement consacré aux dossiers et, en conséquence, un temps significatif consacré à celui-ci par des magistrats chargés de procéder à de telles estimations.

Cette charge et l'incertitude du résultat pouvant être atteint est par ailleurs à rapporter à la faible valeur ajoutée qui résulterait d'une estimation automatique du temps de traitement des pourvois.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2

Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	0
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 3 – Détection des précédents et rapprochements de jurisprudence au sein des décisions rendues par la Cour de cassation

Une recherche de jurisprudence est effectuée afin de déterminer s’il existe ou non une position arrêtée de la Cour de cassation sur la question de droit posée.

En l’état d’une position constante et non équivoque sur la question de droit qui lui est soumise, le dossier sera orienté vers des formations réduites (formation restreinte ou formation simple).

Ce travail de recherche est effectué sur la base de données Jurinet, interne à la Cour qui recense plus de 900 000 décisions rendues par la Cour de cassation ainsi que les travaux préparatoires y afférant.

Une solution d’IA peut être envisagée afin d’aider à la recherche de précédents et rapprochements de jurisprudence.

➤ *Evaluation du cas d’usage :*

Ce cas d’usage s’inscrit dans la continuité du projet « Divergences »⁷⁵ dont les résultats intermédiaires ont mis en évidence la faisabilité technique.

La poursuite d’une recherche à moyen terme permettrait ainsi d’améliorer les performances actuelles du SIA développé à partir d’un modèle de langage moyen.

⁷⁵ Voir Partie 1 Chapitre 1 Section 3 pour une description plus complète de ce projet

Il conviendra cependant d’être vigilant afin de garantir que l’orientation vers une formation réduite demeure, de manière effective, un choix strictement humain.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

Section 4 - Cartographie du litige

La cartographie du litige vise notamment à dresser une chronologie de la procédure au fond jusqu’au pourvoi devant la Cour de cassation et à identifier et classer les parties au litige en matérialisant visuellement les relations entre elles (assuré, assureur, sous-traitant, délégataire, intervenant volontaire, etc.). Elle vise également à relier entre elles les affaires présentant une connexité.

Ce travail d’analyse et de synthèse s’appuie actuellement sur les écritures des parties et pourrait être amélioré par l’utilisation d’outils d’IA.

➤ *Evaluation du cas d’usage :*

L’utilisation de l’IA en tant qu’outil d’aide à la cartographie des litiges constituerait manifestement un gain d’efficacité pour les magistrats, y compris du fond.

Cependant, ce cas d’usage présente dans l’ensemble une certaine complexité, dans la mesure où il reposerait sur l’exploitation combinée des métadonnées de la Cour de cassation ainsi que sur l’exploitation du texte des écritures des parties et des décisions rendues en première instance ou en appel. La capacité du SIA à avoir une juste compréhension de l’articulation des parties entre elles, sans pour autant déployer des ressources trop coûteuses, est un point d’incertitude fort.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

Chapitre 3 – Aide à la recherche et exploitation des bases de données documentaires

Les travaux de recherche conduits par les conseillers et avocats généraux pour l’instruction des dossiers qui leur sont confiés s’appuient en premier lieu sur les bases de données jurisprudentielles internes ou externes à la Cour et les bases doctrinales des éditeurs juridiques.

Les outils d’IA pourraient améliorer les capacités de recherche et d’exploitation des bases de données documentaires.

Section 1 – Exploitation de bases de données pour un usage interne

Le SDER tient deux bases de données en application de l’article R. 433-3 du code de l’organisation judiciaire⁷⁶ :

- la première rassemblant « *les décisions et avis de la Cour de cassation et des juridictions ou commissions juridictionnelles placées auprès d’elle, publiés ou non publiés aux bulletins mensuels mentionnés à l’article R. 433-4, ainsi que les décisions présentant un intérêt particulier rendues par les autres juridictions de l’ordre judiciaire* » (Jurinet),
- la seconde devant rassembler « *les décisions des premier et second degrés rendues par les juridictions de l’ordre judiciaire* » (Jurica).

La base Jurinet contient un ensemble de données constituées notamment de plus de 956 000 décisions de justice. La très grande majorité de ces décisions est issue des arrêts rendus par la Cour de cassation mais la base intègre également une sélection de décisions de cours d’appel et dans une moindre mesure de décisions de première instance.

La base Jurica contient 2 807 000 décisions rendues par les cours d’appel en matière civile, sociale et commerciale, à l’exclusion de la matière pénale. Cette base recense l’intégralité des décisions rendues depuis 2008 et une sélection de décisions antérieures.

Les bases Jurica et Jurinet sont actuellement accessibles par deux moteurs de recherche distincts. Il résulte de cette partition de la donnée que la consultation, par les magistrats et juristes de la Cour de cassation mais également des juridictions du fond, de l’état de la

⁷⁶ Dans sa version en vigueur depuis le 1er juillet 2020 suite à sa modification par l’article 4 du décret n° 2020-797 du 29 juin 2020. La base JURINET a été créée par décret n°2005-13 du 7 janvier 2005. La base JURICA a ensuite été créée par le décret n° 2008-522 du 2 juin 2008 et son périmètre portait initialement, jusqu’à la modification de l’article R. 433-3 du code de l’organisation judiciaire par le décret du 29 juin 2020, sur « *l’ensemble des arrêts rendus par les cours d’appel et décisions juridictionnelles prises par les premiers présidents de ces cours ou leurs délégués* ».

jurisprudence se rattachant à une question de droit précise, peut requérir d'interroger deux moteurs de recherche différents, trois s'ils souhaitent connaître l'état de la jurisprudence de première instance via le portail Judilibre.

Il apparaîtrait donc pertinent de regrouper la consultation de ces trois bases au sein d'un moteur de recherche unique.

Il pourrait également être pertinent d'adjoindre à ces bases de données, directement ou par renvoi, les fonds documentaires de la Cour ainsi que d'autres bases de données publiques de législation, réglementation et de jurisprudence, telles que les bases de données Légifrance tenues par la Direction de l'information légale et administrative (DILA), les bases de données de jurisprudence administrative tenues par le Conseil d'Etat, les décisions du Conseil constitutionnel ou encore la jurisprudence de la Cour de justice de l'Union européenne et de la Cour européenne des droits de l'homme. Cela permettrait notamment de disposer d'un chaînage complet des décisions permettant directement de consulter, par exemple, une décision de la Cour de justice de l'Union européenne rendue sur question préjudicielle d'une juridiction de l'ordre judiciaire.

L'exploitation de l'ensemble de ces bases par un outil d'IA permettrait ainsi de mettre ces données en relation et d'améliorer l'exploitation de la jurisprudence du fond ainsi que des travaux préparatoires, en lien avec la législation et la réglementation. En cela, elle constitue un gain significatif de qualité en fournissant aux magistrats de la Cour, ou des juridictions du fond, des références de jurisprudence plus pertinentes et faciliterait l'exploitation des travaux préparatoires antérieurs.

Cette exploitation peut prendre plusieurs formes cumulatives ou alternatives selon les technologies mobilisées qui permettront de moduler la saisie en entrée, la présentation des résultats en sortie et la base interrogée. Dans un contexte *d'open data* et dans le respect de notre système de droit continental, il apparaît essentiel de hiérarchiser les résultats, notamment en prenant en considération le niveau de juridiction, la formation de jugement et le niveau de publication des arrêts de la Cour de cassation.

➤ *Evaluation globale du cas d'usage :*

La mise en œuvre de ce projet peut cependant prendre trois dimensions qui présentent une complexité croissante.

La création d'un moteur de recherche sémantique, de faible complexité mais également d'intérêt moindre doit en effet être envisagée par comparaison à un système de RAG (*Retrieval Augmented Generation*), plus complexe et coûteux mais également plus flexible et riche de potentiel pour les utilisateurs.

Quant à la création d'un outil conversationnel qui assurerait une entrée en langage naturel et une sortie de même nature, en ce qu'elle nécessiterait un grand modèle de langage à l'état de l'art, elle ajoute, par rapport au RAG, un aléa supplémentaire et accroît le coût d'exploitation associé.

Ces cas d'usage ont en commun de pouvoir s'appuyer sur des données intègres, ce qui aura un impact sur le respect du RGPD, et sur des données de qualité qui nécessiteront néanmoins une annotation humaine.

Par ailleurs, quelles que soient ses modalités, un tel outil n'est pas dénué d'impact sur les droits fondamentaux dans la mesure où les résultats de la recherche peuvent avoir une influence sur la décision finale.

§ 1 – Création d'un moteur de recherche sémantique

Les moteurs de recherche tels qu'ils existent actuellement pour Jurica, Jurinet et Judilibre s'appuient sur une interface permettant de combiner champs de métadonnées et champs libres accessibles à la saisie des utilisateurs par mots clés ou expressions combinant ou non des opérateurs booléens. Les résultats en « sortie » (*output*) sont triés hiérarchiquement, par ordre de pertinence ou chronologique ou en combinant ces deux paramètres.

Il peut être envisagé d'améliorer le fonctionnement de ce type de moteurs de recherche en y embarquant des modèles de langage permettant de mieux appréhender le contexte des mots clés ou expressions saisies et intégrant une proximité lexicale si l'utilisateur le souhaite.

Le gain d'efficacité qui se traduit par un gain de temps pour les magistrats de chambre induit par ce moteur de recherche tient à la capacité de l'outil à réaliser une recherche unique au sein des trois bases de données internes et des bases de données publiques susceptibles d'y être associées. Le temps d'exploitation des résultats fournis par cet outil est cependant de nature à diminuer le gain d'efficacité du système par rapport à une technologie de type RAG.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	1
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	3
	Ressources métier	2
	Autres ressources (Designer, Devops)	1

	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – RAG (Retrieval Augmented Generation)

L'implémentation d'un RAG au sein du système de recherche documentaire permettrait de jouer sur les résultats en sortie, qui seraient présentés sous une forme littérale s'appuyant sur une IA générative, à la différence des résultats par liste présentés par les moteurs de recherche classiques.

La limitation de la génération au contexte contenu au sein des bases sans possibilité pour le LLM de se détacher de ces données est un facteur de limitation des erreurs ou hallucinations qui sont fréquemment associées à ces modèles.

L'utilisation de ce type d'outil supposerait cependant de sourcer précisément chaque citation générée par le SIA. Par ailleurs, un réglage de la « température » du modèle pourrait être prévu afin de diminuer la variabilité des réponses générées en réponse à une même requête.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	1

§ 3 – Outil conversationnel

Les membres de la mission s'accordent à reconnaître le potentiel qui pourrait être associé à des requêtes saisies en langage naturel.

Ce paramètre qui viendrait s'agréger au système de RAG « jouerait » ainsi sur le format de la donnée d'entrée comme le RAG « jouait » sur celle de sortie.

L'intérêt de ce modèle conversationnel, dont les caractéristiques pourraient être proches des modèles grand public qui ont été diffusées depuis la fin de l'année 2022⁷⁷, réside dans la souplesse de saisie des termes de la recherche et la facilité d'appréhension de l'outil.

Les faiblesses relevées de ce même type de SIA sont en revanche généralement liées à la faiblesse des algorithmes de recherche qui tient pour les modèles grand public à son caractère généraliste. Il est permis d'espérer que le conditionnement de l'algorithme en RAG et son affinage sur des données juridiques permettraient de dépasser cette limite.

Compte tenu des risques liés à la génération de contenus, il est par ailleurs envisageable de prévoir que la donnée en sortie se limite à présenter des listes de résultats, c'est-à-dire à renvoyer vers du contenu existant. Ces résultats devront prendre en compte la hiérarchisation de la jurisprudence.

L'utilisation d'une couche d'IA générative en sortie n'est pas neutre en tout état de cause sur la pertinence des résultats qui seront générés ni sur la puissance de calcul à mobiliser.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

77 Tel que Claude, ChatGPT, Perplexity, DeepSeek,

Section 2 Exploitation des bases de données pour un usage public.

Les options présentées ci-dessus pour les outils internes d'exploitation des bases de données sont également applicables aux outils pouvant être mis à disposition du public par la Cour de cassation dans le cadre de sa mission de diffusion de la jurisprudence, dans la continuité du moteur de recherche Judilibre accessible sur le site internet de la Cour.

Ce moteur de recherche pourrait ainsi évoluer pour intégrer d'autres bases de données publiques (Légifrance, autres juridictions...).

Les outils d'IA présentés ci-dessus (recherche sémantique, RAG, outils conversationnels) pourraient également être déployés dans les mêmes conditions pour un usage public.

Le groupe de travail constate, en ce qui concerne ces outils publics, que les différentes institutions responsables de ces bases de données travaillent actuellement de manière isolée à la mise en valeur de leurs fonds sans mise en relation de ces bases ni mutualisation des moyens déployés.

Ces travaux répondent cependant à un objectif commun de transparence et d'accessibilité du droit.

L'intégration de ces bases au sein d'un outil public commun, de manière similaire à ce qui a pu être effectuée par des acteurs privés, permettrait d'améliorer l'accès au droit et à la jurisprudence.

Bien que cet outil puisse être développé par la Cour de cassation seule, s'agissant de bases publiques, un partenariat dans le cadre de consortiums au niveau national (par exemple avec le Conseil d'Etat, la DILA, le Conseil constitutionnel) ou au niveau européen serait certainement préférable.

Ces consortiums nationaux ou européens pourraient ainsi avoir pour ambition de créer des modèles ou outils communs accessibles à tous et permettant l'exploitation des bases de données publiques en s'appuyant entre autres sur la richesse des métadonnées dont ils disposent et les compétences des acteurs internes.

➤ *Evaluation globale du cas d'usage :*

L'évaluation de la faisabilité de ce cas d'usage et du gain d'efficacité attendue est quasiment identique à la précédente en ce qu'elle reposerait sur les mêmes modalités et les mêmes technologies.

Toutefois, l'accessibilité des outils développés à des utilisateurs externes à la Cour de cassation, même limités aux professionnels du droit, augmente d'autant le coût d'exploitation des SIA en production.

La puissance de calcul ou l'entretien d'un modèle sollicité par plusieurs milliers ou millions de personnes n'est en effet en rien comparable au coût d'exploitation de ces mêmes outils limités aux usages internes de la Cour de cassation.

S'agissant du RGPD, les risques sont limités compte tenu de l'utilisation de bases publiques essentiellement pseudonymisées.

§ 1 – Création d'un moteur de recherche sémantique

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	1
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	3
	Ressources métier	2
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – RAG

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2

	Ressources métier	2
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	0

§ 3 – Outil conversationnel

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	0

Section 3 Exploitation des travaux parlementaires

La recherche de l'intention du législateur par le magistrat dans l'interprétation des textes représente une charge de travail importante. Ce travail est rendu fastidieux et chronophage par la multiplicité des documents à analyser (exposés des motifs des textes ou des amendements, études d'impact, rapports, débats en commission des lois ou en séance...) et nécessite souvent d'identifier une phrase unique au sein d'un long compte-rendu de débats en séance ou en commission. La spécificité de ces documents et de ce type de recherches ne permet pas, selon le groupe de travail, d'intégrer leur exploitation au sein des outils de recherche documentaire présentés ci-dessus.

L'initiative et la création d'un tel outil relèvent plus naturellement des missions du Parlement dans le cadre de la diffusion de leurs travaux. Sa réalisation serait, de surcroît, plus performante puisqu'il dispose de fonds propres et de métadonnées.

Toutefois, une association de la Cour de cassation, du Conseil constitutionnel et du Conseil d'Etat, qui ont également relevé l'intérêt de ce cas d'usage, constituerait un atout compte tenu de leur expertise juridique et technique.

Les outils de recherche sur les travaux parlementaires, de manière identique aux outils portant sur l'exploitation des bases de données à usage interne ou public, pourraient reposer sur un moteur de recherche sémantique, des RAG ou des modèles conversationnels.

L'évaluation proposée ci-dessous s'appuie sur le postulat d'un accès de la Cour aux bases de données du parlement.

➤ *Evaluation du cas d'usage :*

L'exploitation des travaux parlementaires par un SIA présente un grand intérêt en termes de gain d'efficacité pour les magistrats et juristes, cette opération étant particulièrement chronophage.

Le projet pourrait en outre s'appuyer à la fois sur des données libres et de qualité obéissant à des règles de classification précises.

Le seul point de vigilance tient donc au coût d'exploitation du modèle selon le champ des utilisateurs envisagés.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	3
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	1

Section 4 – Veilles automatiques

Les services supports des chambres au sein de la Cour de cassation réalisent de manière périodique des veilles de jurisprudence, veilles doctrinales et autres veilles médiatiques.

Il en va ainsi du service des relations internationales (SRI), des différents bureaux du SDER et du service de la communication de la première présidence.

Ces veilles sont activement consultées par les conseillers et avocats généraux et sont des outils précieux qui peuvent être réutilisés à l'occasion des travaux préparatoires.

§ 1 – Veille automatique européenne s'appuyant sur la jurisprudence de la Cour européenne des droits de l'homme (CEDH) et la Cour de justice de l'Union européenne (CJUE)

Le SDER et le SRI maintiennent une veille intellectuelle de l'actualité jurisprudentielle et législative européenne. Ces veilles sont notamment réalisées à partir des sites internet de la CEDH et de la CJUE.

Une extraction automatique des contenus pertinents par un SIA et une mise en forme par IA générative avant transmission aux membres de la Cour de cassation permettrait ainsi d'automatiser ces veilles.

➤ *Evaluation du cas d'usage :*

L'intérêt de ce cas d'usage et les gains de qualité ou d'efficacité pouvant résulter de la mise en place d'un SIA sont à apprécier par rapport aux nombreux outils de qualité préexistants et disponibles sur les sites Curia et Hudoc.

En outre, si la faisabilité technique d'un tel projet ne semble pas revêtir une particulière complexité, la réalisation de celui-ci sera cependant intégralement subordonnée à l'autorisation des juridictions européennes d'accéder de manière automatisée à leurs bases de données.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	1
	Gain d'efficacité	1
Critères Techniques	Disponibilité de la donnée	0
	Preuve de faisabilité technique	2

Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – Veille doctrinale avec synthèses et chaînage de la jurisprudence de la Cour

L'éclatement des sources doctrinales au sein des différents sites des éditeurs juridiques représente une charge de travail considérable pour les conseillers, avocats généraux et membres du SDER dans leurs travaux de recherche et la réalisation des veilles doctrinales qui permettent de compiler par chambre les études et commentaires d'arrêts.

La diversité des moteurs de recherche de chaque site, des classifications opérées par chaque éditeur et des identifiants à renseigner sur chaque site contribue au manque d'ergonomie du système actuel.

La création d'un outil de recherche centralisé reposant sur l'IA constituerait donc un véritable gain de temps. Cet outil de recherche, qui pourrait reposer sur les technologies visées ci-dessus au paragraphe 1^{er} (recherche par mots-clés sémantique, RAG et outil conversationnel), pourrait permettre d'identifier les références doctrinales au sein des différentes bases, à l'image de ce que propose aujourd'hui le site « Le Doctrinal », mais permettrait également d'en récupérer directement les contenus et d'exploiter les métadonnées associées à ces documents.

➤ Evaluation du cas d'usage :

L'intérêt qualitatif et quantitatif attaché à la réalisation de veilles doctrinales automatiques apparaît évident et constituerait de surcroît un enrichissement précieux des bases de données jurisprudentielles.

Ce cas d'usage est cependant subordonné à la possibilité pour la Cour d'accéder directement de manière automatisée aux fonds documentaires des éditeurs.

Par ailleurs, outre cette condition nécessaire et préalable, l'analyse technique du cas d'usage met en évidence une certaine complexité du projet et de la détermination de son coût d'exploitation, qui est liée à l'interrogation quotidienne ou du moins périodique de vastes fonds documentaires par un SIA reposant sur un grand modèle de langage.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2

Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	3
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	0
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	2
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	0

§ 3 – Veille médiatique

Les veilles médiatiques réalisées quotidiennement par le service de communication de la Cour et de manière périodique par les services supports de la Cour de cassation constituent un élément d'information de la perception de l'activité judiciaire dans la société et des sujets qui y sont associés.

Ces veilles sont actuellement réalisées humainement en s'appuyant sur les sites des principaux médias et pourraient être extraites à la source et mises en forme par un SIA.

➤ *Evaluation du cas d'usage :*

La réalisation de veilles médiatiques automatiques pose les mêmes questions de disponibilité de la donnée que les veilles doctrinales.

En outre, l'intérêt associé à ce projet est moindre en l'état de solutions commerciales généralistes appropriées aux objectifs poursuivis.

Il n'apparaît donc pas intéressant, financièrement comme techniquement, de concevoir un modèle spécifique interne.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	1

	Gain d'efficacité	1
Critères Techniques	Disponibilité de la donnée	0
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

Section 5 – Identification des rapprochements ou divergences de jurisprudence

La recherche de précédents constitue une phase incontournable des travaux préparatoires du conseiller et de l'avocat général saisis d'un pourvoi. Elle définit l'état de l'interprétation d'un texte ou d'un principe normatif à travers le temps et contribue à la cohérence et à la sécurité juridique qui s'attachent à la jurisprudence de la Cour de cassation.

Ce mode de travail n'est cependant pas propre à l'activité juridictionnelle de la Cour mais correspond à tout travail d'étude et de recherche du magistrat ou juriste saisi d'une question de droit.

Au sein de la Cour de cassation, la recherche de précédents ou de rapprochements de jurisprudence mobilise fortement les services du SDER dans la rédaction d'études, panoramas de jurisprudence et autres notes à destination des juridictions du fond ou des magistrats de la Cour de cassation ainsi que pour l'enrichissement des arrêts publiés au Bulletin. Ces mentions, accessibles au grand public, contribuent à la lisibilité d'une jurisprudence dans le temps et peuvent mentionner expressément la continuité ou la rupture d'un fil jurisprudentiel, dans le cas d'un revirement (voir également, en ce qui concerne les rapprochements avec les arrêts de la Cour de cassation, le paragraphe 3 de section).

La Cour possède déjà une expertise mobilisable dans la génération automatique de rapprochements de décisions de la Cour de cassation, comme cela a été rappelé au sein de la Section 3 du Chapitre 1^{er} de la deuxième partie (« Le projet Divergences »), à laquelle il y a lieu de renvoyer pour de plus amples développements.

Les cas d'usage reposant sur cette technologie pourraient être étendus à différents champs mais obéissent à une logique commune : l'identification par IA de la portée juridique des décisions qui lui sont soumises, l'extraction des éléments composant la question de droit et la recherche au sein d'un corpus d'éléments similaires. Les résultats seraient alors

proposés à l'utilisateur sous forme d'une liste de décisions assorties ou non de résumés ou chaînes de mots clés générés par IA.

Le travail à fournir pour franchir chacune des étapes précitées dépendra de la qualité et de l'homogénéité rédactionnelle des décisions comparées.

Il peut d'ores et déjà être intéressant d'envisager les rapprochements suivants.

§ 1 – Rapprochements jurisprudentiels ou divergences en droit interne et européen

➤ Evaluation du cas d'usage :

Ce cas d'usage présente les mêmes difficultés techniques que celles liées à la création de veilles automatiques européennes.

L'autorisation d'accès automatisé à la donnée européenne constitue en effet la condition nécessaire à la poursuite de ce projet qui pourrait s'appuyer sur les ressources de la Cour.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	0
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – Identification de rapprochements ou divergences de jurisprudence entre les juridictions du fond

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1

Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 3 – Identification de rapprochements ou de divergences entre les décisions des juridictions du fond et la jurisprudence de la Cour de cassation

➤ *Evaluation des cas d'usage § 2 – 3 :*

L'hétérogénéité de mise en forme des décisions des juridictions du fond et des contentieux qu'elles connaissent constitue la principale difficulté technique de ces deux projets qui supposeront en conséquent un fort investissement du métier et des ressources humaines en *data science* pour l'entraînement et la mise en production d'un modèle de langage à l'état de l'art.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2

	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 4 – Identification de rapprochements ou divergences de jurisprudence au sein de la Cour de cassation

➤ *Evaluation du cas d'usage :*

Les difficultés relevées dans les précédents cas d'usages sont ici en partie lissées par la mise en forme homogène des décisions de la Cour de cassation et l'expérience acquise par ses équipes dans le cadre du projet « Divergences ».

Le cas d'usage d'identification de rapprochements ou divergences de jurisprudence au sein de la Cour de cassation apparaît ainsi réalisable à moyen ou court terme grâce à un investissement raisonnable des équipes métiers et techniques dans la poursuite du projet.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	3
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	3
	Coût d'exploitation	3

Section 6 – Analyse de la jurisprudence du fond afin de déterminer des tendances et positions dominantes sur des questions précises

§ 1 – Tendances et positions dominantes au sein des juridictions du fond dans les contentieux chiffrés

Les conséquences attachées à la publication massive des décisions de première et seconde instance ont fait l'objet d'une analyse approfondie dans le rapport de la Cour de cassation sur la diffusion des données décisionnelles et la jurisprudence⁷⁸.

La disponibilité de ces décisions représente une opportunité intéressante pour la Cour de cassation d'accroître sa connaissance et d'affiner son analyse de la jurisprudence du fond en particulier sur les contentieux sur lesquels elle peut adapter l'intensité de son contrôle.

Cette exploitation de milliers de décisions, qui ne peut raisonnablement reposer sur une analyse humaine, est rendue possible par l'utilisation d'outils algorithmiques.

De plus, certains éditeurs juridiques proposent des outils d'analyse appliqués aux contentieux chiffrés tels que la réparation des préjudices corporels, les assurances dommages-ouvrages ou le calcul de prestations compensatoires. Ces outils, qui favorisent certes les modes alternatifs de règlement des litiges, participent néanmoins d'un risque ou d'une tendance à la barémisation de ces contentieux et questionnent la valeur de ces preuves algorithmiques, ainsi que les effets de ces preuves sur la conviction du juge⁷⁹.

Il serait utile que la Cour de cassation puisse se doter d'une expertise objective, impartiale et neutre sur ces questions, pour elle-même comme pour les juridictions du fond.

➤ *Evaluation du cas d'usage :*

La très grande variété des contentieux combinée à l'hétérogénéité des décisions des juridictions du fond nécessiterait pour ce projet de dédier un ou plusieurs experts juridiques à temps complet afin d'annoter de manière intensive un grand nombre de décisions.

L'évaluation automatisée des contentieux chiffrés générée par cet outil est de nature à favoriser les modes alternatifs de règlement des litiges mais également à influencer sur la décision du juge. En conséquence, une vigilance particulière doit être mise en œuvre pour s'assurer que le juge conserve de manière effective toute latitude pour, le cas échéant, faire le choix de s'écarter de ces barèmes.

Critères Ethiques	Conciliation avec les droits fondamentaux	1
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3

⁷⁸ L. Cadiet, C. Chainais et J.-M. Sommer (dir.), S. Jobert et E. Jond-Necand (rapp.), *La diffusion des données décisionnelles et la jurisprudence*, Rapport remis à la première présidente de la Cour de cassation et au procureur général près la Cour de cassation - juin 2022, disponible en ligne sur le site de la Cour

⁷⁹ Etienne Vergès et Geraldine Vial, *Le juge face à la boîte noire : l'intelligence artificielle au tribunal ; La modélisation des décisions de justice à partir des faits ; Justice prédictive - Comment l'IA imite les raisonnements des juges à partir d'une grille de critères*, Recueil Dalloz 2022 p.1920, s.

	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	0
Critères Techniques	Disponibilité de la donnée	1
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	0
	Autres ressources (Designer, Devops)	1
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – Identification de questions nouvelles ou de litiges sériels – soutien aux missions de l'observatoire des litiges judiciaires

La détection des questions de droit nouvelles ou des contentieux émergents participe des missions dévolues à l'observatoire des litiges lancé par la Cour de cassation en juin 2022 suite à une recommandation du rapport « Cour de cassation 2030 »⁸⁰.

Le rapport relevait ainsi « qu'au sein de l'ordre judiciaire, le grand nombre des juridictions du fond et la multitude des affaires qu'elles traitent ne permettent pas facilement le repérage des évolutions qui s'y développent ainsi que les dossiers justifiant une attention particulière. Pourtant, une telle vision permettrait de se doter d'une capacité d'anticipation partagée entre la Cour de cassation et les cours d'appel, notamment pour rapprocher certains dossiers dispersés dans différentes juridictions mais posant un problème similaire, afin de mettre en œuvre des processus consistant à se concentrer provisoirement sur des dossiers pilotes puis, une fois dégagée une orientation satisfaisante, de traiter efficacement tous ceux relevant de la même problématique.

De même, il serait ainsi possible de connaître très rapidement l'urgence particulière marquant certains dossiers, afin de leur réserver un traitement adapté en première instance, en appel puis devant la Cour de cassation. »

L'identification, en amont et au sein d'une très large masse de données, des contentieux émergents dépasse donc les seuls besoins internes de la Cour de cassation et relève d'un enjeu de bonne administration de la justice, le SIA envisagé pouvant bénéficier aux juridictions du fond. En effet, l'exploration de ce champ de recherche permettrait

⁸⁰ Pour aller plus loin, [le rapport de la commission de réflexion sur la cour de cassation 2030 est disponible sur le site internet de la Cour](#)

également à la Cour de se saisir des questions émergentes afin de prévenir les éventuelles divergences de jurisprudence qui pourraient naître au sein des juridictions du fond.

A ce jour, l'observatoire des litiges s'appuie sur les ressources juridiques du SDER et sur son réseau de référents au sein des cours d'appel. Ce traitement humain des décisions ne peut évidemment qu'être parcellaire compte tenu de la charge de travail induite par cette tâche qui s'ajoute à leurs autres activités.

L'entraînement d'un SIA s'appuyant sur les décisions diffusées en *open data* des juridictions du fond mais également sur les bases internes de la Cour de cassation pourrait ainsi représenter une assistance précieuse dans la tâche d'indentification des questions nouvelles.

➤ *Evaluation du cas d'usage :*

Le gain de qualité ou d'efficacité associé à la création d'un SIA en capacité d'identifier des questions nouvelles ou de litiges sériels au sein des juridictions du fond est significatif.

Il s'agit d'un projet d'envergure mais qui peut s'appuyer sur les connaissances acquises par la Cour en matière de rapprochements automatiques à l'occasion du projet « Divergences ».

Il supposera toutefois un fort investissement des équipes de *data science* comme du métier.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 3 – Outil général d'analyse de la jurisprudence du fond, notamment massive

En complément des deux cas d'usage évoqués précédemment et dans ce même contexte d'augmentation du volume des décisions judiciaires accessibles en *open data*, il pourrait être envisagé de concevoir un outil d'assistance du magistrat à l'identification des solutions juridiques ou des critères retenus dans les décisions des juridictions du fond. Cette opération relève en effet des missions de l'observatoire des litiges judiciaires et contribue à accroître la lisibilité de la jurisprudence du fond par la Cour de cassation.

Par ailleurs, et dans le cadre de son activité juridictionnelle, la Cour de cassation est saisie de contentieux de masse impliquant parfois plusieurs centaines de décisions des juridictions de première instance et d'appel dont l'exploitation humaine se révèle particulièrement chronophage.

L'identification par IA des critères retenus par les juridictions du fond au soutien de leur motivation ou tout simplement de la solution retenue dans leur dispositif pourrait ainsi permettre une première exploitation d'un jeu important de données. Cette exploitation serait de nature à favoriser le dialogue des juges, horizontal et vertical puisque chaque juge pourra élaborer son raisonnement en connaissance des positions des autres sur une même question nouvelle.

Ce cas d'usage suppose toutefois de s'appuyer sur un modèle d'IA général puissant et polyvalent compte tenu de la diversité des contentieux susceptibles de se présenter et de l'impossibilité de développer des modèles dédiés affinés sur l'ensemble de ces matières.

Deux options techniques pourraient donc être envisagées. La première pourrait reposer sur un modèle commercial en ligne alimenté par des jeux de données publiques extraits de Judilibre. La seconde supposerait d'héberger localement un grand modèle général afin de pouvoir notamment l'affiner sur du contenu juridique.

➤ Evaluation du cas d'usage :

L'analyse d'une grande masse de décisions du fond, de nature variée, suppose de reposer sur un grand modèle de langage, polyvalent mais présentant également un fort coût d'exploitation même si son usage devait rester limité au sein de la Cour de cassation.

Ce cas d'usage présenterait toutefois un fort gain fonctionnel et faciliterait l'exploitation de vastes jeux de données jurisprudentielles.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	2
	Mise en conformité réglementation données personnelles	3
Critères Fonctionnels	Gain de qualité	3

	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

Chapitre 4 - Aide à la rédaction

La mission IA a identifié quatre types d'usages différents de l'intelligence artificielle pour l'assistance à la rédaction des membres de la Cour de cassation.

Ces cas d'usage reposent en premier lieu sur la capacité de l'IA à générer des résumés ou à identifier au sein des écritures des parties les éléments nécessaires aux travaux préparatoires des conseillers rapporteurs et avocats généraux.

Section 1 – Aide à la rédaction des rapports, avis et décisions

L'assistance à la rédaction ne peut avoir pour effet d'amoindrir le contrôle humain de la rédaction des décisions rendues par la Cour de cassation.

Elle peut en revanche accompagner le processus d'écriture en permettant à l'utilisateur d'identifier et d'intégrer plus rapidement des éléments provenant d'affaires traitées précédemment par la Cour de cassation. Il pourrait s'agir concrètement d'analyser la similarité avec des affaires précédentes pour identifier des blocs de rédaction pertinents dans les précédents rapports, avis et décisions de la Cour de cassation. Cet outil d'aide à la rédaction s'appuierait ainsi sur des techniques proches ou identiques à celles utilisées pour l'aide à la recherche documentaire.

Ces étapes de rédaction sont actuellement réalisées manuellement par les avocats généraux et conseillers rapporteurs. La création d'un SIA permettrait ainsi d'accroître leur efficacité et de faciliter l'exhaustivité de la recherche documentaire.

La compréhension retenue par la mission IA d'un SIA d'assistance à la rédaction est donc avant tout celle d'une assistance au rassemblement de la documentation pertinente, intégrée à l'outil de rédaction, et non une génération automatisée globale et de qualité incertaine.

➤ *Evaluation du cas d'usage :*

Le gain fonctionnel attendu d'une assistance à la rédaction des magistrats par un SIA est significatif. Ce résultat peut être atteint grâce à l'adaptation de grands modèles de langage, performants mais nécessairement plus coûteux, qui permettront d'identifier des similarités entre les affaires pendantes et les décisions rendues antérieurement par la Cour de cassation.

Il s'agit donc de la prolongation logique de plusieurs cas d'usage développés au sein de l'aide à la recherche documentaire. Une particulière attention devra cependant être apportée à l'impact de ce SIA sur le plan éthique et juridique, afin notamment de garantir la maîtrise humaine. Pour cela, comme pour l'entraînement du modèle, un investissement significatif du métier et des équipes techniques sera requis.

Critères Ethiques	Conciliation avec les droits fondamentaux	1
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	1
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	1
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

Section 2 – Outil d'aide à la rédaction pour les litiges sériels

La Cour de cassation est régulièrement saisie de séries de dossiers présentant des connexités matérielles ou intellectuelles. La bonne administration de la justice justifie que ces affaires soient appréhendées globalement par les chambres compétentes.

Pratiquement, cela suppose, pour le conseiller et l'avocat général saisis du dossier, d'exploiter de multiples jeux d'écritures parfois strictement identiques à une ou deux phrases près et mettant en jeu quelques dizaines jusqu'à quelques centaines de parties.

Une analyse par IA de ces écritures afin de déterminer les prétentions communes et parties en présence, mais également les divergences, notamment dans la rédaction des moyens, pourrait à ce titre se révéler précieuse.

➤ *Evaluation du cas d'usage :*

En dépit d'un intérêt certain pour la Cour de cassation, qui est saisie, chaque année, d'environ 650 séries, la grande variété des contentieux actuels et futurs interroge la capacité technique d'un SIA à effectuer une comparaison, qui dépasserait le seul stade formel, des mémoires ampliatifs.

La comparaison formelle des moyens, des arrêts attaqués, la génération de motivations identiques en réponse aux questions identifiées comme identiques sont techniquement envisageables dans les séries présentées comme homogènes.

Dans d'autres séries cependant, lorsque les mémoires présentent entre eux une connexité intellectuelle, ou même factuelle, mais se rattachent à des décisions différentes, mettant en cause des parties pas toutes identiques et avec des mémoires produits par des avocats différents, l'exploitation par un SIA se révèle plus compliquée. Dans ces cas, la création d'un outil *ad hoc* permettrait de surmonter la difficulté à l'instar de ce qui a pu être réalisé dans le cadre du Dieselgate en Allemagne.

Le recours à un SIA général s'appuyant sur un vaste LLM pour traiter une hétérogénéité de séries ne semble donc pas couvrir entièrement les besoins de la Cour même s'il permettrait d'appréhender les séries les plus évidentes. Un des leviers utilisables pour compenser cette faiblesse pourrait être la mobilisation à plein temps d'experts juridiques au sein de la Cour de cassation.

Les gains associés sont toutefois très significatifs car le traitement des séries est très chronophage et les risques d'erreurs matérielles réels.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	1
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	2
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	1

Section 3 – Uniformisation et vérification de la rédaction (formulaires et guides de rédaction, normes de saisie, orthographe, syntaxe, qualité de la langue, complétude).

§ 1 – Intégration des formulaires et guides de rédaction

Les recommandations des formulaires et guides pour la rédaction des décisions et travaux préparatoires pourraient faire l'objet d'applications d'IA pour être intégrées en tant qu'outil d'aide à la rédaction.

Par exemple, le guide actualisé des règles relatives à la structure et à la rédaction des arrêts en matière civile détaille au travers d'un document d'une cinquantaine de pages les usages observés pour la rédaction des arrêts civils ainsi que l'ordre de présentation des différents moyens et l'articulation des réponses apportées.

➤ *Evaluation du cas d'usage :*

Le projet peut s'appuyer sur une documentation précise et exhaustive interne à la Cour de cassation. La très grande technicité associée à la procédure de cassation et à l'articulation des différentes règles présentes dans les guides de rédaction supposerait cependant un investissement soutenu d'un magistrat, conseiller ou avocat général, familier de cette technique en lien constant avec une équipe de *data science*.

Pour sa mise en œuvre, le recours à une IA générative s'appuyant sur un LLM déployé au niveau de la Cour est de nature à entraîner un coût d'exploitation modéré mais réel.

Compte tenu de la proximité du SIA envisagé avec le processus décisionnel (génération de blocs de la décision et risques de biais d'automatisation), il existe un impact faible mais résiduel sur les droits fondamentaux.

De même, l'intégration dans le SIA de blocs intègres, extraits des mémoires ampliatifs ou du corps du projet de décision, impose de veiller au respect du RGPD sauf à ce que les magistrats de la Cour interagissent via des « prompts » abstraits avec le SIA.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – Vérification automatique des normes de saisie

Les normes de saisie en vigueur à la Cour de cassation sont exposées dans un document qui comprend les usages à respecter, notamment pour se référer à des textes de loi, de réglementation, de la jurisprudence et de la doctrine, ainsi que des règles d'écriture et de présentation.

Le groupe de travail constate que le respect de ces normes de saisie, établies en 2012, participe à une logique d'harmonisation de la forme qui facilite l'exploitation automatisée des documents par intelligence artificielle. Il pourrait à cet égard être envisagé, à l'occasion de la mise à jour de ces normes de saisie, qu'il soit proposé aux juridictions du fond de les suivre également, afin de faciliter la valorisation de la jurisprudence dans le cadre du projet d'*open data* des décisions de justice.

Le respect de ces normes de saisie entraîne actuellement une charge de travail non négligeable pour les magistrats et le greffe de la Cour de cassation, alors même que leur documentation précise les prédispose tout particulièrement à l'usage de l'IA.

Cette vérification pourrait inclure celle du contenu des textes ou arrêts auxquels il est fait référence d'où leur impact sur la protection des données personnelles.

➤ *Evaluation du cas d'usage :*

La mission IA a pu identifier l'existence de solutions techniques commerciales proches du cas d'usage, ce qui constitue un bon indice de faisabilité technique du projet, sous réserve d'un investissement raisonnable du métier afin de permettre la retranscription technique des normes de saisie juridiques définies par la Cour de cassation.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	2
	Ressources métier	2
	Autres ressources (Designer, Devops)	2

	Puissance et temps de calcul (entrainement)	3
	Coût d'exploitation	2

§ 3 - Harmonisation des styles rédactionnels

D'autres SIA pourraient également être envisagés en s'appuyant sur la reconnaissance des styles rédactionnels de façon à permettre au rapporteur d'adopter plus facilement les spécificités stylistiques propres à sa section ou à sa chambre.

➤ *Evaluation du cas d'usage :*

La complexité du projet est liée à la diversité des styles rédactionnels en vigueur au sein des différentes chambres de la Cour de cassation ou encore des sections qui les composent.

Il supposerait en conséquence de s'appuyer sur un gros modèle de langage, polyvalent mais dont l'affinage serait difficilement envisageable sur des contenus juridiques en raison de sa dimension.

Les résultats associés à ce cas d'usage demeurent donc incertains et nécessiteraient en tout état de cause un très fort investissement humain, métier et technique.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	2
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entrainement)	1
	Coût d'exploitation	2

§ 4 – Vérification des oublis de moyens, de la cohérence des textes visés et des moyens

La vérification par le magistrat de l'exhaustivité des moyens repris au sein de ses projets de décisions ainsi que de ses travaux préparatoires est un travail nécessaire et chronophage qui pourrait être assisté par IA.

➤ *Evaluation du cas d'usage :*

La mission IA constate la grande sensibilité de ce cas d'usage en raison de sa proximité immédiate avec la fonction juridictionnelle des magistrats de la Cour et du risque d'impact sur la décision qui résulterait d'une erreur d'analyse du SIA mis en place. Ainsi, le signalement par le SIA d'un oubli de moyen conduira le magistrat à effectuer une vérification au sein des mémoires. A l'inverse, l'absence de signalement, justifié ou erroné, tendra à conforter le magistrat dans l'idée qu'il n'y a pas eu d'omission de moyen ce qui est une illustration d'un biais d'automatisation.

Il s'agit par ailleurs d'un cas d'usage complexe et coûteux sur le plan technique au regard des technologies à mobiliser qui s'appuieraient sur des LLM à l'état de l'art afin d'appréhender la variété des contentieux et des documents traités.

Critères Ethiques	Conciliation avec les droits fondamentaux	1
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	1
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	3
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	1
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	1
	Coût d'exploitation	2

Section 4 – Génération automatique de résumés

§ 1 – Proposition de résumés automatiques pour la rédaction des travaux préparatoires et mise en valeur des faits, de la procédure, des moyens, de l'article 700 du code de procédure civile et des avis

L'analyse des écritures des parties nécessite un travail minutieux des conseillers et avocats généraux au stade des travaux préparatoires.

La mission IA a ainsi retenu que s'il ne pouvait en aucun cas être envisagé de procéder à une lecture parcellaire de ces écritures, la génération de résumés des faits constants et pertinents dans les décisions attaquées et des moyens soulevés par les parties pouvait, dans certaines hypothèses, être de nature à faciliter la première approche des magistrats lors de ce travail d'analyse.

Le groupe de travail a toutefois souligné les limites de cette technologie. La génération de résumés pourrait en effet aboutir à une analyse parcellaire des documents sans prise en compte du contexte dans lequel s'inscrivent les éléments extraits par l'IA et compilés et pouvant même donner lieu à des contresens. Dans ces éventualités, le travail de vérification de l'utilisateur pourrait s'avérer plus long que la lecture initiale qu'il aurait pu faire des écritures analysées.

Les auditions conduites par la mission IA ont confirmé l'existence de ces réserves techniques et permis d'imaginer des cas d'usage alternatifs aux limites inhérentes à la génération de résumés.

Ainsi, il pourrait être envisagé d'associer ou de substituer à la génération de résumés un SIA qui mettrait en évidence au sein des corpus juridiques et des écritures les passages qui apparaissent pertinents sous la forme de surlignages.

Cette solution technologique permettrait ainsi de garantir un niveau élevé de contrôle humain sur les résultats du SIA à partir des premiers éléments identifiés et mis en évidence dans le texte.

Il peut également être envisagé de ne résumer par IA que les éléments pertinents préalablement sélectionnés à la lecture par le magistrat⁸¹.

➤ *Evaluation du cas d'usage :*

La génération de résumés automatiques s'appuyant sur une exploitation des écritures des parties paraît devoir être abordée avec beaucoup de prudence au regard des enjeux juridiques et éthiques qui s'y attachent.

⁸¹ Voir IA générative et décision du juge, E. Legrand et M. Popa-Fabre, impact du numérique sur la justice, vo 2, oct. 2024, IERDJ

Par ailleurs, le manque de fiabilité de l'IA et la dégradation de la qualité qui en découlerait compte tenu des technologies existantes au jour du rapport conduisent le groupe de travail à écarter ce cas d'usage en l'état.

La solution alternative identifiée par le groupe de travail et reposant sur une sélection par le conseiller des éléments à synthétiser ou une mise en valeur dans le corps du texte des éléments recherchés semble en revanche poser de moindres difficultés d'ordre éthique comme technique.

Critères Ethiques	Conciliation avec les droits fondamentaux	1
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	1
	Mise en conformité réglementation données personnelles	1
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

§ 2 – Publication : proposition de résumé ou de sommaires pour publication au Bulletin ou à la Lettre de chambre, résumés de vulgarisation en vue d'une diffusion pour le grand public

En complément du cas d'usage précédent la génération de résumés extraits à partir de la Lettre des décisions de justice à des fins de vulgarisation ou d'enrichissement pourrait être également appréhendée.

➤ Evaluation du cas d'usage :

L'analyse du précédent cas d'usage est reproductible ici. Bien que la sensibilité du SIA soit moindre ici compte tenu de son application postérieurement au prononcé de la décision, le manque de précision et de fiabilité des systèmes d'IA générative conduit à exclure ce projet de génération automatique de résumés.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	1
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	0
	Gain d'efficacité	1
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	2
Critères Economiques	Ressources en data science	2
	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

Section 5 - Amélioration du titrage

Le titrage constitue un enrichissement des décisions publiées par la Cour de cassation au bulletin ou au rapport annuel. Cette opération, réalisée manuellement par des juristes du SDER, permet d'associer une chaîne de mots-clés, ou maillons, à un « sommaire » rédigé par le conseiller rapporteur, qui met en exergue la portée juridique de la décision publiée.

Cette opération chronophage est précieuse en ce qu'elle permet une meilleure classification des décisions publiées au sein des bases Jurinet et Judilibre et une mise en relation des décisions présentant les mêmes titrages entre elles.

➤ *Evaluation du cas d'usage :*

Les résultats intermédiaires du projet Divergences ont mis en évidence la capacité du SDER à développer à très court terme une interface de titrage semi-automatique des décisions publiées au bulletin ou au rapport. Cette technologie s'appuie sur de petits modèles de langage, peu énergivores grâce à la qualité des données qui ont pu être exploitées lors de son entraînement.

Cette fonctionnalité constituerait un gain de temps pour les agents du SDER ainsi qu'un gain potentiel de qualité en ce qu'elle permettrait d'optimiser la réutilisation des maillons de titrage préexistant et ainsi d'améliorer la classification des décisions au sein de la base Jurinet. Elle pourrait également être utilisée par les chambres dans la constitution des *feuilletons* internes.

Une phase de test, idéalement en double aveugle, devrait toutefois être prévue afin de vérifier que la qualité des titrages générés automatiquement ou semi-automatiquement n'est pas inférieure à la qualité actuelle des titrages créés humainement.

Critères Ethiques	Conciliation avec les droits fondamentaux	3
	Développement durable : frugalité des algorithmes	3
Critères Juridiques	Mise en conformité réglementation IA	3
	Mise en conformité réglementation données personnelles	3
Critères Fonctionnels	Gain de qualité	2
	Gain d'efficacité	2
Critères Techniques	Disponibilité de la donnée	3
	Preuve de faisabilité technique	3
Critères Economiques	Ressources en data science	3
	Ressources métier	2
	Autres ressources (Designer, Devops)	3
	Puissance et temps de calcul (entraînement)	3
	Coût d'exploitation	3

Chapitre 5 – Cas d'usage du greffe

Section 1 – Aide à la rédaction d'ordonnances de désistement

Le demandeur qui a introduit un pourvoi en cassation peut se désister de son pourvoi et mettre fin à l'affaire en cours.

Le greffe de la Cour prépare en moyenne 2 000 décisions de désistement par an.

Ce travail préparatoire suppose pour le greffier d'identifier le contexte procédural du pourvoi à l'occasion duquel est présentée une requête en désistement. Si les désistements dans des pourvois avec un seul demandeur ne posent guère de difficultés, les vérifications à opérer sont nécessairement plus longues lorsqu'il y a plusieurs demandeurs et/ou défendeurs au pourvoi.

En cas de complexité procédurale, le greffier sollicite les instructions préalables du président puis procède à la formalisation de la décision de désistement.

La préparation de ces décisions de désistement mobilise ainsi fortement le greffe de la Cour.

Par conséquent, la possibilité d'une assistance de l'IA sur cette tâche allègerait la charge de travail des greffiers. En outre, les délais de traitement des pourvois en seraient nécessairement réduits.

➤ *Evaluation du cas d'usage :*

La mise en œuvre de ce cas d'usage relève d'un système expert élaboré. S'il n'apparaît donc pas complexe techniquement, ni coûteux en ce qu'il ne repose pas sur une IA générative de dernière génération, sa mise en œuvre supposera cependant un investissement considérable d'un expert métier afin de retranscrire en système expert l'articulation de l'ensemble des règles juridiques applicables.

Critères Ethiques	Conciliation avec les droits fondamentaux	2
	Développement durable : frugalité des algorithmes	2
Critères Juridiques	Mise en conformité réglementation IA	1
	Mise en conformité réglementation données personnelles	2
Critères Fonctionnels	Gain de qualité	1
	Gain d'efficacité	3
Critères Techniques	Disponibilité de la donnée	2
	Preuve de faisabilité technique	1
Critères Economiques	Ressources en data science	2

	Ressources métier	1
	Autres ressources (Designer, Devops)	2
	Puissance et temps de calcul (entraînement)	2
	Coût d'exploitation	2

Section 2 – Aide à l'établissement par le greffe des certificats de non-pourvoi

Le certificat de non-pourvoi (CNP), délivré en matière civile par le greffe de la Cour de cassation, atteste qu'au jour de sa délivrance il n'y a pas de pourvoi formé ni de demande d'aide juridictionnelle enregistrée en vue de former un tel pourvoi.

Concrètement, le greffe est ainsi saisi, majoritairement par voie électronique, de demandes de CNP à raison de 20 à 30 demandes quotidiennes. Ces demandes ne sont assujetties à aucun formalisme légal.

La tâche de rédaction des CNP s'appuie en grande partie sur des trames mais suppose que le courrier de saisine contienne la copie de la décision attaquée et de sa signification. Or, il apparaît que de nombreux mails incomplets sont adressés au greffe, ce qui suscite une correspondance entre le greffe et les demandeurs.

Un SIA à même d'analyser les pièces jointes adressées électroniquement à la Cour et de générer des relances types en cas d'insuffisance des pièces produites permettrait ainsi d'alléger considérablement la charge de travail associée à cette activité du greffe.

Un cas d'usage similaire a été identifié par le Conseil d'Etat pour la délivrance de certificats de non-recours.

Par ailleurs, la détection par ce même SIA de termes du courrier permettant manifestement de réorienter le justiciable vers une autre autorité compétente permettrait également d'accélérer le processus (identification des termes « tribunal administratif » ou « cour administrative d'appel » par exemple). Il apparaît en effet que de nombreux courriers sont à tort adressés à la Cour de cassation.

➤ *Evaluation du cas d'usage :*

La mise en place d'un SIA permettrait ici de remédier à une communication par courriel sans utiliser le formulaire dédié disponible sur le site de la Cour de cassation.

L'analyse des pièces produites en pièces jointes, qui supposerait un procédé de *computer vision*, apparaît démesurément coûteuse et incertaine quant à la qualité des documents produits.

Critères Ethiques	Conciliation avec les droits fondamentaux	3	
	Développement durable : frugalité des algorithmes	0	
Critères Juridiques	Mise en conformité réglementation IA	3	
	Mise en conformité réglementation données personnelles	1	
Critères Fonctionnels	Gain de qualité	0	
	Gain d'efficacité	3	
Critères Techniques	Disponibilité de la donnée	1	
	Preuve de faisabilité technique	1	
Critères Economiques	Ressources en data science	1	
	Ressources métier	2	
	Autres ressources (Designer, Devops)	2	
	Puissance et temps de calcul (entraînement)	1	
	Coût d'exploitation	1	

CONCLUSION

A l'issue de ses travaux, le groupe a pu identifier des pistes susceptibles d'être priorisées pour le développement de l'IA à la Cour de cassation.

En premier lieu, certains cas d'usage relativement simples et ne posant pas de difficulté juridique ou éthique apparaissent en mesure d'offrir des gains fonctionnels intéressants, sans nécessiter d'investissements lourds.

Il s'agit en particulier du cas d'usage transversal de structuration et d'enrichissement des documents⁸², qui pourrait bénéficier à de nombreux autres cas d'usage, qu'ils soient relatifs à l'exploitation des écritures des parties⁸³, à l'aide à la recherche et l'exploitation des bases de données documentaires⁸⁴ ou encore à l'aide à la rédaction⁸⁵.

Ces cas d'usage trouveraient un complément et un prolongement très utiles dans un outil d'aide à la vérification des règles d'écritures et des normes de saisie en vigueur à la Cour de cassation⁸⁶.

Il peut également s'agir du cas d'usage de moteur de recherche sémantique dans les bases de données documentaires, s'il était préféré ne pas développer d'emblée des RAG ou outils conversationnels de recherche qui répondent aux mêmes finalités.

Dans le cas où tout ou partie de ces projets seraient retenus, ils pourraient faire l'objet de développements internes à intégrer au plan de charge des équipes en place.

En deuxième lieu, certains cas d'usage, plus complexes quant aux critères d'analyse juridiques, éthiques ou techniques, nécessitent un investissement plus lourd, mais qui apparaît pleinement justifié par les résultats attendus de manière établie par la littérature scientifique et l'état de l'art.

Il s'agit, en ce qui concerne l'exploitation des écritures des parties, des cas d'usage de recherche de connexité matérielle et intellectuelle⁸⁷ ainsi que des cas d'usage de détection des précédents et rapprochements de jurisprudence dans l'analyse des mémoires ampliatifs⁸⁸.

Il s'agit également des cas d'usage relatifs à l'exploitation des bases de données documentaires : les RAG ou outils conversationnels de recherche dans les bases de

⁸² Chapitre 1 – Cas d'usage transversaux : structuration et enrichissement des documents

⁸³ Chapitre 2 – L'exploitation des écritures des parties

⁸⁴ Chapitre 3 – Aide à la recherche et exploitation des bases de données documentaires

⁸⁵ Chapitre 4 – Aide à la rédaction

⁸⁶ Chapitre 4, section 3, §1 et §2

⁸⁷ Chapitre 2, section 2 – Détection de connexité matérielle et intellectuelle

⁸⁸ Chapitre 2, section 3, paragraphe 3 – détection des précédents et rapprochements de jurisprudence au sein des décisions rendues par la Cour de cassation

données à usage interne et externe⁸⁹ ou encore les cas d'usage d'aide aux rapprochements de jurisprudence, divergences internes à la Cour de cassation, voire éventuellement divergences entre la jurisprudence de la Cour de cassation et des décisions des juridictions du fond⁹⁰.

De même la recherche dans les bases de données de jurisprudence d'éléments utiles aux travaux de l'observatoire des litiges judiciaires (questions de droit nouvelles, divergences entre décisions du fond, litiges sériels, questions d'intérêt sociétal), quoique présentant une certaine complexité sémantique, constitue une piste très intéressante.

Pour ces cas d'usage, la voie d'action privilégiée reposerait essentiellement sur les ressources internes de la Cour de cassation, qui devraient être renforcées, en particulier en ce qui concerne les ressources nécessaires à la conduite de projet, mais parfois avec l'apport nécessaire de ressources externes.

En outre, le développement de certains cas d'usage pourrait reposer sur des partenariats lorsque les objectifs peuvent être partagés avec d'autres acteurs.

A cet égard, la collaboration d'ordre scientifique et technique sur ces cas d'usage pourrait s'inscrire dans le cadre du partenariat de la Cour de cassation avec *Sorbonne Cluster for Artificial Intelligence* (SCAI) ou encore dans celui envisagé avec d'autres partenaires institutionnels et universitaires sur l'exploitation de la jurisprudence.

Des partenariats avec d'autres acteurs institutionnels seraient de nature à permettre d'élargir les projets susceptibles d'être envisagés. Ainsi en irait-il de l'outil de recherche documentaire pour l'exploitation des travaux parlementaires⁹¹.

Il serait, de façon plus générale, très opportun d'envisager des partenariats avec le Conseil d'Etat, dont les missions juridictionnelles et de diffusion de la jurisprudence sont très proches de celles de la Cour de cassation. Des échanges peuvent également se poursuivre avec d'autres professions du droit pour identifier d'éventuelles synergies.

En troisième lieu, le groupe de travail souligne l'intérêt de cas d'usage présentant à ce stade des problématiques éthiques, juridiques et techniques réelles mais qui sont très prometteurs.

Il s'agit du cas d'usage d'aide à la rédaction reposant essentiellement sur l'analyse des affaires précédentes⁹². Ce cas d'usage pourrait offrir un gain très important de qualité et d'efficacité dans l'exploitation, par la Cour de cassation, de sa propre jurisprudence de la manière la plus adaptée à une affaire nouvelle. Toutefois, cela ne peut passer que par une réflexion juridique et éthique approfondie ainsi que par une participation directe aux avancées scientifiques de l'état de l'art.

⁸⁹ Chapitre 3, sections 1 et 2

⁹⁰ Chapitre 3, section 5

⁹¹ Chapitre 3, section 3

⁹² Chapitre 4, section 1

Ce projet paraît particulièrement important pour permettre à la Cour de cassation de se doter d'outils performants, dans une approche comparable à ce qui est proposé par des acteurs privés à d'autres professions du droit, tout en garantissant sa conformité aux principes juridiques et éthiques applicables, notamment par sa fiabilité et sa transparence.

Un tel projet ne peut être mené seul et sans investissement conjoint. Il répond à des problématiques communes à d'autres hautes juridictions de cassation en Europe. Il paraîtrait particulièrement adapté au groupe de travail de mener le développement de ce cas d'usage dans le cadre de la réponse à un appel à projets de l'Union européenne.

De même, le cas d'usage portant sur l'aide à la rédaction des litiges sériels, s'il présente une certaine technicité, représenterait un gain d'efficacité et de qualité pour la Cour.

Le groupe de travail considère que les autres cas d'usage étudiés, s'ils ne sont pas dénués d'intérêt, sont moins prioritaires à ce jour compte tenu, le plus souvent, de la charge importante qu'ils peuvent représenter sans garantie de résultats satisfaisants ou pour un gain fonctionnel moins important. Cela ne signifie pas pour autant qu'ils ne pourraient pas être développés avec profit, à l'avenir, au regard de l'évolution de certains critères, qu'il s'agisse, notamment, de la disponibilité de la donnée ou de l'état de l'art.

Le groupe de travail souligne aussi particulièrement l'interaction très forte qui existe entre les projets d'IA et le bénéfice d'un système d'information solide, fiable, centré sur les besoins des utilisateurs et prenant en considération des objectifs de bonne structuration et conservation des données utiles. De surcroît, de nombreux projets, pour lesquels un apport de l'IA a pu être envisagé, doivent reposer en réalité sur une amélioration des applicatifs métiers déployés à la Cour de cassation sans recours à l'IA.

Bien que la réflexion du groupe de travail ait été centrée sur les cas d'usage pour les besoins propres de la Cour, plusieurs de ceux dont il a été préconisé le développement pourraient intéresser l'ensemble des juridictions. Ainsi l'usage des SIA documentaires de recherche dans les bases de données interne et externe, comme les SIA d'analyse de données jurisprudentielles massives ont vocation à être accessibles aux magistrats des juridictions du fond. De même les cas d'usage d'aide à la rédaction de travaux à la Cour de cassation pourraient peut-être aboutir à la création de technologies susceptibles de contribuer à l'élaboration d'outils adaptés aux besoins des juridictions du fond.

Par ailleurs, **la méthodologie d'évaluation des cas d'usage suivie par le groupe de travail paraît pouvoir être également utilisée par les juridictions** qui s'inscriraient dans une démarche similaire et souhaiteraient par exemple à leur tour, comme certaines l'ont déjà fait avec l'aide de la DINUM, développer et expérimenter des programmes d'IA pour leurs besoins spécifiques.

A cet égard, le groupe de travail relève que **la démarche éthique doit primer**, en complément de la conformité juridique aux règles directement applicables en matière de protection des données et d'intelligence artificielle, afin de garantir l'utilisation d'une IA vertueuse, pleinement respectueuse des droits fondamentaux, de l'équilibre de la procédure et de l'office du juge. C'est à cette condition, essentielle, que l'intelligence

artificielle, dont le groupe de travail apprécie tout le potentiel au service de la justice, pourra véritablement renforcer la capacité du juge à remplir son office dans un état de droit.

Il paraît par ailleurs important de rappeler que **chaque utilisateur de système d'intelligence artificielle doit veiller à l'utiliser de manière conforme au droit et à l'éthique**, ce qui peut nécessiter une obligation de formation et une vigilance permanente qui peut s'inscrire dans le cadre d'une charte ou d'obligations déontologiques.

En outre, le groupe de travail considère que l'analyse éthique des cas d'usage de l'IA, effectuée dans le présent rapport au stade d'une première réflexion sur les outils à développer, devrait se poursuivre de manière continue tout au long du cycle de vie d'un projet, jusqu'à la décision, le cas échéant, de l'abandonner. **La mise en œuvre de cette méthode d'évaluation « de bout en bout » pose aussi la question de la gouvernance du développement de l'IA judiciaire.**

Le groupe de travail propose ainsi l'instauration d'un comité de suivi du développement des SIA à la Cour composé notamment de magistrats et de scientifiques du laboratoire d'innovation du SDER ainsi que des autres services de la Cour. Un des objets de ce comité pourrait être d'évaluer les projets déployés, le cas échéant au cours de phases expérimentales, au regard de l'ensemble des critères mis en avant par le groupe de travail, notamment celui de l'impact des SIA sur les droits humains. Il pourrait, à cet effet, dans le cadre de ces travaux, faire appel à des chercheurs en droit, en *data science* ou en sociologie.

Il conviendra aussi que la spécificité de l'activité juridictionnelle soit prise en compte dans le cadre du contrôle de la mise en œuvre du règlement IA par les autorités nationales compétentes. En complément, la question de la création, à l'échelle de l'institution, d'un comité consultatif d'éthique pour les SIA judiciaires, se posera, ainsi que la place que devra y tenir la Cour de cassation. Dans la mesure où ce comité serait notamment consulté sur l'impact sur les droits humains des SIA utilisées par les magistrats dans le cadre de leur activité juridictionnelle, son indépendance doit être assurée.

ANNEXES

ANNEXE 1 : LES LARGES MODELES DE LANGAGE : UNE REVOLUTION DANS L'INTELLIGENCE ARTIFICIELLE

L'intelligence artificielle connaît actuellement une révolution majeure grâce à l'émergence des modèles de langage, communément appelés LLM (*Large Language Models*). Ces systèmes sophistiqués transforment profondément notre manière d'interagir avec les technologies numériques, ouvrant des perspectives inédites dans de nombreux domaines, y compris celui du droit. Cette présentation vise à éclairer les professionnels du droit sur la nature de ces technologies, leurs créateurs et leurs implications potentielles pour la pratique juridique.

I. Comprendre les Modèles de Langage

Les modèles de langage représentent l'aboutissement de décennies de recherche en intelligence artificielle. Ces systèmes complexes, qualifiés de "large" en raison de leur taille considérable, sont capables de comprendre et de générer du langage naturel avec une sophistication sans précédent. Leur fonctionnement repose sur l'analyse de vastes corpus de textes, englobant la littérature mondiale, les publications scientifiques, et une multitude de documents techniques et professionnels.

L'architecture qui sous-tend ces modèles, baptisée "Transformer", constitue une innovation majeure introduite par Google en 2017. Cette avancée technique permet aux systèmes de saisir les subtilités du langage en analysant simultanément les relations entre différents éléments d'un texte, quelle que soit leur position dans la séquence. Cette capacité leur confère une compréhension contextuelle remarquable, proche de celle d'un lecteur humain.

Un aspect particulièrement fascinant de ces modèles réside dans leur apprentissage non supervisé. Contrairement aux systèmes traditionnels qui nécessitent une intervention humaine constante pour leur apprentissage, les LLM développent leur compréhension de manière autonome en prédisant les éléments manquants ou suivants dans une séquence de texte. Cette approche leur permet d'acquérir une connaissance générale extraordinairement vaste et nuancée.

II. Les Acteurs Majeurs et leurs Innovations

Le paysage des modèles de langage est dominé par plusieurs acteurs majeurs, chacun apportant sa vision et ses innovations spécifiques. OpenAI, entreprise californienne

fondée en 2015, s'est imposée comme un leader incontesté avec son modèle GPT-4. Initialement créée comme une organisation à but non lucratif, OpenAI a évolué vers un modèle commercial tout en maintenant une mission éthique forte. Son partenariat stratégique avec Microsoft, matérialisé par un investissement de plusieurs milliards de dollars, lui a permis de développer des applications révolutionnaires comme ChatGPT.

Sur la scène européenne, Mistral AI émerge comme un acteur prometteur. Cette entreprise française, fondée en 2023 par d'anciens chercheurs de DeepMind et Meta, adopte une approche novatrice centrée sur l'efficacité et l'open source. Leurs modèles, notamment Mistral 7B et Mixtral 8x7B, démontrent qu'il est possible d'atteindre des performances remarquables avec des architectures plus compactes et économes en ressources.

Hugging Face, entreprise fondée en 2016, s'est imposée comme une infrastructure centrale dans l'écosystème de l'IA en créant la plus grande plateforme de partage de modèles open source au monde. Initialement focalisée sur le traitement du langage naturel, l'entreprise franco-américaine a développé une communauté dynamique de plus d'un million de développeurs et chercheurs. Sa plateforme héberge désormais des dizaines de milliers de modèles et ensembles de données, facilitant ainsi l'accès et le déploiement des innovations en IA. À travers ses outils de développement et sa bibliothèque Transformers, devenue une référence dans le domaine, Hugging Face joue un rôle crucial dans la démocratisation des modèles de langage et promeut activement une approche ouverte et collaborative de l'intelligence artificielle.

En Asie, Alibaba se distingue par le développement de sa série Qwen. Ces modèles, fruit des efforts d'Alibaba Cloud, s'inscrivent dans la stratégie chinoise d'indépendance technologique. Leur particularité réside dans leur excellence en traitement multilingue, notamment pour la langue chinoise, et leurs capacités multimodales avancées qui leur permettent de traiter aussi bien le texte que les images.

III. Implications pour la Pratique Juridique

L'intégration des modèles de langage dans le domaine juridique ouvre des perspectives prometteuses tout en soulevant des questions importantes. Dans le domaine de l'analyse documentaire, ces systèmes permettent d'examiner rapidement des volumes considérables de documentation juridique, identifiant automatiquement les précédents pertinents et extrayant les informations essentielles des contrats et textes de loi. Cette capacité transforme la recherche juridique, permettant aux professionnels de se concentrer sur les aspects stratégiques de leur travail.

L'assistance à la rédaction constitue un autre domaine d'application majeur. Les LLM peuvent suggérer des formulations juridiques appropriées, vérifier la cohérence des arguments, et faciliter la standardisation des documents. Cette assistance permet d'optimiser le temps consacré aux tâches rédactionnelles tout en maintenant un haut niveau de qualité et de précision.

Cependant, l'utilisation de ces technologies dans le contexte juridique nécessite une approche prudente et réfléchie. La question de la fiabilité demeure centrale : les modèles

de langage, malgré leur sophistication, peuvent parfois générer des "hallucinations" ou des erreurs. La supervision humaine reste donc indispensable, particulièrement dans un domaine où la précision et l'exactitude sont cruciales.

La protection des données constitue un autre enjeu majeur. L'utilisation de ces technologies doit s'accompagner de mesures strictes garantissant la confidentialité des informations sensibles et la conformité avec les réglementations en vigueur, notamment le RGPD. Les professionnels du droit doivent donc développer des protocoles adaptés pour intégrer ces outils tout en préservant leurs obligations déontologiques.

IV. Application aux cas d'usage de la Cour

Le développement et la mise en production des différents cas d'usage identifiés au sein de la Cour de cassation nécessitent une étude précise des différents modèles de langage disponibles au grand public afin d'optimiser les performances et les coûts d'exploitation. Plusieurs critères doivent être pris en compte dans le choix du ou des modèles qui seront utilisés :

- Une bonne maîtrise du français, ce qui implique un modèle entraîné uniquement sur la langue française ou un modèle multilingue avec une part significative de données d'entraînement en français ;
- La taille du modèle qui va définir notre capacité à entraîner puis utiliser le modèle sur différentes architectures (ordinateurs personnels, ressources GPUs internes à la Cour, super-calculateur, ...) ;
- Une compréhension approfondie du raisonnement juridique et de la terminologie spécialisée. Cela nécessite que le pré-entraînement du modèle inclut un corpus juridique francophone conséquent. Il est cependant plus probable qu'il soit nécessaire d'effectuer un affinement du modèle choisi sur nos données spécifiques ;
- Une taille de contexte suffisante pour traiter des documents juridiques qui peuvent être longs.

Outre ces critères déterminants pour proposer des applications qualitatives et fonctionnelles, d'autres aspects doivent être pris en compte, notamment dans des soucis d'éthique :

- La licence d'exploitation du modèle, il faut privilégier les licences ouvertes (Apache 2.0, MIT) ou permissives ;
- L'origine des données d'entraînement afin de respecter la conformité aux droits d'auteurs, au règlement RGPD et au règlement IA. Dans le cadre du raisonnement juridique, il peut être judicieux d'éviter d'utiliser un modèle entraîné uniquement sur des données juridiques provenant de pays pratiquant la *common law* ;

- L’auteur du modèle de langage, de nombreux acteurs développent des LLMs, que ce soit des multinationales étrangères, des acteurs privés français ou bien des laboratoires de recherche publique.

Dans ce cadre, un état des lieux non exhaustif présentant différents modèles potentiellement adaptés est proposé ci-dessous. Il convient de mentionner que le développement de LLMs est un secteur très actif ce qui rend l’analyse des différentes publications de modèles de langage difficile à suivre.

V. Panorama des Principaux Modèles de Langage

Dans le paysage actuel des modèles de langage, nous pouvons distinguer trois grandes catégories d’acteurs : les modèles commerciaux propriétaires, les modèles open source, et les modèles spécialisés pour le français. Voici une analyse détaillée de ces différentes solutions.

A. Les Modèles Commerciaux Propriétaires

GPT-4o (OpenAI)

Accessible uniquement via API, ce modèle généraliste se distingue par ses capacités exceptionnelles en compréhension contextuelle et en génération de contenu. Particulièrement performant en français, il est largement utilisé par les éditeurs juridiques et les legal techs dans leurs applications RAG. Cependant, son utilisation nécessite un accord commercial avec OpenAI et présente des coûts d’utilisation significatifs. Sa taille exacte n’est pas publique, mais elle est estimée à plus de 1 000 milliards de paramètres.

Claude 3.5 Sonnet (Anthropic)

Dernier né de la famille Claude, ce modèle se distingue par sa capacité à traiter des contextes très longs (jusqu’à 200 000 tokens) et sa fiabilité accrue. Ses performances en français sont excellentes, et il est particulièrement adapté aux tâches nécessitant un raisonnement rigoureux. Comme GPT-4, il n’est accessible que via API et nécessite un accord commercial.

B. Les Modèles Open Source de Nouvelle Génération

La famille Mistral (Mistral AI)

L’entreprise française Mistral AI a développé des modèles remarquables par leur efficacité. Certains sont accessibles uniquement via une API payante, comme Mistral Large, mais d’autres sont disponibles gratuitement et libre d’utilisation :

- Mistral 7B
 - 7 milliards de paramètres ;

- Modèle ouvert disponible pour la communauté sous licence d'utilisation Apache 2.0 (licence permissive qui permet une utilisation libre, y compris à des fins commerciales, sans restriction majeure) ;
- Ce modèle multilingue (entraîné sur 5% de données en français) a été conçu pour surpasser les modèles de taille similaire, et même des modèles plus grands, dans diverses tâches, y compris le codage et le raisonnement. Il est optimisé pour des performances élevées tout en maintenant une efficacité computationnelle.
- Mistral 8x7B
 - Architecture de 45 milliards de paramètres, avec seulement 12,9 milliards de paramètres actifs grâce à une architecture de type « Mixture of Experts » (MoE) ;
 - Modèle ouvert disponible pour la communauté sous licence d'utilisation Apache 2.0 ;
 - Mistral 8x7B est conçu pour surpasser des modèles plus grands sur la plupart des benchmarks, tout en offrant une inférence plus rapide. Il dispose d'une fenêtre de contexte de 32 000 tokens et offre des capacités multilingues.
- Mistral Small 3
 - 24 milliards de paramètres ;
 - Modèle ouvert disponible pour la communauté sous licence d'utilisation Apache 2.0 ;
 - Lancé le 30 janvier 2025, Mistral Small 3 est le dernier modèle ouvert de la famille Mistral et atteste de gains de performances significatifs par rapports aux précédents.

La famille Qwen2.5 (Alibaba)

Disponible en plusieurs tailles (de 0.5B à 72B paramètres), ce modèle multilingue se distingue par sa polyvalence. La version 72B offre des performances proches de GPT-4 sur certaines tâches. Sa licence Apache 2.0 et son support multilingue en font une option intéressante, bien que ses performances en français soient moins documentées que celles des autres modèles.

La famille OLMo 2 (Allen AI)

La famille OLMo comprend des modèles de 7 milliards (7B) et 13 milliards (13B) de paramètres. Ces modèles sont entièrement open source et disponibles pour la communauté. Ils sont distribués sous la licence Apache 2.0. Les données d'entraînement ainsi que le code source utilisés pour ces modèles sont ouverts et accessibles, permettant aux chercheurs de facilement reproduire les résultats ainsi que d'examiner et d'utiliser les ensembles de données pour leurs propres travaux.

C. Les modèles spécialisés en langues européennes et en français

Modèle Lucie (OpenLLM France)

Lucie-7B est un modèle de 7 milliards de paramètres. Le modèle est disponible en open source pour la communauté. Il est distribué sous la licence Apache 2.0, offrant des libertés d'utilisation étendues. Les données d'entraînement ainsi que le code source utilisés pour ces modèles sont ouverts et accessibles.

Modèle CroissantLLM (CentraleSupélec & des startups françaises)

CroissantLLM est un modèle de 1,3 milliard de paramètres, disponible en open source pour la communauté. Il est distribué sous la licence Apache 2.0. Les ensembles de données utilisés pour l'entraînement, incluant des sources en français et en anglais, sont disponibles publiquement ainsi que le code source complet pour la formation et l'évaluation du modèle, favorisant la transparence et la collaboration.

Modèle EuroLLM-9B (UTTER - Unified Transcription and Translation for Extended Reality, financé par Horizon Europe, le programme-cadre de recherche et d'innovation de l'Union européenne)

EuroLLM est un modèle de 9 milliards de paramètres disponible en open source. Il est distribué sous la licence Apache 2.0. Les informations sur la disponibilité des données d'entraînement et le code source ne sont pas fournies.

Modèle Pharia-1-LLM-7B-control (Aleph Alpha)

Pharia est un modèle de 7 milliards de paramètres disponible en open source. Il est distribué sous la licence Open Aleph, permettant une utilisation à des fins de recherche non commerciale et éducative. Les données d'entraînement ne sont pas disponibles directement mais une description très détaillée de l'obtention et du traitement de celles-ci est disponible. Il est également mentionné que les données sont conformes aux règlements européens et nationaux, notamment pour le respect des droits d'auteur et de la vie privée.

Modèle ARIA 7B

Aria est un modèle à 7 milliards de paramètres disponible en open source. Il est distribué sous la licence Llama 2, permettant une utilisation commerciale conformément aux termes de cette licence. Le modèle a été entraîné sur un ensemble de données de haute qualité en français, mais la disponibilité publique de ces données n'est pas précisée.

Modèle	Paramètres	Performances		Disponibilité en français	Code/Données/Licence	Provenance	
		Mistral 7B	7,3B	Surpasse Llama 2 13B et Llama 1 34B sur plusieurs benchmarks. Excellente	Oui (5% des données)	/Apache 2.0	Privé (Mistral AI)

		performance en code et raisonnement.	multilingues).		
Mistral 3 24B	24B	Performances élevées en compréhension de texte et en génération de contenu.	Oui (données multilingues).	/Apache 2.0	Privé (Mistral AI)
Llama-3 8B	8B	Performances exceptionnelles en compréhension et génération de langage, code, et raisonnement.	Oui (5% des données multilingues)	/Llama 3.1 Community License	Privé (Meta)
OLMo 13B	13B	Surpasse Llama-2 13B et Qwen 2.5 7B sur plusieurs benchmarks, avec une consommation énergétique réduite.	Non spécifié	/Apache 2.0	Public (AI2)
Qwen 2.5 7B	7B	Performances solides en génération de texte et en raisonnement, avec un support multilingue étendu.	Oui (support multilingue)	/Apache 2.0	Privé (Alibaba)
Qwen 2.5 32B	72B	Performances de pointe en compréhension de langage et en génération de code, avec un support multilingue.	Oui (support multilingue)	/Apache 2.0	Privé (Alibaba)
DeepSeek-V3	671B	Performances impressionnantes en raisonnement complexe, mais avec des problèmes de répétition signalés.	Oui (support multilingue)	/MIT	Privé (DeepSeek AI)
OpenLLM-France Lucie 7B	7B	Performances solides en français, optimisé pour les tâches linguistiques spécifiques au français.	Oui (spécialisé en français)	/Apache 2.0	Public (OpenLLM France)

EuroLLM 9B	9B	Performances équilibrées en multilingue, avec une attention particulière aux langues européennes.	Oui (spécialisé en langues européennes).	/Apache 2.0	Public (Utter Project)
Aleph - Alpha Pharia 7B	7B	Performances solides en génération de texte et en contrôle de sortie.	Non spécifié	/Open Aleph License (utilisation non commerciale)	Privé (Aleph Alpha)
Faradaylab ARIA 7B	7B	Performances compétitives en génération de texte et en compréhension de langage.	Non spécifié	/ Llama 2 Community License (utilisation commerciale)	Privé (Faraday Lab)
Croissant LLM Base	1.3B	Performances solides en français, optimisé pour les tâches linguistiques spécifiques au français.	Oui (spécialisé en français)	/MIT	Public

ANNEXE 2 : ETUDES DU SERVICE DES RELATIONS INTERNATIONALES



SERVICE DES RELATIONS INTERNATIONALES

L'utilisation de l'intelligence artificielle dans le système judiciaire en Argentine⁹³

Contexte et objectifs

Avec le développement de l'outil Prometea, l'Argentine a été pionnière dans l'utilisation de l'intelligence artificielle (IA) par le système judiciaire. Cet outil a été développé par le Laboratoire de l'Innovation et de l'Intelligence Artificielle de la faculté de droit de l'Université de Buenos Aires en collaboration avec le bureau du parquet de Buenos Aires, et il est utilisé judiciairement depuis octobre 2017. Son objectif principal est d'accélérer les processus bureaucratiques et décisionnels dans les affaires judiciaires en fournissant des solutions automatisées à des litiges de différentes natures. Le système repose sur une base de données juridiques volumineuse et utilise des algorithmes pour proposer, par similarité, des solutions à des litiges.

Prometea a été adapté et déployé dans d'autres pays, notamment sous le nom de PretorIA en Colombie, démontrant l'efficacité prometteuse de cette technologie dans diverses juridictions.

Fonctionnalités et efficacité

Prometea a été entraîné sur 2400 jugements et 1400 avis juridiques préalables du parquet. L'outil fonctionne de la manière suivante. Dès qu'un dossier arrive dans le bureau d'un procureur, ce dernier charge son numéro dans l'outil. Par suite, Prometea recherche le

⁹³ [CORVALAN Juan G., LE FEVRE CERVINI Enzo Maria, "Prometea experience. Using AI to optimize public institutions", CERIDAP, 2020, Fascicolo 2/2020; Explora Latam, "Prometea".](#)

numéro de dossier et l'associe à un autre dossier. Il lit ensuite les décisions de première et de deuxième instance du système judiciaire de la ville et analyse en tout plus de 300 000 documents. Grâce à la recherche et la détection de certains mots-clés prédéfinis, Prometea génère un modèle d'avis sur la manière dont l'affaire devrait être résolue. Le document est ensuite vérifié par l'huissier de justice compétent qui peut l'éditer, le télécharger et l'imprimer.

Prometea propose des solutions automatisées pour les litiges touchant par exemple à l'attribution de logements ou d'aides sociales, ou encore aux accidents de la route. À titre d'illustration, concernant ces derniers cas, Prometea a été notamment utilisé par la Cour de Morón (province de Buenos Aires) afin de déterminer les liens de causalité dans les accidents de la route, démontrant ainsi la flexibilité de l'outil dans le traitement de différentes catégories de litiges judiciaires.

En somme, son taux de réussite est très élevé car l'outil est capable de proposer, en vingt secondes, une **solution correcte dans 96 % des cas**. Cela signifie que la solution proposée correspond à celle qui avait été donnée dans des cas passés substantiellement identiques.

L'efficacité de Prometea réside non seulement dans sa rapidité d'exécution, mais aussi dans sa capacité à générer automatiquement le projet de décision correspondant, réduisant **de 99 % les erreurs de typographie**.

De plus, ce système d'IA a été considéré comme un modèle pour d'autres pays d'Amérique latine, tels que la Colombie, où il a été adapté sous le nom de PretorIA. Ce dernier a été déployé à la Cour constitutionnelle colombienne pour analyser plus rapidement les décisions quotidiennes de *tutela* (réclamations constitutionnelles), ce qui permet d'analyser 2700 décisions par jour en cinq secondes, alors que cela prend 202 jours de travail en manuel.

Statistiques et impact

Les résultats de l'outil Prometea sont particulièrement impressionnants :

- 96% de taux de succès pour les solutions proposées dans des affaires similaires.
- 99% de réduction des erreurs de typographie dans les projets de décisions générés.
- Gain de temps significatif : là où la production de documents nécessitait deux heures et quarante minutes auparavant, Prometea réduit ce temps à environ seize minutes.
- Il permet de rédiger 1000 sentences judiciaires relatives au droit au logement en seulement 45 jours, alors qu'il faudrait 174 jours manuellement.
- La rectification de 6000 dossiers administratifs a été faite en 2 mois avec Prometea quand la procédure normale prenait 8 mois.

L'outil est devenu un atout stratégique pour le système judiciaire argentin, réduisant considérablement les délais de traitement des affaires tout en maintenant un haut niveau de précision.

L'Argentine, à travers Prometea, s'est positionnée à l'avant-garde de l'utilisation de l'intelligence artificielle dans la justice en Amérique Latine. Ce système offre un soutien significatif dans la gestion des affaires judiciaires, en particulier pour les litiges de masse et les questions sociales, tout en garantissant des décisions rapides et fiables. Son succès a conduit à des adaptations dans d'autres pays, renforçant ainsi dans la région l'idée que l'IA peut devenir un outil central dans l'administration de la justice.



SERVICE DES RELATIONS INTERNATIONALES

L'Utilisation de l'Intelligence Artificielle et des nouvelles technologies par le système judiciaire en Colombie⁹⁴

Contexte Général

En Colombie, le système judiciaire a entrepris l'intégration de l'intelligence artificielle (IA) afin de moderniser ses processus et améliorer l'efficacité de l'administration de la justice. L'un des projets les plus notables est PretorIA, un outil développé pour la Cour constitutionnelle colombienne. Ce projet s'inscrit dans une démarche plus large de transformation numérique, visant à optimiser le traitement des affaires judiciaires. Il a été présenté fin juillet 2020 par la Cour constitutionnelle de Colombie.

PretorIA : Présentation et fonctionnement

PretorIA a été adapté à partir du système Prometea, initialement développé en Argentine par le parquet de Buenos Aires. Ce programme est le fruit du travail conjoint du Laboratoire d'Innovation et d'Intelligence Artificielle de la faculté de droit de l'Université de Buenos Aires, qui a créé Prometea, et de l'Université del Rosario en Colombie. PretorIA repose sur un modèle d'apprentissage supervisé, utilisant des systèmes experts, et a été entraîné sur un ensemble de 1 200 affaires labellisées par une équipe de dix juristes.

L'objectif principal de PretorIA est d'aider la Cour constitutionnelle à gérer efficacement le grand volume de dossiers soumis chaque jour, notamment les décisions de *tutela* (recours constitutionnels urgents), qui sont nombreuses en Colombie. À titre d'illustration, la Cour constitutionnelle recevait environ 12 000 dossiers de *tutela* par semaine, soit environ 2 700 par jour, en 2019.⁹⁵

Le logiciel est entraîné pour examiner les critères spécifiques à la Cour constitutionnelle, notamment en matière de droits fondamentaux et de constitutionnalité des décisions. Le système est conçu pour analyser les affaires de manière standardisée, garantissant ainsi une uniformité dans le traitement des dossiers.

Grâce à PretorIA, le temps de traitement des affaires a considérablement diminué. Là où un traitement manuel nécessitait 202 jours de travail par homme, PretorIA effectue la même

⁹⁴ LEFEBVRE Arnaud, « En Colombie une intelligence artificielle va aider la justice », *Cinquième pouvoir*, 5 août 2020 [[En Colombie, une intelligence artificielle va aider la justice - Cinquième-pouvoir.fr](#)] ; SAAVEDRA Victor, UPEGUI Juan Carlos, « PretorIA and automating the processing of human rights cases », *Derechos Digitales*, Mars 2021 [https://www.derechosdigitales.org/wp-content/uploads/05_Informe-Colombia-EN_180222.pdf]

⁹⁵ [05 Informe-Colombia-EN_180222.pdf](#)

analyse en cinq secondes, avec un taux de succès de 95% sur 13 critères d'examen (parmi les 33 définis par la Cour). De plus, la production d'un document par PretorIA prend en moyenne une minute, contre deux heures et quarante minutes auparavant.

Réduction des erreurs

L'un des bénéfices majeurs de PretorIA est la réduction du taux d'erreurs humaines. L'outil génère des documents juridiques avec une précision accrue, minimisant les erreurs typographiques et procédurales qui peuvent survenir lors de la rédaction manuelle.

Gains de Productivité

PretorIA permet d'économiser un temps précieux pour les juges et les greffiers en accélérant la production de documents et l'analyse des décisions.

La transformation numérique induite par cet outil contribue également à alléger la charge de travail des magistrats et à améliorer la réactivité du système judiciaire colombien.

PretorIA s'intègre aux processus juridiques existants, facilitant l'examen des affaires urgentes et complexes. Avec un **taux de réussite de 95 %** sur les critères d'examen, PretorIA se montre fiable pour automatiser des tâches de routine tout en laissant les décisions finales aux magistrats.

Avantages et défis

Avantages : Rapidité, réduction des erreurs humaines, standardisation des processus, et gains de productivité.

Défis : Comme tout système d'IA, PretorIA doit continuellement évoluer pour faire face à la complexité croissante des affaires juridiques, et les juges doivent toujours valider manuellement les résultats.

PretorIA représente un exemple probant de l'utilisation de l'intelligence artificielle dans le domaine judiciaire en Colombie. En automatisant les tâches répétitives et en accélérant le traitement des affaires, cet outil semble permettre à la justice colombienne de répondre aux défis posés par l'augmentation du volume des affaires tout en garantissant la précision et l'efficacité des décisions judiciaires. Il reste cependant essentiel de maintenir une supervision humaine pour éviter toute dérive ou automatisation excessive dans un domaine aussi sensible que la justice.



SERVICE DES RELATIONS INTERNATIONALES

L'utilisation de l'intelligence artificielle et autres technologies par le système judiciaire en Espagne

L'Espagne a pris des mesures significatives pour intégrer les technologies de l'information et l'intelligence artificielle (IA) dans le domaine de la justice. Cette intégration s'inscrit dans le cadre plus large de la transformation numérique du pays, portée par des stratégies nationales et des mesures législatives spécifiques. Parmi les principaux développements figure l'« Agenda España Digital » ainsi que le décret-loi royal 6/2023 qui a formalisé l'utilisation de l'IA au sein de l'administration de la justice.

La stratégie nationale d'intelligence artificielle (IA)

L'un des piliers de la transition numérique en Espagne est la stratégie nationale d'intelligence artificielle, partie intégrante de l'« Agenda España Digital », un vaste plan visant à moderniser l'administration publique d'ici 2026. Ce plan est composé de huit stratégies spécifiques, dont l'une est dédiée à l'IA. Son objectif est de promouvoir une intelligence artificielle « inclusive, durable et centrée sur le citoyen ». Le gouvernement espagnol entend ainsi faire de l'IA un levier majeur pour améliorer l'efficacité des services publics, y compris celui de la justice.

Le décret-loi royal 6/2023 et ses dispositions sur l'IA dans la justice

Le décret-loi royal 6/2023, adopté le 19 décembre 2023, constitue un cadre législatif important pour l'utilisation de l'IA dans l'administration de la justice. Ce texte fait partie intégrante du « Plan de redressement, de transformation et de résilience de l'économie espagnole » qui est conçu pour répondre aux défis post-Covid-19 et moderniser les services publics.

Le décret-loi établit clairement que les outils d'IA doivent avoir un caractère instrumental et de support, sous le contrôle humain, en conformité avec les garanties procédurales et constitutionnelles. Cela signifie que l'IA ne peut remplacer les décisions humaines, mais uniquement les assister.

Trois types d'actions liées à l'IA sont identifiées dans ce cadre :

- Les actions automatisées : Les actions totalement automatisées concernent des tâches simples ne nécessitant pas d'interprétation juridique, comme la numérotation des dossiers, l'archivage, la production de copies, etc.
- Les actions proactives : ce sont des actions initiées automatiquement par les systèmes informatiques, sur la base d'informations contenues dans des dossiers, pour générer des avertissements ou des effets directs dans une procédure.

- Les actions assistées : L'IA peut générer des projets de documents complexes qui serviront de support à une décision judiciaire. Toutefois, ces projets ne prennent effet qu'après validation par un juge ou un magistrat.

Encadrement éthique et juridique

L'utilisation de l'IA dans la justice espagnole doit respecter des principes éthiques stricts, notamment ceux énoncés dans la Charte des droits numériques adoptée en 2021. Bien que non contraignante, cette charte sert de cadre de référence pour garantir la protection des droits des citoyens dans un environnement numérique. Le décret-loi 6/2023 affirme que les dispositions relatives aux juridictions sont guidées par les principes contenus dans cette charte.



SERVICE DES RELATIONS INTERNATIONALES

L'utilisation de l'IA et des technologies par le système judiciaire en Suède

La Suède se positionne comme un acteur important dans le développement de technologies visant à améliorer la transparence, l'accessibilité et l'efficacité de son système judiciaire, tout en respectant les principes éthiques et juridiques liés à la protection des données. Bien que ces initiatives soient encore en cours de développement, elles s'inscrivent dans une dynamique plus large à l'échelle européenne et internationale, visant à intégrer l'intelligence artificielle et d'autres technologies numériques dans le domaine de la justice.

À cet égard, comme illustré ci-dessous, plusieurs initiatives et outils sont en cours de développement.

Anonymisation et pseudonymisation des décisions de justice

Un des projets majeurs en Suède est la création d'un outil automatisé pour l'anonymisation et la pseudonymisation des décisions judiciaires semblable à celui mis en place par le SDER de la Cour de cassation dans le cadre du projet *open data*. Le projet suédois s'inscrit dans un mouvement plus large en Europe visant à protéger la vie privée des personnes impliquées dans des litiges lors de la publication de jugements. L'objectif est de réduire les atteintes potentielles à la protection des données personnelles dans les décisions rendues publiques. Des initiatives similaires sont également en cours dans d'autres pays européens tels que l'Autriche, le Danemark, la Finlande (outil « Anoppi »), et l'Espagne. En Suède, cet outil est en phase de conception.

Chatbot pour l'information des justiciables

Le Domstolsverket (l'Agence nationale des tribunaux suédois) a lancé un projet de chatbot, destiné à répondre aux questions fréquentes des justiciables concernant l'avancement de leurs dossiers judiciaires. Ce chatbot en cours de déploiement dans les juridictions suédoises est conçu pour faciliter l'accès à l'information judiciaire, en permettant aux justiciables de suivre le statut de leur dossier sans avoir à solliciter directement les greffes des tribunaux. Ce projet s'inscrit dans une tendance plus large en Europe où d'autres pays tels que l'Autriche travaillent également sur des chatbots pour les juridictions. Le déploiement du chatbot suédois permet ainsi d'automatiser une partie des interactions avec le public, réduisant la charge de travail administrative des juridictions tout en améliorant l'expérience des justiciables.

Outil de traduction automatique des décisions de justice

La Suède développe un outil de traduction automatique des arrêts de la Cour suprême vers d'autres langues, notamment l'anglais. De tels projets permettent de renforcer l'accessibilité

des décisions judiciaires suédoises à un public international et à promouvoir la transparence des décisions dans un contexte transnational. Cet outil sera fondé sur un système d'apprentissage machine entraîné à partir de textes juridiques préalablement traduits, ce qui permettra d'améliorer progressivement la qualité des traductions. En outre, il s'efforcera de respecter les exigences strictes en matière de protection des données. Enfin, il participe aux efforts de convergence et d'harmonisation du droit européen.

Autres initiatives européennes

Il est intéressant de noter que d'autres pays européens travaillent également sur des outils innovants. Par exemple :

- La Finlande : développement d'un outil de recouvrement automatisé des amendes pénales, utilisant la « Robot Process Automation » pour relier les paiements aux sanctions, facilitant ainsi la gestion des paiements et des défauts de paiement ;
- La Croatie et l'Estonie : mise en place d'outils de reconnaissance vocale pour automatiser la rédaction des procès-verbaux des audiences, un domaine que la Suède pourrait explorer à l'avenir.



SERVICE DES RELATIONS INTERNATIONALES

Utilisation de l'intelligence artificielle au sein de la Justice Brésilienne

Note réalisée avec Mme Meggie TCHOUTLA, magistrate de liaison au Brésil

Le SRI attire l'attention du groupe de travail plus spécialement sur l'applicatif « ATHOS ».

En 2022, environ 78 millions de procès étaient en cours dans les 91 tribunaux répartis sur le territoire brésilien, ce qui place le pays à la première place du classement des pays où le nombre de procès est le plus élevé au monde, selon le rapport « Justiça en numeros » du Conselho Nacional da Justiça (CNJ). Le recours à l'intelligence artificielle est devenu une nécessité pour éviter l'engorgement judiciaire et accélérer les processus judiciaires.

Des chercheurs de la Fondation Getulio Vargas (FGV) ont réalisé une étude sur l'utilisation de l'IA dans les tribunaux brésiliens et sur l'aide qu'elle apporte au travail quotidien des fonctionnaires⁹⁶. L'objectif était de comprendre quels problèmes ces systèmes cherchent à résoudre et comment ils ont facilité le travail des acteurs du monde judiciaire.

Les résultats de cette étude montrent que la plupart des outils présents dans le système judiciaire brésilien sont capables d'aider les fonctionnaires à réaliser un tri pour orienter les procédures. Caroline Tauk, juge au tribunal régional fédéral, estime que dans le cas du Tribunal suprême fédéral (STF), l'analyse de la réception d'un recours par un fonctionnaire, qui prend en moyenne 44 minutes, peut être résolue par un système d'intelligence artificielle en 5 secondes.

Pour réaliser l'enquête « *Intelligence artificielle : technologie appliquée à la gestion des conflits dans le système judiciaire* », la Fondation Getulio Vargas a envoyé des questionnaires aux 91 tribunaux du pays, qui comprennent les Cours de justice, les Cours fédérales régionales, les Cours régionales du travail, les Cours militaires et les Cours électorales régionales, ainsi que la Cour suprême de justice (STJ) et le Conseil national de justice (CNJ). Parmi ces organes, 44 tribunaux et le CNJ ont indiqué qu'ils disposaient d'un système d'IA, allant de simples programmes d'automatisation à des programmes plus complexes.

A cet égard, **la Cour suprême de justice possède l'un des secteurs d'intelligence artificielle les plus avancés et les plus spécialisés du pays**. Compte tenu de l'importance de ce tribunal, ces outils deviennent essentiels, car le STJ a un volume élevé d'affaires. Parmi ces outils, il y a notamment « ATHOS », un programme capable d'identifier et de contrôler

⁹⁶ [SALOMAO Luis Felipe, « Tecnologia aplicada a gestao dos conflitos no ambito do poder judiciario brasileiro », FGV Conhecimento \]](#)

les contentieux les plus récurrents jugés par ce tribunal (pour identifier la jurisprudence applicable à ces affaires) et d'examiner la recevabilité des recours.

Le TJ de Sao Paulo a pour sa part triplé l'usage de robots en 2023 : 15,5 millions de tâches ont été exécutées sur 84 applications disponibles dans toutes les unités judiciaires de l'état. Cela représente une augmentation de traitement plus de 254% par rapport à la moyenne des trois années antérieures qui était de 4,4 millions de tâches par an.

Il est important de noter que 91% de ces outils d'IA ont été développés par les services informatiques des tribunaux, et non par le secteur privé.

I. Les différents systèmes d'IA et leurs fonctions

Afin de bien comprendre comment les intelligences artificielles s'insèrent dans les activités du pouvoir judiciaire, ces systèmes peuvent être divisés en deux types de technologie (elles-mêmes sous divisées en deux catégories distinctes). Il y a, d'une part, un premier usage qui vise à rationaliser le temps passé sur des tâches simples, répétitives et de classification, et, d'autre part, un second usage qui est plus proche d'une assistance pour le magistrat à la prise de décision ou à l'orientation des procédures.

1) Les technologies déployées dans les tribunaux relatives à la rationalisation de tâches non juridictionnelles

a) Pour les activités d'accueil et d'information au public

Certains outils aident à l'administration d'un tribunal, comme le recours à des chatbots qui aident, orientent les justiciables et répondent à des questions simples.

Le TJ de Bahia a mis en place « o Judi », une intelligence artificielle. Cet assistant virtuel capable de converser avec les citoyens se limite à fournir des informations administratives. Dans le même ordre d'idées, il existe également un autre outil, appelé « AMON », qui recueille les photos des personnes qui entrent fréquemment dans les tribunaux, de sorte à ce que ces personnes n'aient pas à passer aux rayons X tous les jours pour entrer dans cette unité.

b) Pour la réalisation de tâches administratives

La plupart des outils technologiques utilisés par le pouvoir judiciaire au Brésil appartiennent à cette catégorie qui englobe les tâches administratives telles que la retranscription des audiences. Ces technologies soutiennent directement la gestion des cabinets des juges en effectuant des activités administratives pour aider les juges dans leur processus de prise de décision.

Le système ATHOS du STJ entre dans cette catégorie, et la Cour de justice du Paraná dispose d'un système similaire appelé LARRY et qui regroupe les demandes relevant du même type de contentieux. Le système est capable d'identifier les affaires présentant le même type de demande réparties dans l'État.

L'utilisation de l'intelligence artificielle (IA) pour automatiser les tâches répétitives dans le système judiciaire a permis d'accélérer le traitement des procédures. Lors d'un déjeuner organisé par l'Institut des avocats de São Paulo (IASP) le 28 août 2023, la « ministre » (Présidente du STJ) a déclaré qu'un projet visant à cataloguer et à regrouper les affaires

présentant des similitudes sémantiques a permis de diminuer le temps de traitement de 2 390 000 affaires, soit par des retraits, soit par des abstentions d'appel.

La présidente du STJ a indiqué que la Cour a identifié, à partir du regroupement des affaires, des hypothèses dans lesquelles les prétentions des plaideurs étaient clairement contraires aux précédents de la Cour. En conséquence, ils ont conclu un partenariat avec le Ministère public de la Confédération (MPC) qui a accepté de renoncer aux prétentions ayant de faibles chances de succès.

Depuis le début de l'accord en 2020, le nombre d'appels spéciaux (AREsps) déposés par l'AGU (Avocat Général de l'Union) a diminué de 28%. Environ 4 850 affaires qui avaient déjà fait l'objet d'un recours devant le STJ ont également été retirées.

2) Les technologies déjà déployées dans les tribunaux d'assistance au magistrat

a) Pour filtrer ou prioriser les affaires : sous supervision humaine

Le troisième type d'IA aide également à rendre la justice, mais en plus de classer et de filtrer les affaires, ces technologies sont plus proches des juges et les aident à rédiger des décisions, des sentences, ces technologies peuvent mettre en évidence certains précédents qui aident les juges à prendre une décision.

L'une des utilités de ces outils est qu'ils permettent d'identifier les affaires judiciaires susceptibles d'avoir des répercussions générales. Il s'agit d'affaires qui dépassent les intérêts des deux parties impliquées dans le jugement, qui touchent un grand nombre de personnes ayant des intérêts politiques, sociaux, économiques, et qui sont susceptibles d'avoir d'éventuelles répercussions médiatiques.

Certains systèmes, comme « VICTOR », sont capables d'analyser en 5 secondes si une affaire a des chances d'être classée comme ayant une "répercussion générale" (au-delà d'un simple cas d'espèce), ce qui permet au fonctionnaire de consacrer ses efforts à une autre activité.

Un autre exemple d'outil allant dans ce sens est ELIS, de la Cour de justice de Pernambuco (TJ PE), qui s'occupe de l'application des taxes, des processus de recouvrement des dettes, du triage des affaires, de l'information sur le fait que le recouvrement d'une certaine dette est toujours possible, si ce recouvrement a déjà dépassé les cinq ans autorisés par la loi pour que l'individu puisse être accusé.

Ces outils permettent d'identifier les délais de prescription fiscale, la répercussion générale, entre autres types de cas, mais toujours sous une double supervision humaine, à partir d'un fonctionnaire, et ensuite du juge lui-même.

b) Pour la conciliation et l'assistance au magistrat

Enfin, le quatrième groupe d'IA dans le système judiciaire brésilien se réfère à une minorité qui se concentre sur la phase de conciliation ou d'assistance dans la prise de décision.

Utilisant des informations provenant d'affaires judiciaires antérieures similaires, cette IA a pour but de vérifier quelles affaires ont le plus de chances de se terminer par une conciliation entre les parties impliquées dans l'affaire.

Lorsqu'un juge reçoit la proposition, ce système analyse le type de recours et le justiciable, entre autres informations, pour indiquer les affaires qui ont le plus de chances d'aboutir à une conciliation, et ceci afin de mieux informer le juge pour qu'il puisse prendre sa décision. Pour les conciliations les plus simples, les juridictions peuvent avoir recours à un chatbot.

Concernant l'assistance à la prise de décision par le juge, il existe plusieurs types d'initiative ayant recours à l'IA :

- Le Tj du Mato Grosso développe depuis 2020 en partenariat avec Amazon un robot pour assister les juges dans la production de minutes de jugements ;
- Dans la justice du travail, il existe un « Monitor do Trabalho Decente » qui analyse les décisions rendues en matière de travail des enfants, harcèlement sexuel, etc, afin de permettre aux tribunaux de savoir quels types de décisions sont rendues sur ce type de contentieux ;
- Gemini permet de trouver des jurisprudences sur des cas similaires pour appuyer les décisions rendues ;
- Le TJ de Santa Catarina utilise une IA qui « esquisse » les décisions judiciaires.

II. Une multiplication du recours à l'intelligence artificielle dans le monde judiciaire brésilien

De nombreux exemples de l'usage de l'intelligence artificielle existent au niveau de la justice nationale et étatique :

- PJe : Processo judicial Eletrônico

PJe offre une numérisation totale des actes via une plateforme utilisée par les juges, avocats, et justiciables, laquelle permet de déposer une demande dans un tribunal et de suivre en temps réel l'avancée de cette demande.

En vertu de la résolution 335/2020 du CNJ, le pouvoir judiciaire a lancé la Plateforme numérique du pouvoir judiciaire (PDPJ-BR) qui vise non seulement à encourager l'utilisation de l'outil PJe afin d'accroître la numérisation, mais aussi à permettre une plus grande collaboration et intégration entre les applications existantes dans les 90 (quatre-vingt-dix) tribunaux du pays.

Il s'avère qu'en plus de l'intégration des systèmes, il existe un mouvement clair pour appliquer davantage d'outils technologiques à l'application et à la réalisation de la justice dans le pays.

- Athos (STJ)

Le projet Athos date de 2018. Il a pour but d'offrir une sécurité juridique et de permettre de gagner en rapidité dans le traitement des procédures.

L'ensemble des documents du STJ y sont compilés, soit plus de 18 millions de documents et 3,5 millions de décisions. Le logiciel trie les documents et aide au travail d'analyse des procédures. Il recherche des cas similaires, les regroupe et permet de dégager une solution médiane entre les différents cas, pour préparer les suivants. Il permet aussi d'établir des statistiques, notamment afin de connaître le nombre de décisions similaires rendues.

L'ensemble des personnes intervenant sur la procédure au tribunal travaille avec le logiciel Athos. L'ensemble du processus décisionnel est fait avec le logiciel. Cela a permis d'accélérer le travail et de diminuer le temps pour rendre une décision.

Plus de 90% des quelques 300 000 affaires traitées chaque année par le STJ sont numériques. Le succès de cette initiative a inspiré l'adoption de la même pratique par les tribunaux d'origines, qui envoient numériquement leurs dossiers au STJ, évitant ainsi les coûts élevés de gestion et de transport du papier.

- **Socrates**

Socrates est l'outil du STJ qui aide les juges à identifier à l'avance les controverses juridiques dans les appels spéciaux. Pour un grand nombre d'affaires, l'application s'avère d'une valeur indéniable. Athos et e-Juris sont deux autres algorithmes appliqués à la Cour des citoyens pour identifier à l'avance les recours avec un système d'attribution et pour extraire des références législatives et jurisprudentielles.

- **Rafa et Victor (développé en partenariat avec l'UnB, Université de Brasilia)**

Rafa et Victor sont des applications gérées par le STF. Victor, utilisé depuis 2017 par la Cour suprême, est chargé d'analyser les thèmes de répercussion générale pour filtrer les recours extraordinaires déposés dans tout le pays. Lors d'une cérémonie dirigée par la juge Rosa Weber, le STF a également lancé le plus récent outil de la cour, appelé VictorIA, une application utilisant l'IA pour regrouper et classer les affaires judiciaires.

Rafa 2030 (Artificial Networks Focused on Agenda 2030) est un outil d'intelligence artificielle utilisé pour classer les actions en fonction des objectifs de développement durable (ODD) de l'Agenda 2030 de l'Organisation des Nations unies (ONU).

Dans l'Agenda 2030 de l'ONU, tous les pays se sont engagés à adopter dix principes dans les domaines des droits de l'homme, du travail, de l'environnement et de la lutte contre la corruption. Le pouvoir judiciaire brésilien a décidé de participer directement à cet agenda afin de rendre les institutions plus efficaces, responsables et inclusives, et par conséquent de rendre les processus judiciaires plus rapides, capables de fournir des réponses plus rapides, et de connecter les activités des tribunaux avec les principes des ODD

- **JusticIA** : le développement de l'IA dans les bases de données juridiques

Intégré à la vaste base de données Jus, l'assistant JusticIA est en mesure d'interpréter et de simplifier les termes juridiques, de répondre aux questions et d'aider à la rédaction de documents.

JusticIA, logiciel gratuit, est le premier service juridique d'IA disponible pour un large public au Brésil. Développé par des spécialistes, JusticIA est une étape innovante pour l'entreprise Jus Navigandi qui est une référence en matière d'information juridique sur l'internet brésilien. Pionnier du monde juridique sur l'internet, Jus possède la collection la plus vaste et la plus complète de doctrine juridique en portugais. Il compte plus de 100 000 articles rédigés par les plus grands experts brésiliens sur les sujets traités.

- **Le Saref**

Le Saref (système de présentation par reconnaissance faciale) au TJ du District Federal est utilisé par le juge d'application des peines pour les condamnés en milieu ouvert (le condamné utilise son téléphone pour se prendre en photo et est géolocalisé pour confirmer sa présence)

- **Bem-te-vi** (STF) : système d'appui au cabinet notamment dans l'activité de tri des procédures ;
- **Sigma (TRF)**: aide à la production de minutes et au classement de décisions ;
- **Analise Legal Inteligente – Alei (TRF)** : assistant intelligent pour l'élaboration de décisions en lien avec la jurisprudence antérieure ;
- **Toth** qui lit la demande initiale, vérifie qu'elle est complète et, en fonction des éléments renseignés, peut orienter la demande.

Si l'application de l'IA dans l'environnement judiciaire brésilien est déjà une pratique réelle, les défis que pose son utilisation restent importants, notamment en ce qui concerne la formation et l'adaptation des fonctionnaires et magistrats à la technologie ainsi que la nécessité d'inscrire l'usage de l'intelligence artificielle dans un cadre éthique, notamment en ce qui concerne la confidentialité des données, la responsabilité des décisions prises et la discrimination algorithmique qui ne doit pas en créer de sociales. Les nouveaux moyens technologiques ne sauraient s'émanciper du contrôle d'un juge.



SERVICE DES RELATIONS INTERNATIONALES

Étude synthétique comparative : l'utilisation de l'intelligence artificielle dans les systèmes judiciaires étrangers

15 octobre 2024

L'intelligence artificielle (IA) est devenue une composante essentielle de la modernisation du travail des acteurs économiques et juridiques, notamment dans le secteur privé. Avec l'émergence d'outils commerciaux et publics performants et opérationnels, les institutions judiciaires s'intéressent aussi à l'utilisation de ces outils pour faire face aux défis que posent l'accroissement des contentieux de masse, les délais de traitement des dossiers et la complexité croissante des contentieux techniques. Ainsi, à l'instar de la Cour de cassation française, de nombreuses juridictions étrangères ont entrepris des initiatives pour intégrer des outils d'IA.

Dans un premier temps, cette étude comparative non-exhaustive propose un état des lieux de l'utilisation de l'IA dans huit pays — l'Allemagne, l'Argentine, la Colombie, l'Espagne, le Brésil, Singapour, l'Inde et la Suède — en regroupant les initiatives en fonction de thématiques clés afin d'offrir une vue d'ensemble cohérente.

Puis, dans un second temps, les besoins et pistes identifiés par le groupe de travail IA de la Cour de cassation sont mis en parallèle avec les initiatives étrangères existantes.

Partie I

I. Assistance à la prise de décision des juges

L'un des usages les plus novateurs de l'IA dans les systèmes judiciaires consiste à assister les juges dans la prise de décision, sans pour autant se substituer à l'humain.

Un exemple phare se trouve en Argentine avec l'outil Prometea développé par le parquet de Buenos Aires. Cet outil, un modèle d'efficacité, permet par exemple de résoudre des litiges sociaux et liés au logement en quelques secondes, avec un taux de réussite de 96%. En générant automatiquement des projets de décisions, il réduit également les erreurs typographiques de 99%.

Ce système a permis de réduire de manière spectaculaire les délais de traitement des affaires judiciaires. Par exemple, la production de 1000 décisions relatives aux droits de logement, qui prenait auparavant 174 jours, est maintenant réalisée en 45 jours grâce à Prometea. Dans

d'autres domaines, comme le droit du travail, le délai est passé cette fois de 83 jours à 5 jours. Prometea a également été testé par la Cour interaméricaine des droits de l'homme, et les résultats de son utilisation à la Cour constitutionnelle de Colombie mettent en lumière la réduction des délais de traitement des dossiers prioritaires de 96 jours à 2 minutes, ainsi que l'augmentation de l'efficacité dans la création de documents de 937%.

En Colombie, une version de Prometea, appelée PretorIA, est utilisée par la Cour constitutionnelle pour analyser les décisions de *tutela* (réclamations constitutionnelles urgentes). PretorIA est capable d'examiner environ 2700 décisions par jour, et met cinq secondes à réaliser des analyses qui nécessitaient auparavant 202 jours de travail, contribuant ainsi à la réduction des délais de traitement des dossiers. Il produit en une minute un document dont la rédaction humaine nécessite en moyenne 2h40.

En Allemagne, des projets comme Codefy, mis en place dans les Länder tels que Bade-Wurtemberg, offrent une aide précieuse aux juges dans le traitement des contentieux complexes à travers une structuration automatique des dossiers ainsi qu'une assistance dans l'analyse du contenu de ces derniers. L'idée est également de créer une « plateforme de connaissances » avec notamment un recueil de modèles de décisions. Il peut s'agir d'affaires avec un volume important ou une complexité particulière (ex : droit privé de la construction).

Selon le magistrat de liaison français en Allemagne, le système allemand OLGA (*OberLandesGerichtAssistent*) pourrait apprendre à partir des documents analysés pour mettre en évidence les divergences entre les mémoires et les développements dans l'argumentation pour identifier de nouveaux paramètres. Le but est également de créer une grande base de données avec des documents consultables grâce à une indexation automatique.

Un autre projet, MAKI, mis en place en Basse-Saxe, compare les nouveaux dossiers avec ceux existants pour générer des propositions de décisions que les juges peuvent personnaliser.

En Inde, l'outil SUPACE (Supreme Court Portal for Assistance in Court's Efficiency) a été lancé en 2021 pour aider les juges de la Cour suprême à organiser les dossiers, extraire les informations pertinentes et rédiger des documents préparatoires. Cet outil, bien que facilitant le travail du juge dans la prise de décision, ne prend aucune décision autonome. Ainsi, il aide le juge, tout en garantissant que ce dernier conserve le contrôle total sur le jugement final. C'est une caractéristique clé des initiatives que l'on retrouve chez les autres cours étrangères.

En Espagne, le décret-loi royal 6/2023, qui fait partie du Plan de redressement post-Covid-19, encadre l'utilisation de l'IA dans l'administration publique et judiciaire. Ce cadre permet l'utilisation de l'IA dans différentes catégories de tâches qui varient en technicité et complexité, mais aussi en impact pour le justiciable. Si le cadre prévoit l'utilisation des IA allant jusqu'à la rédaction d'un projet de décision (tout en exigeant que toutes les actions générées par l'IA soient validées par un juge), la plupart des utilisations relève de l'automatisation des tâches administratives.

II. Automatisation des tâches et gestion des contentieux de masse

L'automatisation des tâches administratives grâce à l'IA est un moyen efficace de soulager les juridictions des tâches répétitives et d'accélérer le traitement des contentieux de masse.

Au Brésil, le système ATHOS, utilisé par la Cour suprême de justice (STJ), est à mi-chemin entre la gestion de contentieux de masse et l'aide à la prise de décision : il permet de regrouper les dossiers similaires et d'analyser des millions de documents pour offrir aux juges des solutions fondées sur la jurisprudence. Cela accélère le processus décisionnel tout en garantissant la cohérence des décisions.

D'autres juridictions au Brésil ont aussi embrassé l'utilisation de l'IA : le TJ de Sao Paulo a pour sa part triplé l'usage de robots en 2023. Ainsi, 15,5 millions de tâches ont été exécutées sur 84 applications disponibles dans toutes les unités judiciaires de l'état.

Il est important de noter que 91% de ces outils d'IA ont été développés par les services informatiques des tribunaux, et non par le secteur privé.

Il ressort des informations recueillies sur les initiatives allemandes, notamment par le biais du magistrat de liaison, que l'Allemagne, avec sa justice fédérale, est un vivier de différents projets (déployés ou en cours d'élaboration) en matière de gestion de contentieux de masse ou très similaires (ex : plaintes liées au transport aérien).

Le logiciel « OLGA » (*OberLandesGericht.Assistent*) a notamment été utilisé dans des affaires telles que celle du Dieselgate. Ce projet, qui combine traitement du langage naturel (NLP) et apprentissage automatique, vise à minimiser les ressources humaines dans le cadre de la gestion des procédures de masse. Il s'agit d'abord de numériser l'ensemble des jugements et des mémoires de première instance afin de pouvoir affecter automatiquement, selon certains paramètres, les dossiers en appel à des groupes thématiques et conséquemment de faciliter leur traitement.

Le logiciel Codefy, mentionné dans la partie précédente a pour objectif un meilleur traitement des contentieux de masse à travers une structuration automatique des dossiers ainsi qu'une assistance dans l'analyse du contenu de ces derniers.

De nombreux outils sont également utilisés pour automatiser des tâches comme la rédaction de procès-verbaux (à partir de reconnaissance vocale), et l'anonymisation des décisions judiciaires. Le projet GOTTBERT (Ministère de la justice bavarois en coopération avec l'université de Erlangen-Nürnberg), utilise l'IA pour garantir l'anonymisation des jugements publiés, tout en maintenant un haut niveau de précision. En Suède, un projet similaire vise à protéger les données personnelles en anonymisant les décisions avant leur publication.

Notons qu'un grand nombre de pays, non concernés par cette étude, ont des outils (ou les envisagent) capables de « pseudonymiser » les décisions : Autriche, Danemark, Suède, Finlande (outil « Anoppi »), Estonie (outil « Texta »), République tchèque, Croatie, Espagne, et Luxembourg.

III. Accessibilité et suivi des procédures judiciaires

L'utilisation de l'IA est également étudiée et envisagée pour améliorer l'accès à la justice et faciliter le suivi des procédures.

À Singapour, plusieurs chatbots ont été développés pour aider les justiciables à comprendre leurs droits et suivre leurs affaires. Divorce AIDE, par exemple, aide à estimer la répartition

des biens lors d'un divorce, tandis qu'un autre chatbot, fondé sur l'IA générative, est en phase de test au Tribunal des petites créances pour fournir des conseils juridiques simples.

En Inde, des initiatives comme le système eCourts services app permettent aux citoyens d'accéder à leurs dossiers judiciaires et de suivre leur évolution en temps réel. Les cours virtuelles, qui permettent de juger certaines affaires en ligne, simplifient l'accès à la justice pour les populations éloignées des centres judiciaires, et augmentent la publicité des débats. Ces outils font souvent appel à un système partiel ou total d'intelligence artificielle, dépassant la simple base de données ou plateforme. Le gouvernement indien a récemment approuvé le projet de développement d'une nouvelle version de eCourt (« Phase iii ») qui s'appuiera davantage sur l'IA.

IV. Gestion des données et traduction automatique

Dans certains pays, l'IA est utilisée pour améliorer la gestion des données et pour surmonter les barrières linguistiques dans le cadre de la justice.

Depuis 2017, l'Allemagne a créé un « Service de traduction automatique de la justice » fondé sur un réseau neuronal. Ce projet a été mis en œuvre avec l'aide d'une entreprise privée française.

En Inde, le système SUVAS (Supreme Court Vidhik Anuvaad Software) est utilisé pour traduire automatiquement les documents judiciaires entre l'anglais et les langues locales. Cela est particulièrement crucial dans un pays où l'anglais, bien qu'étant la langue judiciaire officielle, n'est parlé que par 14% de la population. Ainsi, ce logiciel permet de contribuer à une meilleure accessibilité de la justice.

De même, en Suède, un projet est en cours pour traduire automatiquement les décisions de la Cour suprême vers d'autres langues, notamment l'anglais. Cette initiative vise à renforcer la transparence et à faciliter la diffusion des décisions judiciaires suédoises à l'international.

V. Défis éthiques et juridiques de l'utilisation de l'IA

Il ressort des divers conférences, publications et textes juridiques que l'utilisation croissante de l'IA dans les systèmes judiciaires soulève des questions éthiques et juridiques importantes, notamment en matière de responsabilité des professionnels du droit et de biais algorithmiques.

En Allemagne, le cadre constitutionnel est particulièrement précis en ce qui concerne l'IA. Les articles 92 et 101 de la Loi fondamentale prévoient que seules les personnes physiques peuvent exercer le pouvoir judiciaire. L'IA ne peut donc être utilisée que pour assister les juges, sans jamais se substituer à eux.

À Singapour, l'affaire *Quoine Pte Ltd v B2C2 Ltd* a mis en lumière les défis posés par l'utilisation de systèmes d'IA. La Cour d'appel a jugé que, dans le cas des systèmes déterministes, la responsabilité revient au programmeur, mais l'application de ce principe aux systèmes plus complexes reste un sujet de débat.

Au Brésil, bien que l'IA est utilisée pour faciliter la prise de décision et organiser les dossiers, le système judiciaire prévoit la garantie que les décisions finales restent sous le contrôle humain. Par exemple, le TJ de São Paulo a considérablement augmenté l'utilisation de robots

pour automatiser les tâches administratives, mais toutes les décisions sont toujours supervisées par un juge.

Partie II

VI. Comparaison et liens avec les besoins identifiés par le Groupe de travail IA en septembre 2024

L'analyse croisée, d'une part, des besoins exprimés par les chambres de la Cour de cassation française en matière d'intelligence artificielle et, d'autre part, des technologies déjà mises en œuvre dans d'autres juridictions à travers le monde permet de dégager des points de convergences. Il est possible d'envisager des pistes pour la modernisation du système judiciaire français à partir des solutions déjà déployées ailleurs ou en cours d'étude.

Tout d'abord, il ressort du GT IA de la Cour de cassation un souhait de renforcer l'orientation automatique des dossiers et le besoin de détection des litiges sériels.

À cet égard, plusieurs systèmes utilisés à l'étranger présentent des similarités avec ce que la Cour souhaiterait mettre en place. Par exemple, en Colombie, l'outil PretorIA aide la Cour constitutionnelle à traiter un volume massif de recours constitutionnels (tutelas), permettant d'identifier et de standardiser les décisions pour des affaires similaires. PretorIA est capable d'analyser environ 2700 décisions en cinq secondes, un gain de temps considérable qui reflète ce que la Cour de cassation pourrait attendre d'un outil d'orientation et de gestion des litiges sériels.

Autre exemple pertinent, le système ATHOS, utilisé par la Cour suprême de justice du Brésil, trie les contentieux les plus récurrents et propose des solutions fondées sur la jurisprudence existante, accélérant ainsi le processus décisionnel dans les affaires de masse.

De même, en Allemagne, le **projet MAKI**, testé en Basse-Saxe, joue un rôle similaire en comparant automatiquement les nouvelles affaires avec celles déjà jugées, facilitant ainsi la gestion des contentieux de masse. **Le système OLGA**, utilisé dans la gestion des dossiers de l'affaire Dieselgate, peut aussi être intéressant pour la Cour de cassation en matière d'affectation automatique des dossiers en appel à des groupes thématiques afin de faciliter leur traitement.

Un second besoin identifié par la Cour concerne l'assistance à la recherche juridique et les veilles automatisées. Plusieurs chambres et le parquet général ont souligné la nécessité d'un système centralisé permettant de consulter rapidement les bases de données comme Jurinet ou Judilibre, tout en automatisant les veilles jurisprudentielles.

À cet égard, le système SUPACE utilisé par la Cour suprême de l'Inde est un exemple concret d'une IA assistante. SUPACE aide les juges à organiser les informations complexes, à effectuer des recherches juridiques avancées et à préparer des documents. Cette IA augmente l'efficacité des juges en allégeant la charge de travail liée à la recherche d'informations, leur permettant ainsi de se concentrer davantage sur le raisonnement juridique et la rédaction.

En Espagne, le décret-loi royal 6/2023 prémentionné met également en place des outils d'IA destinés à assister les juges dans la rédaction de documents et dans la recherche d'informations, tout en veillant à ce que l'IA reste un simple soutien sous supervision

humaine. Ces systèmes démontrent que les outils de recherche automatisée et de veille que la Cour de cassation souhaite développer existent déjà dans d'autres juridictions et offrent des résultats intéressants, même si leur utilisation reste expérimentale pour certains, et absolument toujours encadrée par une personne physique.

Le système OLGA qui est capable **d'identifier des divergences** pourrait aider à apporter des éclairages pour la Cour de cassation qui a relevé ce besoin des chambres.

En matière d'automatisation des tâches administratives, notamment pour la « pseudonymisation » des décisions avant leur publication, la Cour de cassation pourrait s'intéresser aux projets en cours dans plusieurs pays européens afin d'identifier d'éventuelles pistes d'améliorations et parfaire son propre système mis en place au sein du SDER dans le cadre du projet *open data*. En Suède, un outil d'anonymisation et de « pseudonymisation » des décisions judiciaires est en cours de développement, garantissant la protection des données personnelles tout en facilitant la publication des jugements. Un projet similaire est en cours en Allemagne, où une IA fondée sur l'apprentissage profond est utilisée pour anonymiser automatiquement les décisions judiciaires dans les Länder de Bavière et Bade-Wurtemberg.

Plus généralement, des juridictions de länder allemands pourraient inspirer la Cour de cassation avec la grande panoplie d'expérimentations. Il y a par exemple un outil de reconnaissance vocale par IA qui permet **la retranscription automatique** d'interrogatoires, un logiciel permettant de réaliser une traduction rapide de documents, un Chatbot au soutien des greffes des tribunaux, ou encore un logiciel permettant de créer des « arbres à décision » pour représenter des processus de travail répétitif et uniformes. Ces arbres à décision permettent ensuite à une IA d'exécuter de manière automatisée ces tâches.

L'étude des différents systèmes de traduction automatique de documents déployés à l'étranger pourraient être une source d'enseignements pour la Cour de cassation française, aidant ainsi à répondre aux besoins de traduction automatique juridique exprimés par la première présidence, le parquet général, la 1^{ère} civ et la chambre commerciale. Il conviendrait de se rapprocher des institutions allemandes, indiennes, et suédoises.

Enfin, l'assistance à la rédaction des documents représente un domaine où la Cour de cassation pourrait tirer des enseignements des pratiques internationales qui visent globalement à assurer un juste équilibre entre gain d'efficacité et qualité juridique et rédactionnelle.

Le décret-loi royal 6/2023 en Espagne encadre l'utilisation de l'IA pour générer des projets de documents complexes, tout en s'assurant que ces documents soient validés par un juge avant de produire des effets. Les essais de système générant des propositions de décisions en Inde ou en Allemagne, tout comme le chatbot fondé sur l'IA générative singapourien, actuellement testé dans le Small Claims Tribunal, aident à rédiger des conseils juridiques dans des cas simples tels que les litiges de voisinage et les testaments.

Toujours dans l'esprit de l'aide à la rédaction des documents, le GT intelligence artificielle a recueilli le souhait d'exploiter l'IA pour favoriser/aider à **l'uniformisation des documents** (Vérification et correction automatique des normes de saisies, des oublis de moyens (Civ 1) (Soc), (Crim) (Com), de la cohérence des textes visés et des moyens (Crim) : le système Prometea, développé par le parquet de Buenos Aires, joue un rôle de correcteur automatique, éliminant 99% des erreurs typographiques. Avec d'autres systèmes qui proposent des projets

de rédactions/de décisions, il pourrait servir de modèle illustrant les possibilités de l'IA pour uniformiser/harmoniser des documents.

Les expériences de Chatbot de plusieurs cours (Inde, Allemagne, Autriche, Suède, Singapour) peuvent servir d'inspiration pour la mise en place de systèmes semblables à la Cour, ou encore la **création d'un « avatar »** pour l'institution. La retranscription automatique, souhaitée par plusieurs chambres ainsi que par la première présidence et le parquet général, est déjà déployée dans des juridictions allemandes (logiciel Dictasoft). La Croatie (outil « Speech-to-text » en service dans les juridictions croates depuis décembre 2018), l'Estonie, ou encore la Hongrie développent ou ont développé des outils de création automatisée des procès-verbaux d'audiences utilisant l'apprentissage machine (reconnaissance vocale).

La chambre sociale a exprimé un besoin de créer un outil **pouvant analyser de manière systématique et automatique des données de la Cour notamment statistiques** pour permettre de dégager plus facilement et rapidement des tendances. Il serait opportun de se rapprocher de nos homologues italiens pour mieux connaître le projet italien « Aut Dedere, Aut Judicare » dont le but est de **fournir des statistiques** dans le domaine de la coopération judiciaire pénale par l'exploitation automatisée de données dans les productions judiciaires (mandats d'arrêts, extraditions, transfèrements...).

Il semble probable que certains des autres besoins exprimés par les représentants des chambres dans le cadre du GT IA sauraient trouver des solutions/pistes de réflexions dans les logiciels déployés par les institutions judiciaires, mais que les publications consultées pour cette note n'ont pas permis de les mettre en évidence : il paraît donc essentiel d'interroger plus en détail nos homologues étrangers afin de creuser davantage les opportunités existantes (zonage automatique, cartographie du litige, analyse de la jurisprudence du fond afin de déterminer des tendances et positions dominantes sur des questions précises, analyse statistique, etc.) ou qui sont à l'horizon car actuellement à l'étude.

ANNEXE 3 : LISTE DES AUDITIONS

COMMISSION EUROPEENNE POUR L'EFFICACITE DE LA JUSTICE (CEPEJ)

- M^{me} Muriel Decot, secrétaire exécutive de la CEPEJ ;
- M. Daniel Schmidt, secrétaire du groupe de travail de la CEPEJ sur la cyberjustice et l'IA ;
- M. Matthieu Quiniou, membre du bureau consultatif sur l'IA de la CEPEJ et maître de conférences à l'Université Paris 8 ;

CONSEIL D'ÉTAT

- M. Thomas Andrieu, président de la huitième chambre du contentieux du Conseil d'État ;

CONSEIL CONSTITUTIONNEL

- M. Jean Maia, secrétaire général du Conseil constitutionnel ;

DIRECTION INTERMINISTÉRIELLE DU NUMÉRIQUE (DINUM)

- Mme Stéphanie Schaer, directrice interministérielle du numérique ;
- M. Pierre Pezziardi, conseiller de la directrice interministérielle du numérique ;
- M. Jules Pondard, superviseur du développement d'Albert API ;

ÉCOLE NATIONALE DE LA MAGISTRATURE

- Mme Alexandra Moreau, coordinatrice de formation pour les magistrats détachés à l'École nationale de la magistrature ;
- M. Tarik Lakssimi, professeur à l'Université de Lyon 2 en détachement à l'École nationale de la magistrature au département recherche et documentation ;

JURIDICTIONS DE L'ORDRE JUDICIAIRE

- M^{me} Isabelle Raynaud-Gentil, magistrate, première avocate générale près la Cour d'appel de Paris ;
- M^{me} Emmanuelle Wachenheim, magistrate, secrétaire générale du premier président de la cour d'appel de Paris ;

LEGALTECH ET EDITEURS JURIDIQUES

- Case Law Analytics :
 - M. Jacques Lévy Véhel, directeur de recherches à l'Inria et co-fondateur de l'outil *Case Law Analytics* ;
- Dalloz :
 - M. Michael Benesty, directeur R&D ;
 - Capucine Hurel, responsable relation publique de Lefebvre Dalloz ;
 - Erwan Royer, directeur éditorial ;
 - Candice Tran Dai, directrice Sécurité des systèmes d'information de l'entreprise Lefebvre Sarrut ;
- Lexbase :
 - M. Fabrizio Papa Techera, directeur Produit & Innovation ;
 - Fabien Girard, directeur de l'Information et du Développement de Lexbase ;
- Lexis Nexis :
 - M. Christophe Bonnet, responsable grands comptes Lexis Nexis ;
 - M. Charles Daneau, responsable satisfaction client ;
 - M. Jean-Pierre Sirot, directeur de produit (CPO) de LexisNexis France ;
- Predictice :
 - M. Louis Larret-Chahine, co-fondateur de la société Predictice, ancien membre du Comité scientifique du pôle numérique du Club des Juristes ;

MINISTERE DE LA JUSTICE

- Direction des Services Judiciaires (DSJ) :
 - Mme Marion Bouhier, cheffe du pôle de l'innovation de la DSJ
 - M. Julien Guet, chef du bureau des infrastructures techniques et de l'innovation de la DSJ ;
 - Mme Agnès TALON, chargé de mission DSJ pour l'*open data* ;
- Secrétariat Général :
 - M^{me} Audrey Farrugia, cheffe du service de l'expertise et de la modernisation au secrétariat général du Ministère de la justice
- M. Yannick Meneceur, magistrat, inspecteur général de la justice, maître de conférences en droit du numérique à la faculté de droit de Strasbourg, ancien membre du Conseil de l'Europe et de la CEPEJ ;

ORDRE DES AVOCATS AU CONSEIL D'ETAT ET A LA COUR DE CASSATION

- Me Thomas Lyon-Caen, président de l'Ordre des avocats au Conseil d'État et à la Cour de cassation
- Me Géraud Mégret, avocat au Conseil d'État et à la Cour de cassation.

PRESIDENCE DE LA REPUBLIQUE

- M. Arno Amabile, conseiller de la Présidence de la République pour le Sommet de l'action sur l'IA

THINK TANK

- M^e Kami Haeri, avocat associé au cabinet White & Case, président de la commission « *Impact de l'intelligence artificielle sur les professions du Droit* » au sein du Club des Juristes.

UNIVERSITAIRES

- M. Samuel Dahan, professeur de droit à la *Cornell Univerity* de New York et la Faculté de droit de l'Université *Queens's* de Kingston, fondateur du *Conflict Analytic Lab* et président du *Deel Lab* ;
- M. Samir Merabet, professeur de droit privé à l'Université des Antilles ;
- M. Frédéric Rouvière, professeur de droit civil et de théorie du droit, directeur du Laboratoire de théorie du droit, du master 2 de théorie du droit et de l'Académie de droit de l'Université d'Aix-Marseille ;
- M. Etienne Vergès, Professeur à l'Université Grenoble Alpes (Centre de Recherches Juridiques)

ANNEXE 4 : GLOSSAIRE

Agent conversationnel (*Chatbot*) : Programme informatique spécialisé dans le langage naturel avec un humain qui est capable de répondre à des questions ou de déclencher l'exécution de tâches. L'agent conversationnel dialogue avec son utilisateur (par exemple, les services de conversation automatisés dans la relation au client).

Agent intelligent : Système autonome utilisant l'intelligence artificielle, capable de percevoir son environnement, d'analyser les données collectées et d'interagir avec celui-ci dans le but de réaliser une tâche prédéfinie.

Ajustement (*Fine Tuning*) : Technique pour adapter un modèle d'intelligence artificielle déjà entraîné à une tâche spécifique. Cela consiste à le réentraîner légèrement – à l'ajuster – avec des données spécifiques pour qu'il réponde mieux aux besoins de cette tâche, tout en utilisant des connaissances déjà apprises.

Algorithme :

- Suite finie et structurée de règles ou d'étapes formelles (opérations logiques, instructions) permettant de résoudre un problème énoncé et d'obtenir un résultat sur la base d'éléments fournis en entrée. Il peut être exécuté de manière automatisée et s'appuyer, le cas échéant, sur des modèles d'apprentissage machine, garantissant une résolution efficace en un nombre fini d'opérations.
- Ensemble d'instructions ou de règles définies utilisées pour résoudre un problème ou effectuer une tâche spécifique.

Annotation (*Data Labeling*) : Tâche qui consiste à attribuer des labels (étiquettes), valeurs, ou des descriptions spécifiques à des données (ex : marquer « pomme » sur l'image d'un panier de fruits). Ce processus permet à l'intelligence artificielle d'apprendre à reconnaître, classer et distinguer les différentes catégories de données et ainsi, d'adapter en conséquence ses paramètres entraînaux afin qu'elle se rapproche le plus possible de la « bonne réponse ».

Anonymisation : Processus consistant à traiter des données à caractère personnel afin d'empêcher totalement et de manière irréversible l'identification d'une personne physique ou morale. Le processus d'anonymisation suppose qu'il n'y ait plus aucun risque de lien possible entre l'information concernée et la personne à laquelle elle se rattache. L'identification devient impossible.

Apprentissage : Phase durant laquelle un système d'intelligence artificielle ajuste ses paramètres en analysant des données d'entraînement, dans le but de développer sa capacité à produire des réponses pertinentes face à des questions futures.

Apprentissage automatique (*Machine learning*) : Approche d'IA qui consiste à construire un modèle algorithmique capable d'apprendre à partir de données sans être explicitement programmé pour chaque tâche. Au fur et à mesure de son apprentissage et

selon la nature des données d'entrée, le modèle d'IA s'ajuste et produit de meilleurs résultats.

Il existe trois types d'apprentissage automatique :

- L'apprentissage **supervisé**. Les données utilisées pour entraîner le modèle d'IA sont étiquetées, c'est-à-dire que des métadonnées leur sont rattachées et fournissent des informations sur leur nature. Le modèle d'IA dispose des données d'entrée et du résultat attendu ; il adapte alors ses paramètres entraînables pour se rapprocher le plus possible de la « bonne réponse ».
- L'apprentissage **non supervisé** : technique d'apprentissage automatique ne recourant pas à des données annotées/étiquetées. L'IA est nourrie de masses de données importantes (*big data*) et détermine elle-même les caractéristiques communes à certaines données, ce afin de les catégoriser. L'apprentissage non supervisé est caractérisé par un plus faible degré d'explicabilité des modèles d'IA.
- L'apprentissage **par renforcement** : le modèle d'IA, nourri de données d'entrée, prend des décisions qui seront sanctionnées positivement et négativement en fonction du résultat attendu. La somme des récompenses permet au modèle de s'approcher d'un comportement décisionnel optimal.

Apprentissage profond (*Deep Learning*) : Sous-ensemble de l'apprentissage automatique qui repose sur l'utilisation de réseaux de neurones artificiels à plusieurs couches. Le nombre de couches, fixé par l'humain, permet de modéliser et de traiter des informations complexes. Pendant l'entraînement, la machine ajuste les poids des neurones (leur pondération) pour réduire les erreurs et améliorer l'exactitude du modèle. Tel est le cas du système Horus de reconnaissance visuelle destiné aux aveugles.

Automatisation robotisée des processus (*Robotic Process Automation*, dit **RPA)** : Ensemble de technologies conçues pour remplacer l'exécution humaine de tâches normées, souvent répétitives, par une exécution automatisée via un logiciel (« ro(bot) »).

Bac à sable : Environnement technique isolé, sécurisé et contrôlé, conçu pour développer, tester et valider un système d'intelligence artificielle sans risque d'impact sur les données en production.

Bande passante : Mesure du débit maximal d'information transmissible entre deux ordinateurs, exprimée en bits par seconde (bps). Une bande passante élevée permet d'échanger un grand volume de données rapidement, tandis qu'une bande passante trop faible peut ralentir les connexions (par exemple, un « wifi qui rame »).

Banc d'essai (*benchmark*) : Comparaison des solutions existantes sur des critères objectifs. Dans le cadre des modèles d'IA, il existe des jeux de données annotés, publics correspondant à des tâches précises (classification de textes, d'images, ...). Pour les comparer, les modèles sont évalués sur ces jeux de données.

Biais algorithmique :

- Un écart systématique par rapport à la réalité ou à l'objectivité dans un système d'intelligence artificielle, influençant ses performances et/ou ses décisions. Cela se produit lorsque les données utilisées pour l'entraîner sont incomplètes, déséquilibrées ou biaisées, conduisant ainsi le modèle à produire des résultats inexacts ou non représentatifs.
- Erreurs systématiques dans un système d'IA qui peuvent conduire à des résultats injustes ou discriminatoires.

Big data (mégadonnées, données massives ou grand jeu de données) : Terme désignant des ensembles de données extrêmement volumineux et complexes qui nécessitent des outils et des méthodes spécifiques pour leur collecte, stockage, traitement et analyse dans de courts délais. Ces données, structurées ou non, sont issues de sources hétérogènes (moteurs de recherche, *open data*, réseaux sociaux, données propriétaires, données achetées commercialement). Pour les données issues de l'activité judiciaire, le *Big Data* peut inclure des données statistiques, des traces de connexion dans les logiciels des métiers (logs des applications), des bases de données des décisions des tribunaux. Par extension, il désigne l'utilisation à grande échelle d'une quantité massive de données pour en extraire des informations stratégiques, notamment par le secteur privé. Il présente des critères de volume, de variété et de vélocité des données.

Chaînage : Processus de raisonnement utilisé en intelligence artificielle permettant d'enchaîner des étapes logiques pour atteindre un but. Il s'appuie sur des règles et des données, et peut fonctionner selon deux mécanismes de base (cf. Chaînage avant et Chaînage arrière). Ces mécanismes sont essentiels pour résoudre des problèmes et imiter le raisonnement humain.

- **Chaînage avant :** Un raisonnement guidé par les données disponibles. A partir d'informations initiales, le moteur d'inférence applique des règles dans le sens « condition vers conclusion », enrichissant à chaque étape les données en mémoire. Dans le domaine médical, le moteur pourrait partir de symptômes observés (fièvre et toux) pour conclure qu'il s'agit peut-être d'une grippe.
- **Chaînage arrière :** Un raisonnement guidé par un objectif ou une hypothèse. Le moteur d'inférence part d'une conclusion potentielle et remonte vers les conditions nécessaires pour la valider ou l'infirmer. Les étapes successives permettent de vérifier si les prérequis sont remplis. Toujours dans le domaine médical, un moteur fonctionnant par chaînage arrière postule que « le patient a une grippe », puis cherche à confirmer cette hypothèse en vérifiant si les symptômes requis (fièvre et toux) sont présents.

Code source : Ensemble de fichiers textes contenant les instructions qui composent un programme sous une forme lisible pour les humains. Il constitue une traduction informatique du modèle algorithmique, servant de base à la création d'un logiciel exécutable.

Code source ouvert (Open Source) : Désigne un logiciel dont le code source est ouvert, autrement dit, librement accessible, permettant à toute personne de l'utiliser, de l'étudier, de le copier, de le modifier et de le redistribuer, y compris sous une version modifiée. Cette liberté est garantie à la fois sur le plan technique et juridique, conformément aux licence *Open Source*.

Contexte : Désigne l'ensemble des données prises en compte par un modèle d'intelligence artificielle lors de la prédiction. Pour associer une catégorie à un mot, on pourra prendre en compte la phrase contenant le mot, les phrases précédentes et les phrases suivantes. Plus la taille du contexte d'un modèle (exprimée généralement en tokens), plus le modèle est en mesure de prendre en compte des éléments de contexte.

Destination (du système d'IA) : Usage prévu pour ce système, incluant sa finalité, son contexte d'utilisation spécifique et les conditions dans lesquelles il est conçu pour fonctionner (ex : un système destiné à traduire en anglais des textes anciens écrits en français).

Données : Représentation d'une information en vue d'un traitement automatique. Les algorithmes ne s'appliquent qu'à des informations qui peuvent être traduites adéquatement en termes de « données » (et s'intégrer au cycle entrée / sortie des algorithmes).

Données à caractère personnel : Toute information concernant une personne physique identifiée ou identifiable (dite « personne concernée »), directement ou indirectement. Parmi celles-ci, les données sensibles concernent les données relatives à des infractions, des procédures et des condamnations pénales et des mesures de sûretés connexes.

Données ouvertes (*open data*) : Désigne la mise à disposition gratuite et publique de bases de données structurées, téléchargeables dans leur intégralité. Ces données peuvent être réutilisées librement, sous réserve de respecter les conditions définies par une licence spécifique. Il ne faut pas confondre l'*open data* avec les informations publiques accessibles individuellement sur des sites internet, mais dont les bases complètes ne sont pas téléchargeables.

Entraînement : Processus par lequel un modèle d'apprentissage automatique apprend à partir de données pour améliorer ses performances.

Exploration de données (*Data Mining*) : Processus d'analyse de grand volume de données pour identifier des modèles, des corrélations, des tendances ou des informations utiles, dans le but d'enrichir les systèmes d'intelligence artificielle, d'améliorer la prise de décision ou de découvrir des connaissances cachées.

Génération augmentée par récupération (*Retrieval Augmented Generation, dit RAG*) : Système d'intelligence artificielle qui combine deux étapes : la récupération d'informations pertinentes à partir d'une base de données ou d'un ensemble de documents (*retrieval*), puis la génération de texte fondé sur ces informations récupérées (*generation*). Cela permet au modèle de produire des réponses plus précises et contextuelles en intégrant des informations à jour et spécifiques aux requêtes posées.

Grand modèle de langage (*Large Language Model*) : Système d'intelligence artificielle reposant sur le *Deep Learning*, entraîné sur une quantité massive de données textuelles. Cette ingestion de données permet au modèle d'apprendre les règles du langage, de maîtriser la syntaxe, l'orthographe et de développer une représentation contextuelle du

langage, ce qui le rend capable de comprendre et de répondre des requêtes complexes. Chat GPT est un exemple de LLM.

Hypertrucage (deepfake) : Technique de synthèse de contenu multimédia reposant sur l'IA. Elle permet de modifier des fichiers image, vidéo ou audio afin de créer des contenus contrefaits.

Inférence : Capacité de l'intelligence artificielle ayant achevé son processus d'apprentissage à raisonner et à tirer des conclusions à partir de données nouvelles.

Intelligence artificielle : Les termes d'intelligence artificielle regroupent l'ensemble des sciences, théories et techniques ayant pour but la conception et le développement de machines capables de reproduire les capacités cognitives d'un être humain.

La notion d'intelligence artificielle comporte un certain nombre de sous-catégories, parmi lesquelles :

- L'intelligence artificielle **connexionniste (réseau de neurones)** : retranscription, dans la forme de l'IA, des phénomènes mentaux par le biais de réseaux de neurones formels. Le système d'IA est alors composé de plusieurs couches : une couche d'entrée qui capte les données brutes, une ou plusieurs couches dites « cachées » qui traitent les informations en fonction d'un ensemble de paramètres et une couche de sortie qui formule une réponse de sortie.

L'IA connexionniste se distingue de l'IA **symbolique**, fondée sur la programmation logique, et dont l'intégralité du fonctionnement est lisible par l'homme, facilitant son explicabilité.

Certains modèles d'IA, dits « **hybrides** », allient la logique inhérente aux IA symboliques avec l'efficacité des méthodes d'apprentissage des IA connexionnistes.

- L'intelligence artificielle **en périphérie (Edge IA)** : système d'IA intégré à un appareil au plus proche de son utilisateur, contrairement aux *Cloud IA* hébergées sur un datacenter distant. Google assistant, Siri ou Alexa sont des exemples d'Edge IA.
- L'intelligence artificielle **générationnelle** : IA dont la finalité est la création de nouveau contenu texte ou multimédia en réponse à des requêtes. Elle est source de défis éthiques et socio-économiques, génératrice de questionnements philosophiques.
- L'intelligence artificielle **frugale** : rattachable à la notion de sobriété, l'IA frugale désigne une méthode de développement de systèmes d'IA prenant en compte différentes problématiques (sociales, environnementales) sans se concentrer exclusivement sur l'efficacité du système. Cela se traduit par une redéfinition des usages et des besoins : nouvelles contraintes dans l'utilisation des ressources (plus petites quantités de données) et compromis sur le niveau de précision des résultats obtenus.

- L'intelligence artificielle **forte** (*artificial general intelligence*) : modèle hypothétique de système d'IA doté de toutes les caractéristiques de l'intelligence humaine. L'IA forte bénéficie d'une autonomie de conscience dans la résolution de problèmes et la méthode d'apprentissage.

L'IA forte se distingue de l'IA **faible** (*narrow AI*), programmée et entraînée pour appliquer l'intelligence exclusivement à la résolution d'un problème spécifique et très performante dans son domaine d'entraînement, bien que moins flexible.

Enfin, l'intelligence artificielle à **usage général**, bien que pensée comme une IA faible destinée à exécuter des fonctions génériques, a pu être utilisée et adaptée à un large éventail d'applications pour lesquelles elle n'a pas été conçue intentionnellement et spécifiquement.

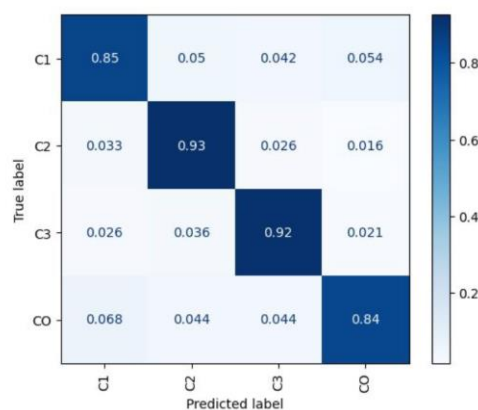
- L'intelligence artificielle d'approche **probabiliste (ou statistique)** : système d'IA dont la finalité est la production de données de sortie assorties de probabilités quant à leur exactitude.

L'IA probabiliste se distingue de l'IA **déterministe**, dont le degré d'inférence est plus élevé. Bien que plus fiable, sa pertinence est limitée : elle permet d'automatiser des tâches répétitives dont le degré d'erreur est moindre.

Interface de programmation d'application (API) : Solution informatique permettant à des applications de communiquer entre elles et de s'échanger des données.

Logiciel : Ensemble des programmes, procédés et règles, et éventuellement de la documentation, relatifs au fonctionnement d'un traitement de données. Le logiciel intègre un ou plusieurs programmes informatiques.

Matrice de confusion : Mode de visualisation des résultats d'un système d'IA afin d'en déterminer le degré d'inférence.



Exemple :

Métadonnées : Données secondaires qui accompagnent des données structurées et fournissent des informations sur celles-ci (date et lieu de création...).

Moteur d'inférence : Logiciel intégrant une solution d'IA permettant à un système expert de conduire des raisonnements logiques et de dériver des conclusions.

Partitionnement de données (*Clustering*) : Technique d'apprentissage, non supervisé, consistant à regrouper des données similaires en sous-ensembles appelés *clusters*. Ces regroupements sont faits en fonction de caractéristiques communes, souvent définies par des critères de proximité ou de similarité, mesurés à l'aide de notions de distance entre les objets. L'objectif est de diviser un ensemble de données en groupes homogènes afin de faciliter leur analyse et leur interprétation.

Processeur graphique, unité de traitement graphique ou carte graphique (*graphics processing unit* ou **GPU)** : Matériel informatique intégré à un système informatique et assurant une fonction de calcul d'image. Son importante force de calcul l'a rendu indispensable à l'exécution des systèmes d'IA connexionnistes, de leur phase d'apprentissage jusqu'à leur inférence.

Profilage : Traitement utilisant les données personnelles d'un individu en vue d'analyser et de prédire son comportement. Le profilage repose sur l'établissement d'un profil individualisé, concernant une personne en particulier : il vise à évaluer certains de ses aspects personnels, en vue d'émettre un jugement ou de tirer des conclusions sur elle.

Pseudonymisation : Pratique consistant à remplacer les données directement identifiantes (nom, prénoms, etc.) d'un jeu de données par des données indirectement identifiantes (alias, numéro séquentiel, etc.).

Requêtes (*prompt*) : Informations données à une IA générative pour qu'elle produise la sortie correspondant à l'attente de son utilisateur.

Règlement sur l'intelligence artificielle (règlement IA ou *IA Act*) : Règlement de l'Union européenne adopté en 2024 et visant à encadrer le développement et l'utilisation des systèmes d'intelligence artificielle en les classifiant par niveau de risque. Les exigences du règlement IA relèvent de la transparence des systèmes d'IA, du respect des droits fondamentaux (données personnelles) mais aussi de la lutte contre la concurrence déloyale et de la protection des droits de propriété intellectuelle.

Scientifique des données (*data scientist* ou ingénieur IA) : Personne chargée de développer, d'entraîner et d'exploiter les systèmes d'IA.

Surapprentissage (*overfitting* ou surentraînement) : Méthode d'apprentissage d'IA qui mène à l'apparition d'un système qui correspond trop précisément à une collection particulière de données utilisées pour l'entraînement. Cette analyse risque de ne pas correspondre à des données utilisées en phase de production et donc de ne pas permettre une utilisation fiable du système d'IA.

Système expert : Logiciel capable de répondre à des questions en reproduisant les mécanismes cognitifs d'un expert dans un domaine particulier. Un moteur d'inférence peut être intégré à un système expert afin d'améliorer son efficacité.

Technologie juridique (*Legal Technology*) : Usage de la technologie pour développer, proposer ou fournir des produits ou des services relatifs au droit et à la justice, ou permettre l'accès des usagers du droit, professionnels ou non, à de tels produits ou services.

Température (d'un modèle d'IA générative) : Mesure de la variabilité des réponses d'un SIA à une même entrée.

Traitement automatique du langage naturel (TAL) ou *natural language processing* (NLP) : Domaine multidisciplinaire impliquant la linguistique, l'informatique et l'intelligence artificielle visant à créer des outils capables d'interpréter et de synthétiser du texte pour diverses applications.

Transformeur (*transformers*) : Type d'architecture de réseau de neurones qui intègre un mécanisme d'attention permettant de prendre en compte un contexte de grande dimension. Principalement utilisé pour générer du texte cohérent et pertinent.



COUR DE CASSATION