

Strasbourg, le 19 décembre 2025

CEPEJ(2025)18Final

**COMMISSION EUROPÉENNE POUR L'EFFICACITÉ DE LA JUSTICE  
(CEPEJ)**

**LIGNES DIRECTRICES SUR L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE  
GÉNÉRATIVE POUR LES TRIBUNAUX**

*Document adopté lors de la 45ème réunion plénière de la CEPEJ  
(Strasbourg, 4 et 5 décembre 2025).*

## Table des matières

<b>Résumé</b> .....	3
<b>1. Introduction</b> .....	4
1.1. Préambule.....	4
1.2. IA générative.....	4
1.3. L'IA générative dans le contexte du secteur judiciaire.....	5
<b>2. Les lignes directrices</b> .....	7
2.1. Aspects opérationnels de l'IA générative dans les tribunaux.....	7
2.1.1. Commencer par identifier le problème, et non la solution technique.....	7
2.1.2. Identifier les domaines adaptés à l'utilisation de l'IA générative.....	7
2.1.2.1. Gestion des affaires et administration de la justice.....	7
2.1.2.2. Gestion des documents et traitement de l'information.....	7
2.1.2.3. Services d'information et assistance aux utilisateurs.....	8
2.1.2.4. Aide à la rédaction et documentation judiciaire.....	8
2.1.3. Être conscient des applications inappropriées.....	8
2.1.4. Préférer des outils personnalisés et sécurisés aux solutions génériques et prêtes à l'emploi.....	9
2.1.5. Garantir une gestion saine dans le déploiement des systèmes d'IA générative	11
2.1.6. Être conscient de la manière dont les parties utilisent l'IA générative.....	13
2.1.7. Sensibiliser et former les utilisateurs de l'IA.....	13
2.2. Aspects normatifs de l'IA générative dans les tribunaux.....	14
2.2.1. Droit à un procès équitable et indépendance judiciaire (article 6 de la CEDH) .	14
2.2.1.1. Garantie d'un procès équitable.....	14
2.2.1.2. Principe de légalité, de sécurité juridique et de prévisibilité.....	15
2.2.1.3. Exclusivité de l'autorité judiciaire.....	16
2.2.1.4. Responsabilité et supervision humaines obligatoires.....	16
2.2.1.5. Indépendance judiciaire et nature non contraignante de l'IA.....	16
2.2.1.6. Transparence et raisonnement adéquat.....	17
2.2.1.7. Droit d'accès à un tribunal.....	17
2.2.2. Droit à la liberté et à la sécurité (article 5 de la CEDH).....	17
2.2.3. Droit à la vie privée et à la protection des données (article 8 de la CEDH).....	17
2.2.4. Droit à un recours effectif (article 13 de la CEDH).....	18
2.2.5. Interdiction de la discrimination (article 14 de la CEDH et protocole n° 12).....	18
2.2.5.1. Principe de non-discrimination.....	18
2.2.5.2. Neutralité et prévention active des biais.....	18
2.2.6. Mesures de protection supplémentaires.....	18
2.2.6.1. Explicabilité et raisonnement transparent.....	18
2.2.6.2. Proportionnalité et prudence dans le déploiement de l'IA générative.....	19
2.2.6.3. Responsabilité de l'État pour les dommages causés par l'utilisation de l'IA générative.....	19
2.2.6.4. Principe procédural de d'équité procédurales et d'efficacité.....	19
<b>3. GLOSSAIRE</b> .....	19

## RESUMÉ

Ces lignes directrices fournissent un cadre pour la mise en œuvre de l'Intelligence artificielle (IA) générative dans l'administration de la justice, en mettant l'accent sur la garantie de la sécurité juridique, l'indépendance judiciaire avec un contrôle humain efficace, la transparence et la traçabilité.

Initialement, les outils à usage général proposés par les grandes entreprises technologiques ne constituent qu'une première étape ; il est ensuite prévu des solutions spécialisées et, le cas échéant, dans des « Large Language Models » (LLM) ou dans des déploiements avec un contrôle souverain sur les données et les infrastructures. Il est essentiel que les outils soient adaptés aux processus judiciaires et déployés sous contrôle public.

La mise en œuvre de ces solutions doit être conforme à la Convention européenne des droits de l'homme (CEDH), à la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle (CETS 225), à la Convention 108+ pour la protection des personnes à l'égard du traitement des données à caractère personnel et à la Charte éthique européenne de la CEPEJ sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement.

Ces lignes directrices reposent sur les principes fondamentaux suivants : l'exercice de l'autorité judiciaire relève de la compétence exclusive des tribunaux ; l'accès effectif à un juge humain est garanti ; les résultats produits par l'IA générative ne sont jamais contraignants ; la transparence doit être assurée quant aux parties de toute décision qui ont été assistées par l'IA ; et l'utilisation de l'IA générative dans les activités judiciaires et la rédaction des décisions doit être transparente. Les exigences en matière de protection de la vie privée imposent la pseudonymisation des données et le contrôle exclusif des données et des infrastructures par l'autorité publique compétente pour la fourniture de services informatiques aux institutions judiciaires. Il est essentiel de prendre en considération la légalité, l'indépendance judiciaire, la non-discrimination, le droit à un recours effectif, ainsi que la traçabilité et l'explicabilité du raisonnement.

Une mise en œuvre progressive est proposée, avec des projets pilotes, des indicateurs et la possibilité de faire marche arrière si nécessaire. Cette approche repose sur le principe de subsidiarité, qui met l'accent sur l'identification des besoins fonctionnels et la prise en compte en priorité des solutions technologiques conventionnelles. L'IA générative ne devrait être déployée que lorsque les solutions alternatives sont jugées inefficaces ou inefficaces.

Son utilisation peut être bénéfique dans divers contextes professionnels, notamment la gestion des dossiers, le traitement des documents, les services d'information publique et l'aide à la rédaction de documents judiciaires courants. Lorsqu'il est utilisé pour aider les autorités judiciaires à rechercher et à interpréter les faits et le droit, à appliquer le droit à des ensembles concrets de faits, ou dans toute autre fonction de soutien à la décision judiciaire, il convient de respecter des exigences strictes. Pour garantir une mise en œuvre réussie, il est essentiel de procéder à des évaluations *ex ante* approfondies des risques et des impacts avant tout déploiement. En outre, la mise en place de mécanismes de suivi et d'observations continus permettra à la fois au tribunal et à l'administration de la justice d'apporter des ajustements et des améliorations. Enfin, la formation de l'ensemble du personnel judiciaire et des utilisateurs du système est essentielle pour garantir l'utilisation efficace de celui-ci.

Enfin, l'État devrait rester responsable de tout dommage résultant de l'utilisation de systèmes d'IA générative dans l'administration de la justice, en veillant à ce que les usagers conservent des voies de recours et des modes de réparation efficaces en cas de préjudice.

# 1. INTRODUCTION

## 1.1. Préambule

1. Les présentes lignes directrices visent à fournir des conseils pratiques sur la manière de mettre en œuvre en toute sécurité l'intelligence artificielle (IA) générative dans l'administration de la justice. Elles s'adressent aux autorités responsables de l'administration de la justice et aux professionnels du droit. Elles devraient être périodiquement revues et mises à jour en fonction des évolutions technologiques récentes.

2. L'utilisation de l'IA dans le domaine de la justice doit respecter les droits garantis par la Convention européenne des droits de l'homme (CEDH), 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 (CETS225), la protection des données à caractère personnel à la lumière de la Convention du Conseil de l'Europe pour la protection des données à caractère personnel (CETS108), ainsi que les principes énoncés dans la Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement (la Charte de la CEPEJ) et, le cas échéant, le règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées en matière d'intelligence artificielle (loi sur l'IA).

3. Ces lignes directrices ont été élaborées par le groupe de travail de la CEPEJ sur la cyberjustice et l'IA (CEPEJ-GT-CYBERJUST) sur la base de la contribution d'un expert<sup>1</sup>. Elles tiennent compte des travaux préliminaires de la CEPEJ, tels que la « Note d'information sur l'utilisation de l'IA générative par les professionnels de la justice dans un contexte professionnel », les « Réflexions du Bureau consultatif de la CEPEJ sur l'intelligence artificielle sur la manière de mettre l'intelligence artificielle au service de la justice » et le « 1<sup>er</sup> Rapport sur l'utilisation de l'intelligence artificielle (IA) dans le système judiciaire basé sur les informations contenues dans le Centre de ressources sur la cyberjustice et l'IA de la CEPEJ » et sont illustrées par des outils présentés dans le Centre de ressources de la CEPEJ sur la cyberjustice et l'IA.

## 1.2. IA générative

4. L'IA générative est une branche de l'IA qui vise à apprendre aux machines à créer de nouveaux contenus imitant de manière convaincante les productions humaines. Au lieu de se contenter de classer ou de prédire des données, ces systèmes apprennent à partir de grands ensembles de données à identifier des modèles et des structures sous-jacents, ce qui leur permet de générer des textes, des images, de la musique, des vidéos et d'autres types de contenus<sup>2</sup>.

5. Dans le domaine de la génération de texte, l'évolution des outils de « traitement du langage naturel » a permis aux logiciels automatisés d'interpréter correctement et de générer le langage de manière naturelle. Auparavant, la compréhension du langage humain représentait un défi pour les logiciels d'IA en raison de son ambiguïté et de sa complexité, mais grâce aux progrès de l'apprentissage profond, les modèles sont désormais capables d'apprendre les modèles de notre langage et de produire des réponses ou des textes cohérents. Cela a entraîné un changement de paradigme dans la manière dont les machines traitent et génèrent du contenu à partir du contenu textuel.

---

<sup>1</sup> Javier Ercilla Garcia, juge, Espagne.

<sup>2</sup> Warankar, M., & Patil, R. (2024). Intelligence artificielle générative. *Revue internationale de recherche scientifique en ingénierie et gestion*, 8(04). [doi :10.55041/IJSREM31146](https://doi.org/10.55041/IJSREM31146)

6. La technologie qui sous-tend l'IA générative repose sur l'apprentissage profond et des architectures de réseaux neuronaux avancées. En substance, ces modèles, appelés *Large Language Models* ou *LLM*, sont entraînés à partir d'importants volumes de données afin d'identifier des schémas complexes (qu'il s'agisse de séquences de mots, de pixels d'images ou de notes de musique) ce qui permet de générer de nouveaux contenus qui imitent les caractéristiques du matériel d'origine.

7. Malgré son efficacité dans un large éventail de tâches et son adoption dans les *chatbots*, l'IA générative présente actuellement des limites, parmi lesquelles les « hallucinations ». Dans la génération de langage naturel, une hallucination est un résultat produit par le modèle sans soutien fidèle dans les données d'entrée ou les connaissances de référence : elle peut contredire explicitement le texte source (intrinsèque) ou ajouter des données impossibles ou fausses (extrinsèque). Il s'agit essentiellement d'informations inventées que le système présente comme des faits<sup>3</sup>.

8. D'autres limites notables concernent les problèmes potentiels de droits d'auteur découlant des données d'entraînement originales de ces modèles, ainsi que la divulgation potentielle non sollicitée de données d'entrée à des tiers et leur réutilisation pour entraîner le modèle d'IA générative, ce qui pourrait constituer une violation de la réglementation applicable en matière de protection des données.

### 1.3. L'IA générative dans le contexte du secteur judiciaire

9. L'utilisation de solutions d'IA générative par les professionnels de la justice est en pleine expansion, que ce soit pour la génération de résumés, la correction de textes, la rédaction de bases juridiques fondées sur la jurisprudence, l'interaction avec les justiciables (explication des jugements aux clients), etc. Selon le rapport *Future Ready Lawyer 2024*<sup>4</sup>, qui a interrogé 712 professionnels du droit aux États-Unis et dans neuf pays européens, 68 % des cabinets d'avocats ainsi que 76 % des services juridiques d'entreprise utilisent l'IA générative au moins une fois par semaine, et plus d'un tiers l'utilisant quotidiennement (35 % dans les services juridiques d'entreprise et 33 % dans les cabinets d'avocats). En Europe, les Pays-Bas sont en tête de l'utilisation quotidienne de l'IA générative avec 42 %, suivis par l'Allemagne avec 38 %. Ces outils sont principalement utilisés pour l'automatisation des documents (67 %), la gestion du temps (66 %) et la planification stratégique (62 %). En outre, 58 % des cabinets d'avocats prévoient d'augmenter leurs investissements dans l'IA au cours des trois prochaines années, tout comme 73 % des services juridiques d'entreprise.

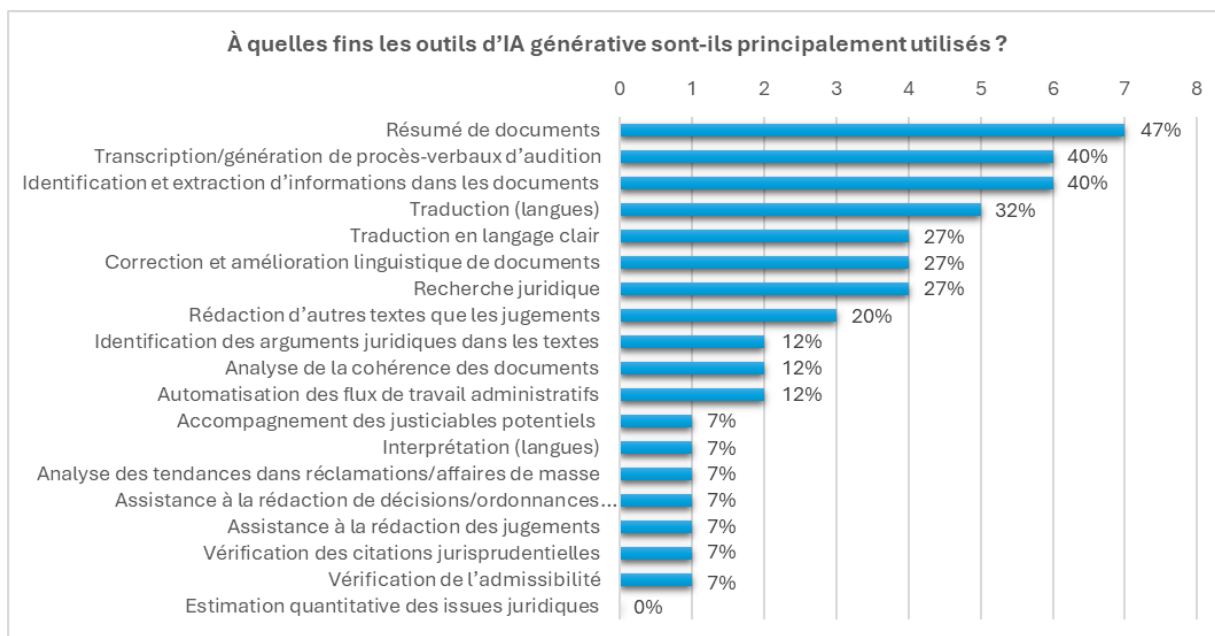
10. Selon une enquête menée en octobre 2025 auprès des membres du Réseau européen de cyberjustice de la CEPEJ, 46 % des répondants<sup>5</sup> ont confirmé l'utilisation de l'IA générative dans les tribunaux, dont 74 % font référence à des applications standard/prêtes à l'emploi (par exemple Microsoft Copilot, Chat GPT, DeepL) et 52 % à des applications personnalisées/sur mesure (les deux options étant possibles). Les utilisations les plus fréquemment mentionnées de l'IA générative concernent la synthèse de documents et d'autres fonctions de soutien éditorial. Aucun répondant n'a mentionné l'utilisation de l'IA générative pour l'estimation quantitative des résultats juridiques (voir figure 1).

---

<sup>3</sup> Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Chen, D., Dai, W., Chan, H. S., Madotto, A., & Fung, P. (2022). Enquête sur les hallucinations dans la génération de langage naturel. arXiv. <https://arxiv.org/abs/2202.03629>

<sup>4</sup> Wolters Kluwer. (2024). Rapport d'enquête 2024 sur les avocats prêts pour l'avenir : Innovation juridique : saisir l'avenir ou prendre du retard ? Wolters Kluwer Legal & Regulatory. <https://www.wolterskluwer.com/en/know/future-ready-lawyer-2024>

<sup>5</sup> n = 33 répondants sur 43 États membres du Conseil de l'Europe interrogés. Décimales arrondies à des nombres entiers.



**Figure 1 : Utilisation de l'IA générative dans les tribunaux.**

11. Plusieurs pays facilitent l'utilisation d'outils d'IA générative dans l'administration publique et les tribunaux grâce à des accords de confidentialité, à la protection des données d'utilisation et à des garanties que ces données ne seront pas réutilisées pour le réentraînement des modèles<sup>6 7</sup>.

12. Dans le domaine juridique, l'utilisation de l'IA générative a donné lieu à des poursuites judiciaires pour fausses citations juridiques<sup>8 9</sup>, qui ont parfois conduit à des sanctions pour ses utilisateurs<sup>10</sup>. Étant donné que les « hallucinations » sont inévitables<sup>11</sup> ou, au mieux, atténuables<sup>12</sup>, en état actuel de la technique, il s'agit d'un risque à prendre en compte lors de l'utilisation de l'IA générative dans le domaine judiciaire, sans garanties adéquates en matière de supervision humaine.

<sup>6</sup> Artificial Lawyer. (24 avril 2025). Les tribunaux britanniques déploient Microsoft Copilot pour les juges et mettent à jour les règles relatives à l'IA générative. <https://www.artificiallawyer.com/2025/04/24/uk-courts-roll-out-microsoft-copilot-for-judges-update-genai-rules/>

<sup>7</sup> El País (2025, 22 September). Do you already know what carpeta.jusitcia.es is? Retrieved from <https://cincodias.elpais.com/extras/2025-09-22/ya-sabes-lo-que-es-carpetajusticiaes.html>

<sup>8</sup> Weiser, B. (27 mai 2023). Voici ce qui se passe lorsque votre avocat utilise ChatGPT. The New York Times. <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>

<sup>9</sup> MIT Sloan Teaching & Learning Technologies. (12 novembre 2024). Quand l'IA se trompe : lutter contre les hallucinations et les biais de l'IA. <https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/>

<sup>10</sup> Carrick, D., & Kesteven, S. (24 juin 2023). Cet avocat américain a utilisé ChatGPT pour faire des recherches dans le cadre d'un mémoire juridique, avec des résultats embarrassants. Nous pouvons tous tirer des leçons de son erreur. ABC News. <https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068>

<sup>11</sup> Xu, Z., Jain, S., & Kankanhalli, M. (2024). L'hallucination est inévitable : une limitation inhérente aux grands modèles linguistiques. arXiv. <https://arxiv.org/abs/2401.11817>

<sup>12</sup> Nie, F., Yao, J.-G., Wang, J., Pan, R., & Lin, C.-Y. (juillet 2019). Une recette simple pour réduire les hallucinations dans la réalisation de surfaces neuronales. Dans les actes de la 57e réunion annuelle de l'Association for Computational Linguistics (pp. 2673-2679). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1256>

## **2. LES LIGNES DIRECTRICES**

### **2.1. Aspects opérationnels de l'IA générative dans les tribunaux**

#### **2.1.1. Commencer par identifier le problème, et non la solution technique**

13. Selon le principe de neutralité technologique, les administrations publiques devraient se concentrer sur les besoins fonctionnels et, sur cette base, sélectionner des solutions technologiques qui peuvent être adaptées au fil du temps. Cette approche vise à minimiser les dépendances technologiques, d'éviter d'imposer des mises en œuvre ou des produits techniques spécifiques et de rester adaptables dans un environnement technologique en rapide évolution. Ce principe devrait s'étendre à l'utilisation de l'IA générative dans l'administration de la justice par le biais d'un **principe de subsidiarité**, selon lequel les administrations de la justice identifient d'abord clairement les besoins fonctionnels et envisagent ensuite différentes solutions technologiques, conventionnelles ou nouvelles, avant d'opter pour la plus appropriée et la moins risquée. Cela ne devrait pas empêcher les tribunaux de s'efforcer en permanence d'améliorer l'efficacité de la justice avec l'appui des technologies lorsque cela est approprié, ni de s'abstenir d'introduire l'IA générative dans des cas d'utilisation bien établis qui répondent à des besoins fonctionnels bien identifiés.

14. Les tribunaux et les administrations de la justice sont confrontés à divers défis, notamment le manque de personnel, le volume élevé de demandes collectives, les tâches répétitives de rédaction, la gestion inefficace des dossiers, les retards procéduraux, les difficultés à trouver la jurisprudence ou les problèmes de communication avec les utilisateurs du système judiciaire et le manque d'interopérabilité entre les systèmes informatiques actuellement utilisés.

15. Une fois les besoins et les problèmes opérationnels identifiés au sein de l'administration, l'étape suivante consiste à sélectionner des solutions techniques appropriées, conformément aux principes de neutralité technologique et de subsidiarité de l'IA.

16. Lors de la sélection des solutions techniques les plus appropriées, il convient de tenir compte de la rentabilité. Cela devrait inclure des facteurs tels que la consommation d'énergie, les mesures de cybersécurité, les licences, la maintenance, les coûts de développement et de conseil, en plus du prix d'achat initial.

#### **2.1.2. Identifier les domaines adaptés à l'utilisation de l'IA générative<sup>13</sup>**

##### **2.1.2.1. Gestion des affaires et administration de la justice**

17. L'IA générative pourrait effectuer une gestion prédictive de la charge de travail en analysant les tendances historiques et créer des référentiels de jurisprudence organisés par type d'affaire et par sujet. Ces outils pourraient faciliter l'identification de schémas argumentatifs dans des demandes similaires, permettant ainsi de détecter les procédures qui peuvent être résolues de manière uniforme lorsqu'elles soulèvent la même question juridique, ou aider à répartir les affaires. L'IA générative pourrait également permettre la reconnaissance automatisée des questions litigieuses en comparant les plaintes et les réponses.

##### **2.1.2.2. Gestion des documents et traitement de l'information**

18. L'IA générative pourrait générer automatiquement des procès-verbaux d'audience, établir des indicateurs thématiques et chronologiques pour des dossiers volumineux et

---

<sup>13</sup> La liste suivante n'est pas exhaustive.

extraire des données pertinentes à partir de documents non structurés. De même, elle pourrait être utilisée pour obtenir des informations clés, telles que des noms, des dates et des montants, et les intégrer automatiquement dans des modèles prédéfinis.

### **2.1.2.3. Services d'information et assistance aux utilisateurs**

19. L'IA générative pourrait permettre la création de chatbots spécialisés par sujet juridique et de systèmes conversationnels d'aide à la procédure pour les citoyens. Ces applications doivent être conçues en accordant une attention particulière à l'exactitude des informations fournies et à la clarté de la communication avec les utilisateurs non spécialisés, en utilisant un langage simple.

### **2.1.2.4. Aide à la rédaction et documentation judiciaire**

20. L'IA générative pourrait être utilisée pour la création automatique de documents procéduraux standard, pour l'aide à la rédaction d'ordonnances et la vérification de la cohérence des documents. Elle pourrait générer des résumés structurés des jugements, des documents procéduraux et des appels, ne nécessitant vérification judiciaire pour s'assurer que le résumé correspond fidèlement à l'original. Elle pourrait faciliter également des documents préparatoires, tels que des rapports, des plans ou des mind-maps qui présentent une vue d'ensemble des points d'accord et de désaccord.

21. En outre, elle pourrait être utilisée pour l'analyse des plaintes et des réponses avec des suggestions de jurisprudence connexe, l'identification et la synthèse de la jurisprudence pertinente pour chaque affaire, l'analyse préliminaire des voies de recours concernant le respect des exigences formelles et techniques, et la mise en relation des arguments entre différents écrits procéduraux, produisant des documents organisés dans lesquels la réponse du défendeur est systématiquement associée à chacun des arguments du plaignant.

### **2.1.3. Être conscient des applications inappropriées**

22. Le règlement (UE) 2024/1689 relatif à l'intelligence artificielle (Loi sur l'IA) établit des interdictions spécifiques concernant les pratiques d'IA qui sont particulièrement pertinentes pour l'administration de la justice et l'IA potentiellement générative. L'article 5 du règlement énonce les pratiques d'IA interdites suivantes, qui peuvent également servir d'illustration pour les pays ne relevant pas de son champ d'application : Systèmes d'IA déployant des techniques subliminales ou manipulatrices (article 5, paragraphe 1, point a)) ; Systèmes d'IA exploitant des vulnérabilités (article 5, paragraphe 1, point b)) ; Systèmes de notation sociale (article 5, paragraphe 1, point c)) ; Systèmes d'évaluation des risques pour prédire les infractions pénales (article 5, paragraphe 1, point d)) ; Systèmes d'extension des bases de données de reconnaissance faciale (article 5, paragraphe 1, point e)) ; Systèmes d'inférence des émotions (article 5, paragraphe 1, point f)) .

23. En outre, bien qu'ils ne soient pas explicitement interdits, la réglementation européenne sur l'IA classe certains systèmes d'IA dans le domaine judiciaire comme présentant un risque élevé dans l'annexe III, en particulier ceux destinés à aider les autorités judiciaires à rechercher et à interpréter les faits et le droit, et à appliquer le droit à des ensembles concrets de faits. Ces systèmes doivent respecter des exigences strictes, mais ne sont pas purement et simplement interdits.

24. Si les LLM peuvent être utiles pour la collecte de preuves au stade de l'enquête, ils ne devraient pas remplacer l'évaluation judiciaire des preuves, une fois qu'elles ont été introduites dans le procès. Bien qu'il existe des courants d'IA logique et neuro-symbolique qui intègrent

des modèles causaux<sup>14</sup>, l'IA générative n'a pas la capacité de « *théorisation prospective et de développement d'une logique causale* »<sup>15 16</sup>.

25. La prise de décision devrait rester sous contrôle humain. Les affaires nécessitant une interprétation normative novatrice ou une analyse de principes juridiques abstraits ne devraient pas être confiées à l'IA, étant donné qu'elle manque de « *compréhension intrinsèque de ce qui est vrai ou faux dans la réalité* »<sup>17</sup>. De même, l'utilisation de l'IA générative devrait être évitée dans les affaires nécessitant une interprétation évolutive des réglementations récentes, dans la mesure où l'IA générative peut montrer des limites dans l'adaptation de sa compréhension à des réglementations innovantes si elle n'a pas été spécifiquement formée à celles-ci<sup>18</sup>.

26. L'analyse de la constitutionnalité ou de la conformité avec les traités internationaux nécessite un niveau d'interprétation axiologique qui dépasse les capacités actuelles de l'IA. Même si cela venait à changer à l'avenir, cette question ne peut être laissée qu'au jugement humain, car elle affecte la manière dont les citoyens ont décidé de se gouverner. Par conséquent, « l'air du temps » influencera le « contrôle de constitutionnalité et de conformité », et cela ne peut être laissé à la discrétion d'une IA.

#### **2.1.4. Préférer des outils personnalisés et sécurisés aux solutions génériques et prêtes à l'emploi**

27. La distinction entre les outils génériques commerciaux d'IA générative et les solutions judiciaires personnalisées d'IA générative est très pertinente pour l'administration de la justice car elle permet de prendre en compte non seulement les spécifications techniques, mais aussi des questions fondamentales telles que la protection des données, l'indépendance judiciaire et l'intégrité des procédures judiciaires.

28. Les outils d'IA générative commerciaux gratuits présentent des risques importants. Lorsque les professionnels du droit utilisent des services d'IA générative non officiels, qu'il s'agisse de solutions commerciales gratuites ou payantes, ils risquent, par inadvertance, de procéder à un transfert international de données, principalement vers des centres de données situés aux États-Unis hébergés par les entreprises fournissant ces services. Ces données sont ensuite utilisées à des fins de réentraînement des modèles, ce qui les soustrait de facto au contrôle administratif judiciaire. De telles pratiques violent les principes fondamentaux de protection des données établis dans la Convention 108+<sup>19</sup> et créent des vulnérabilités inacceptables dans l'écosystème de l'information judiciaire.

29. Les solutions commerciales génériques d'IA générative, même lorsqu'elles sont proposées dans le cadre d'abonnements payants, ne disposent généralement pas de la formation spécialisée nécessaire au raisonnement juridique et à la documentation judiciaire. Ces outils, conçus pour un usage grand public, font preuve d'une compréhension limitée de la terminologie juridique, des exigences procédurales et des schémas de raisonnement nuancés caractéristiques du travail judiciaire. L'absence d'intégration avec des bases de

---

<sup>14</sup> Colelough, B. C., & Regli, W. (2025). L'IA neuro-symbolique en 2024 : une revue systématique. arXiv. <https://arxiv.org/abs/2501.05435>

<sup>15</sup> Felin, T., & Holweg, M. (2024). La théorie est tout ce dont vous avez besoin : IA, cognition humaine et prise de décision. Disponible à l'adresse : <https://ssrn.com/abstract=4737265>.

<sup>16</sup> Li, X., Cai, Z., Wang, S., Yu, K., & Chen, F. (2025). Une enquête sur l'amélioration de la capacité de raisonnement causal des grands modèles linguistiques. arXiv. <https://arxiv.org/abs/2503.09326>

<sup>17</sup> Ibid.

<sup>18</sup> Ercilla García, J. (2025). Justicia automatizada: entre las inteligencias artificiales que fingen y las que persuaden. Lex Social, Revista De Derechos Sociales, 15 (1), 1-39. <https://doi.org/10.46661/lexsocial.11652>

<sup>19</sup> Conseil de l'Europe. (2018). Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel [Convention 108 modernisée], Série des traités du Conseil de l'Europe - n° 223. Strasbourg.

données juridiques vérifiées augmente le risque d'« hallucinations », c'est-à-dire la génération de dispositions juridiques inexistantes ou de jurisprudence erronée, ce qui constitue une menace grave pour la sécurité juridique.

30. En cas de recours à des solutions commerciales, le déploiement de l'IA générative dans l'administration de la justice devrait être exclusivement régi par des accords contractuels formels entre les autorités publiques et les prestataires de services d'IA. En cas de recours à des solutions internes, le déploiement de l'IA générative au sein de l'administration de la justice devrait être encadré par une politique sur l'utilisation de l'IA dans l'administration de la justice<sup>20</sup> ou par un système plus large de sécurité de l'information<sup>21</sup>. Ces contrats ou ces politiques devraient établir des cadres complets traitant des points suivants :

- Des protocoles stricts de gouvernance des données garantissant qu'aucun transfert non autorisé vers des tiers n'aura lieu
- Des procédures obligatoires de destruction des données après des périodes de conservation spécifiées
- L'interdiction explicite d'utiliser des données judiciaires pour le réentraînement de modèles sans autorisation expresse
- Une délimitation claire des responsabilités en cas de violation des données ou de défaillance du système
- La conformité aux réglementations nationales et européennes en matière de protection des données
- Des audits de sécurité réguliers et exigences en matière de tests de pénétration
- Des procédures transparentes de réponse aux incidents
- Des mécanismes permettant une implication continue des parties prenantes (par exemple, consultations publiques, comités de surveillance, évaluations périodiques).

31. Les modèles d'IA générative spécialisés dans le domaine juridique constituent une exigence essentielle pour un déploiement efficace de l'IA dans le contexte juridique. Le langage juridique est un domaine technique qui utilise une terminologie spécialisée souvent très éloignée de l'usage courant. Des termes tels que « *contrepartie* », « *préjudice* », « *ordre public* » ou « *bonne foi* » ont des significations juridiques spécifiques qui diffèrent considérablement de leurs définitions courantes. Cette complexité sémantique nécessite des modèles d'IA générative spécialement formés sur des corpus juridiques provenant de chaque juridiction, afin de garantir une utilisation précise de la terminologie juridique technique et une interprétation contextuelle correcte.

32. Les solutions d'IA générative judiciaires personnalisées peuvent intégrer :

- Une intégration avec les bases de données juridiques officielles et les sources jurisprudentielles vérifiées
- Une formation spécifique à la juridiction, sur les textes juridiques, les décisions de justice et les documents procéduraux
- Des capacités multilingues reflétant la diversité linguistique des systèmes judiciaires européens
- Des modules spécialisés pour différents domaines du droit (civil, pénal, administratif, constitutionnel)
- Des systèmes intégrés de vérification des citations croisant les registres juridiques officiels
- La détection et le signalement automatisés des incohérences juridiques ou des anachronismes potentiels

---

<sup>20</sup> Par exemple, le Comité technique national espagnol pour l'administration judiciaire électronique (2024). Politique sur l'utilisation de l'intelligence artificielle dans l'administration de la justice. Une traduction en anglais est disponible à l'adresse suivante : <https://www.administraciondejusticia.gob.es/cteaje/normativa-complementaria>.

<sup>21</sup> Par exemple, décret royal espagnol 311/2022, qui régit le système de sécurité nationale.

33. La souveraineté des infrastructures est un élément essentiel à prendre en considération. Les systèmes d'IA générative judiciaires devraient fonctionner sur la base d'infrastructures relevant du contrôle exclusif des autorités judiciaires ou des entités gouvernementales lorsqu'ils optent pour des modèles open source fonctionnant sur des serveurs publics, ou en collaboration avec des entreprises fournisseurs lorsqu'ils optent pour des modèles commerciaux.

#### **2.1.5. Garantir une gestion saine dans le déploiement des systèmes d'IA générative**

34. La mise en œuvre de systèmes d'IA générative dans l'administration de la justice devrait suivre une approche globale qui comprend plusieurs éléments clés : une stratégie de déploiement progressive avec des processus de gestion du changement appropriés et des phases pilotes avec des évaluations à des fins d'examen, un examen attentif des modèles de licence (technologie propriétaire ou open source), un soutien continu des fournisseurs et des exigences de formation, des audits réguliers, un suivi et un examen des performances des systèmes d'IA et des risques associés.

35. Afin de garantir une meilleure compréhension de la technologie, de réduire les obstacles potentiels à son adoption et de faciliter la gestion du changement, il est recommandé que la mise en œuvre de l'IA générative dans l'administration de la justice suive une approche progressive par étape et différenciée en fonction de la complexité et de la nature des questions traitées :

**Phase préparatoire** : avant toute mise en œuvre, des audits de préparation institutionnelle devraient être réalisés, une formation de base devrait être dispensée au personnel, des protocoles de sécurité devraient être établis et des comités de surveillance devraient être créés.

**Première phase** : limitée aux fonctions administratives et de soutien de base qui n'ont pas d'incidence directe sur la prise de décision judiciaire. Cette phase devrait se concentrer sur la gestion des documents et le traitement de l'information, comme l'extraction de données pertinentes à partir de documents non structurés, l'obtention d'informations clés (noms, dates, montants) et leur intégration automatique dans des modèles prédéfinis. De même, elle pourrait inclure la gestion de base de la charge de travail grâce à l'analyse des tendances historiques et à l'aide à la répartition des affaires. Des services d'information de base et une assistance aux utilisateurs peuvent également être mis en place, tels que des systèmes d'orientation procédurale pour les usagers utilisant un langage simple et des chatbots spécialisés pour les demandes procédurales simples. Ces informations devraient être complétées par une explication concernant les risques encourus. La transcription non simultanée des audiences pourrait également être envisagée à ce stade. Les tribunaux et l'administration de la justice devraient évaluer l'impact sur la charge de travail, surveiller la satisfaction des utilisateurs et mesurer le temps d'adaptation.

**Deuxième phase** : extension à des fonctions d'analyse et d'aide à la rédaction plus sophistiquées. Cette phase devrait englober des capacités avancées de gestion des documents et de traitement de l'information, notamment la génération automatique des procès-verbaux d'audience et la mise en relation des arguments entre les différents écrits procéduraux. Cette phase pourrait également s'étendre à l'aide à la rédaction et à la documentation judiciaire, avec la création automatique de documents procéduraux standard, l'aide à la rédaction d'ordonnances courantes et la vérification de la cohérence des documents. De même, avec la création de référentiels jurisprudentiels organisés par type d'affaire et par sujet, la reconnaissance automatisée des questions

litigieuses par comparaison des plaintes et des réponses, ainsi que l'établissement d'indicateurs thématiques et chronologiques pour les dossiers volumineux. L'analyse jurisprudentielle et l'aide à la décision devraient être introduites pour l'analyse préliminaire des recours concernant le respect des exigences formelles et techniques. Les tribunaux devraient effectuer des tests A/B pour comparer l'efficacité, évaluer les biais décisionnels et analyser les modes d'utilisation.

**Troisième phase** : introduction de fonctions analytiques et consultatives plus avancées, tout en maintenant un contrôle judiciaire strict. Cette phase nécessite une évaluation positive des phases précédentes, une approbation spécifique pour chaque sous-phase et des protocoles de retour en arrière clairement définis. Cette phase peut être subdivisée en trois sous-phases :

**Sous-phase initiale** : se concentrer sur la recherche jurisprudentielle et l'organisation des documents grâce à l'analyse des plaintes et des réponses avec des suggestions de jurisprudence connexe, et l'identification et l'organisation systématique de la jurisprudence pertinente pour chaque affaire à l'examen judiciaire. L'aide à la rédaction devrait inclure la production de résumés structurés des jugements, des documents de procédure et des recours, nécessitant une vérification judiciaire complète. Des fonctions avancées de gestion des dossiers pourraient faciliter l'identification des schémas argumentatifs dans des demandes similaires et la détection des procédures susceptibles de soulever des questions juridiques similaires.

**Sous-phase intermédiaire** : amélioration de la recherche jurisprudentielle et de l'organisation des documents afin de faciliter la préparation judiciaire, en fournissant des précédents juridiques systématiquement organisés et une jurisprudence pertinente pour examen judiciaire. L'aide à la rédaction pourrait inclure la préparation de documents préparatoires tels que des rapports, des résumés et des mind-maps présentant une vue d'ensemble des points d'accord et de désaccord entre les parties.

**Sous-phase avancée** : dans le cadre de procédures hautement standardisées avec des précédents juridiques bien établis, l'IA générative pourrait apporter son aide en fournissant une analyse complète de la jurisprudence avec reconnaissance des schémas dans des affaires similaires, en identifiant les précédents procéduraux et les incohérences potentielles pour le contrôle judiciaire, et en proposant des projets intégrant la jurisprudence pertinente. De même, elle pourrait aider à rédiger des modèles structurés pour les documents procéduraux courants sur la base des pratiques judiciaires établies, tandis que l'analyse et la prise de décision humaines obligatoires resteraient en place pour toutes les décisions juridiques de fond et l'approbation finale de chaque résultat.

36. Tout ce qui précède devrait être appliqué dans le cadre de projets pilotes dans les procédures de paiement, les demandes d'indemnisation évaluées pour des montants déterminés, les procès-verbaux pour des montants modestes, les amendes administratives standardisées ou des réclamations en matière de droit du travail portant sur des montants non litigieux, dans lesquels toutes les preuves sont documentaires.

37. Les questions touchant aux droits fondamentaux, impliquant des mineurs ou des personnes incapables, ou entraînant des mesures privatives de liberté nécessitent une intervention judiciaire humaine complète, limitant l'IA générative à des fonctions de soutien technique.

### **2.1.6. Être conscient de la manière dont les parties utilisent l'IA générative**

38. Les tribunaux et l'administration de la justice devraient établir des protocoles clairs pour surveiller et gérer l'utilisation de l'IA générative par les praticiens du droit et les parties aux procédures. Cela peut impliquer notamment d'exiger des parties et de leurs représentants légaux qu'ils divulguent l'utilisation de systèmes d'IA générative dans la préparation de documents juridiques, la recherche de jurisprudence ou la compilation de preuves, le cas échéant.

39. Les règles de procédure devraient préciser l'obligation de signaler l'utilisation de l'IA générative dans la préparation des dossiers, avec des sanctions appropriées en cas de non-divulgation ou de recours excessif à des contenus générés par l'IA sans vérification adéquate.

40. Les tribunaux devraient être particulièrement vigilants à l'égard des multiples formes de fautes professionnelles liées à l'IA, notamment : la présentation de preuves fabriquées, de citations juridiques inexistantes ou de contenus générés par l'IA qui n'ont pas été suffisamment examinés par des professionnels qualifiés ; les tentatives de manipulation des systèmes d'IA par le biais d'invites ou d'instructions cachées intégrées dans des documents, telles que du texte invisible, une mise en forme inhabituelle ou des instructions microscopiques intégrées conçues pour compromettre l'analyse assistée par l'IA ; et l'utilisation de l'IA générative pour créer des preuves vidéos ou graphiques frauduleuses. Toutes ces pratiques doivent être réglementées tant sur le plan procédural que pénal.

41. Afin de maintenir l'équité procédurale et l'égalité des armes, les tribunaux devraient veiller à ce que toutes les parties aient un accès égal aux informations relatives à l'utilisation de l'IA générative dans les procédures et qu'aucune partie ne bénéficie d'un avantage indu grâce à une assistance IA non divulguée.

### **2.1.7. Sensibiliser et former les utilisateurs de l'IA**

42. Les juges, les procureurs, le personnel administratif et tout le personnel judiciaire devraient recevoir une formation adéquate pour comprendre, gérer et évaluer de manière critique les outils d'IA générative à leur disposition et ceux disponibles sur le marché. Une formation spécialisée et une sensibilisation accrue au potentiel et aux limites de ces outils garantiront une utilisation éthique et efficace de la technologie, en préservant toujours le jugement humain dans la prise de décision.

43. La mise en œuvre de l'IA dans l'administration de la justice devrait être caractérisée par la transparence et l'automatisation. Pour garantir la certitude et la cohérence, les juges ne devraient pas avoir à effectuer d'opérations spécifiques pour activer le système ; son utilisation devrait être prédéterminée par les lois procédurales.

44. Les compétences fondamentales des juges devraient se concentrer sur le maintien et l'application de leur jugement critique, une qualité inhérente à la fonction judiciaire et attendue de celle-ci, en évaluant rigoureusement chaque résultat fourni par l'IA. Cela implique d'éviter une attitude complaisante ou passive à l'égard des suggestions technologiques et d'adopter plutôt une attitude critique, voire dissidente, à l'égard de certaines alternatives ou propositions générées par ces systèmes.

45. Si les juges doivent être équipés de systèmes d'IA générative avec lesquels ils peuvent interagir directement, ils devraient recevoir une formation spécialisée à leur utilisation (ingénierie des invites, ingénierie du contexte, etc.). Ils devraient acquérir une solide

compréhension technique des limites inhérentes à ces modèles : l'étendue de leurs capacités, les domaines dans lesquels ils ont tendance à échouer, leurs points forts et les types de requêtes qui ne sont pas techniquement viables ou appropriées.

46. Les juges devraient être capables d'analyser et de remettre en question les résultats de l'IA générative, en identifiant les biais, les incohérences ou les limites potentiels dans le raisonnement. Ils devraient élaborer des critères pour évaluer quand et comment il est approprié d'utiliser l'IA générative dans différents types d'affaires et acquérir la capacité de passer efficacement des sources traditionnelles (telles que les bases de données jurisprudentielles et les articles doctrinaux) aux systèmes d'IA générative afin d'obtenir les informations les plus pertinentes et les plus fiables.

47. Les juges devraient être formés pour éviter ce que l'on appelle « l'effet bulle », qui survient lorsqu'un juge, impressionné par l'efficacité et la précision apparentes des systèmes d'IA générative, développe une confiance excessive dans la jurisprudence, les rapports ou les résumés qu'ils fournissent, au risque d'abandonner la consultation indépendante des bases de données jurisprudentielles originales et de négliger les documents sous-jacents.

48. Une politique de formation continue devrait être élaborée pour les nouveaux outils d'IA générative à mesure qu'ils apparaissent, ainsi que pour les progrès réalisés par l'IA elle-même dans différents domaines touchant à la justice (génération d'images, vidéos, voix, etc.). Les institutions judiciaires devraient mettre en place des mécanismes systématiques pour tenir les professionnels de la justice informés des nouvelles capacités de l'IA et de leurs implications potentielles pour l'administration de la justice. Cela devrait inclure la compréhension des nouvelles formes de preuves générées par l'IA, la détection des deepfakes et des contenus synthétiques, et l'adaptation aux méthodes évolutives de manipulation numérique susceptibles d'affecter les procédures judiciaires. Cette formation devrait être régulièrement mise à jour afin de refléter l'état des lieux en matière de technologie d'IA générative et ses applications dans le contexte juridique. Pour garantir la cohérence de l'application, des programmes de formation standardisés pourraient être envisagés.

## **2.2. Aspects normatifs de l'IA générative dans les tribunaux**

49. Le déploiement de systèmes d'IA générative dans le cadre judiciaire nécessite une évaluation approfondie au regard des principes juridiques fondamentaux établis par les cadres juridiques européens et les dispositions constitutionnelles nationales. Afin de garantir le respect de ces principes juridiques, l'outil d'évaluation de la CEPEJ<sup>22</sup> et la méthodologie d'analyse d'impact HUDERIA<sup>23</sup> pourraient être utilisés.

### **2.2.1. Droit à un procès équitable et indépendance judiciaire (article 6 de la CEDH<sup>24</sup>)**

#### **2.2.1.1. Garantie d'un procès équitable**

50. Les systèmes d'IA générative peuvent compromettre l'équité procédurale en raison de processus décisionnels opaques, de recommandations automatisées inexplicables ou

---

<sup>22</sup> Commission européenne pour l'efficacité de la justice (CEPEJ). (4 décembre 2023). Outil d'évaluation pour la mise en œuvre de la Charte européenne d'éthique pour l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement. Conseil de l'Europe.

<sup>23</sup> Conseil de l'Europe, Comité sur l'intelligence artificielle. (28 novembre 2024). Méthodologie pour l'évaluation des risques et des impacts des systèmes d'intelligence artificielle du point de vue des droits de l'homme, de la démocratie et de l'État de droit (méthodologie HUDERIA). Conseil de l'Europe.

<sup>24</sup> Convention de sauvegarde des droits de l'homme et des libertés fondamentales, 4 novembre 1950, ETS n° 5, 213 U.N.T.S. 221. [https://www.echr.coe.int/documents/d/echr/convention\\_ENG](https://www.echr.coe.int/documents/d/echr/convention_ENG)

d'avantages procéduraux inégaux résultant d'un accès différentiel aux outils technologiques avancés entre les parties<sup>25</sup>.

51. Les garanties techniques découlant de ces dispositions devraient inclure la mise en œuvre de systèmes d'IA explicables<sup>26</sup>, la traçabilité complète des processus de raisonnement de l'IA<sup>27</sup>, l'égalité d'accès pour toutes les parties aux résultats générés par l'IA et des mécanismes techniques garantissant une véritable égalité des armes<sup>28</sup>. Les systèmes d'IA qui soutiennent les décisions judiciaires devraient fournir des résultats dont la fiabilité et l'explication peuvent être démontrées, et qui sont compréhensibles pour les experts non techniques. Ils devraient également mettre en œuvre des mesures visant à minimiser les inexactitudes, à garantir que le raisonnement juridique correct est suivi et à protéger la sécurité des données.

#### **2.2.1.2. Principe de légalité, de sécurité juridique et de prévisibilité**

52. Les systèmes d'IA générative qui produisent des « hallucinations juridiques » ou font référence à des dispositions juridiques inexistantes posent de graves risques pour la sécurité juridique et l'État de droit.

53. La mise en œuvre de l'IA ne devrait pas compromettre la prévisibilité et la cohérence des décisions judiciaires. Les sources d'information, les critères et les méthodologies utilisés par les modèles d'IA générative devraient être minutieusement documentés et clairement justifiés afin d'éviter toute incertitude ou arbitraire dans les processus judiciaires.

54. Les garanties techniques devraient inclure la validation automatique des citations juridiques, l'utilisation exclusive de sources juridiques officielles certifiées et des systèmes d'alerte pour les contenus générés qui ne peuvent être vérifiés à l'aide de bases de données juridiques faisant autorité<sup>29</sup>.

---

<sup>25</sup> L'article 6, paragraphe 1, de la CEDH garantit le droit à « *ce que sa cause soit entendue équitablement, publiquement et dans un délai raisonnable, par un tribunal indépendant et impartial, établi par la loi* ».

<sup>26</sup> Niveau de compréhension de la manière dont le système basé sur l'IA est parvenu à un résultat donné, voir ISO/IEC TR 29119-11:2020

<sup>27</sup> Exiger des organisations qu'elles conservent des journaux détaillés des opérations d'IA, y compris les entrées de données, les décisions des modèles, les erreurs et les interactions avec les utilisateurs, voir ISO/IEC 42001:2023.

<sup>28</sup> L'outil d'évaluation de la CEPEJ identifie comme risqué n° 8 : « *Risque d'avantage injustifié pour l'une des parties au procès* ».

<sup>29</sup> L'outil d'évaluation de la CEPEJ identifie comme risqué n° 11 : « *Risque de génération et d'utilisation de dispositions juridiques inexistantes par l'IA générative* ».

### **2.2.1.3. Exclusivité de l'autorité judiciaire**

55. Le pouvoir juridictionnel appartient exclusivement aux juges, conformément aux dispositions constitutionnelles et légales. La prise de décision, le raisonnement juridique et l'évaluation des preuves ne devraient pas être délégués à des systèmes d'IA générative ; ces systèmes ne peuvent être utilisés que pour des tâches auxiliaires ou préparatoires qui soutiennent, mais ne remplacent en aucun cas, le jugement humain de l'instance judiciaire.

### **2.2.1.4. Responsabilité et supervision humaines obligatoires**

56. Les juges conservent la responsabilité ultime et exclusive de toutes les décisions judiciaires, quelle que soit l'assistance technologique utilisée dans le processus décisionnel. Les magistrats devraient comprendre pleinement les limites de tout système d'IA générative utilisé et examiner minutieusement tous les résultats générés par l'IA avant de les intégrer dans des actes ou des décisions judiciaires.

### **2.2.1.5. Indépendance judiciaire et nature non contraignante de l'IA**

57. L'indépendance judiciaire<sup>30</sup> peut être compromise lorsque les juges subissent des pressions directes ou indirectes pour suivre les recommandations générées par l'IA, ce qui limite leur liberté de jugement, leur interprétation juridique et leurs processus décisionnels autonomes. Cette indépendance devrait s'étendre à la liberté de ne pas subir de coercition technologique ou de biais systématique introduit par les recommandations de l'IA. Les juges devraient conserver le droit de refuser de recevoir des suggestions de l'IA.

58. Les recommandations, suggestions ou résultats générés par les systèmes d'IA ne doivent jamais être contraignants pour les juges, ni leur utilisation obligatoire, afin de préserver pleinement l'indépendance judiciaire dans l'interprétation juridique, le raisonnement et la prise de décision. Les juges ne doivent pas être tenus de justifier spécifiquement leur décision de s'écarter des propositions ou recommandations générées par les systèmes d'IA.

59. La conception des modèles d'IA générative ne devrait pas privilégier la maximisation de leurs capacités de persuasion afin de minimiser le risque qu'une influence psychologique indue et un biais d'automatisation aient un effet négatif sur l'impartialité et la liberté judiciaires lors de l'évaluation des avis judiciaires générés par l'IA<sup>31</sup>.

60. Les garanties techniques devraient être proportionnées aux risques du système et cela inclut la conception de modèles d'IA générative strictement comme des outils auxiliaires non contraignants, l'élimination de tout mécanisme obligeant ou contraignant les juges à justifier leur écart par rapport aux suggestions algorithmiques, la réalisation d'évaluations rigoureuses avant toute utilisation déployée d'outils d'IA, la conduite d'audits techniques indépendants et la mise en place de mécanismes et de formations continues visant à prévenir les biais d'automatisation et à préserver le jugement humain autonome<sup>32</sup>.

---

<sup>30</sup> L'article 6, paragraphe 1, de la CEDH exige un jugement rendu par un « tribunal indépendant et impartial établi par la loi ».

<sup>31</sup> Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2024). On the Conversational Persuasiveness of Large Language Models: A Randomized Controlled Trial. Prépublication [arXiv arXiv:2403.14380](https://arxiv.org/abs/2403.14380)

<sup>32</sup> L'outil d'évaluation de la CEPEJ identifie comme risque n° 12 : « Risque de perte de pouvoir et de limitation de la responsabilité du juge par l'utilisation d'une IA non explicable ». Le cadre de HUDERIA permet une évaluation récurrente des impacts négatifs potentiels sur l'autonomie judiciaire.

### **2.2.1.6. Transparence et raisonnement adéquat**

61. L'utilisation de l'IA générative dans tout aspect de la prise de décision judiciaire devrait être clairement documentée<sup>33</sup>. Les parties ont le droit de savoir quels éléments d'une décision judiciaire ont été assistés par des systèmes d'IA générative.

### **2.2.1.7. Droit d'accès à un tribunal**

62. Ce droit fondamental est susceptible d'être compromis par des systèmes automatisés qui remplacent ou entravent l'accès à la prise de décision judiciaire humaine, affectant ainsi le droit au jugement humain dans les affaires judiciaires.

63. Les garanties techniques devraient garantir un accès effectif à un juge humain, limiter l'IA générative à des fonctions strictement d'assistance et interdire l'automatisation totale des décisions judiciaires<sup>34</sup>.

## **2.2.2. Droit à la liberté et à la sécurité (article 5 de la CEDH)**

64. Les décisions affectant la liberté individuelle exigent le plus haut niveau de contrôle et de justification par l'humain. Les systèmes d'IA générative qui influencent la détention, la garde à vue ou les mesures de précaution sans contrôle humain adéquat ni justification transparente présentent des risques importants pour les libertés fondamentales<sup>35</sup>.

65. L'introduction de l'IA dans les fonctions judiciaires devrait suivre une approche progressive et incrémentale, initialement limitée à des tâches administratives ou à des procédures standardisées dont les résultats sont prévisibles. La précision, l'efficacité et la sécurité de l'IA générative devraient être rigoureusement évaluées avant d'être étendues à des fonctions judiciaires plus sensibles.

66. Les garanties techniques devraient inclure des systèmes d'IA générative qui favorisent les alternatives à l'emprisonnement, des explications obligatoires lisibles par l'humain des recommandations générées par l'IA affectant la liberté, et un examen humain complet de chaque décision assistée par l'IA où une privation de liberté pourrait être appliquée.

## **2.2.3. Droit à la vie privée et à la protection des données (article 8 de la CEDH)**

67. La vie privée et la protection des données à caractère personnel devraient être garanties en cas d'utilisation de systèmes d'IA générative. Toutes les technologies et données utilisées devraient rester sous le contrôle exclusif du pouvoir judiciaire, afin d'empêcher tout accès non autorisé, toute fuite ou toute ingérence extérieure et de préserver la souveraineté de l'institution sur les informations. Les données judiciaires devraient rester sous le contrôle exclusif des autorités judiciaires. Les tribunaux devraient conserver la possibilité de changer de fournisseur d'IA sans perte de fonctionnalité, ainsi que le droit de vérifier les algorithmes et les données d'entraînement par le biais de processus d'audit continus.

68. Les garanties techniques peuvent inclure la pseudonymisation locale obligatoire des données à caractère personnel avant leur chargement dans le système d'IA, l'exécution des modèles d'IA sur des infrastructures contrôlées par les pouvoirs publics ou sur une

---

<sup>33</sup> Commission européenne pour l'efficacité de la justice. (31 janvier 2025). Réflexions de l'AIAB sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires (CEPEJ-AIAB(2025)1 Rev5).

<sup>34</sup> L'outil d'évaluation de la CEPEJ identifie comme risque n° 6 : « *Le risque que l'IA remplace l'accès au juge* ».

<sup>35</sup> L'article 5, paragraphe 1, de la CEDH protège le droit à la liberté, sous réserve d'exceptions strictement définies et légales. L'article 5, paragraphe 2, établit le droit d'être informé « *des raisons de son arrestation* ». L'article 5, paragraphe 4, garantit le droit à un contrôle judiciaire de la légalité de la détention.

infrastructure disponible dans le commerce ; dans ce dernier cas, la médiation est assurée par des accords contractuels formels entre les pouvoirs publics et les fournisseurs de services technologiques, et la mise en œuvre de protocoles stricts de protection des données<sup>36</sup>.

#### **2.2.4. Droit à un recours effectif (article 13 de la CEDH)**

69. Les systèmes automatisés basés sur l'IA ne devraient pas entraver l'accès effectif à des recours ou à un contrôle juridictionnel<sup>37</sup>.

70. Les garanties techniques devraient assurer un contrôle humain obligatoire des décisions assistées par l'IA et établir des procédures de recours spécifiques pour contester les décisions automatisées ou « *influencées par l'IA* ».

#### **2.2.5. Interdiction de la discrimination (article 14 de la CEDH et protocole n° 12)**

##### **2.2.5.1. Principe de non-discrimination**

71. Les biais algorithmiques dans les systèmes d'IA générative présentent des risques importants de reproduire ou d'amplifier la discrimination existante à l'égard des caractéristiques protégées, ce qui pourrait violer les principes fondamentaux d'égalité<sup>38</sup>.

72. Les garanties techniques devraient inclure des audits périodiques des biais, le nettoyage et l'équilibrage des données, des conditions de régularisation de l'équité pendant la formation (telles que la parité démographique ou les contraintes d'égalité des chances), des tests contrefactuels/A-B et un suivi continu des indicateurs de discrimination<sup>39</sup>.

##### **2.2.5.2. Neutralité et prévention active des biais**

73. Les systèmes d'IA devraient être évalués périodiquement afin de détecter et d'atténuer les biais, en veillant à ce qu'ils ne perpétuent pas les discriminations existantes ou ne créent pas de nouvelles formes d'inégalité dans la prise de décision judiciaire. Une surveillance systématique des biais pour toutes les caractéristiques protégées est obligatoire.

#### **2.2.6. Mesures de protection supplémentaires**

##### **2.2.6.1. Explicabilité et raisonnement transparent**

74. Les systèmes d'IA utilisés dans le domaine judiciaire devraient fournir des explications claires et compréhensibles sur le cheminement suivi pour chaque résultat ou recommandation, afin que les opérateurs judiciaires puissent les vérifier de manière critique. À cette fin, les systèmes devraient divulguer leur processus d'inférence interne de manière transparente et vérifiable, en garantissant un alignement substantiel et en évitant les

---

<sup>36</sup> L'outil d'évaluation de la CEPEJ identifie comme risque n° 2 : « *Risque de divulgation de données à caractère personnel ou de secrets commerciaux* ». Le cadre d'analyse d'impact sur la vie privée de HUDERIA traite spécifiquement des risques de divulgation non autorisée de données sensibles grâce à sa méthodologie d'analyse des risques COBRA.

<sup>37</sup> L'article 13 de la CEDH établit que « *Toute personne dont les droits et libertés reconnus dans la présente Convention ont été violés, a droit à l'octroi d'un recours effectif devant une instance nationale* ».

<sup>38</sup> L'article 14 de la CEDH établit que « *La jouissance des droits et libertés reconnus dans la présente Convention doit être assurée, sans distinction aucune, fondée notamment sur le sexe, la race, la couleur, la langue, la religion, les opinions politiques ou toutes autres opinions, l'origine nationale ou sociale, l'appartenance à une minorité nationale, la fortune, la naissance ou toute autre situation* ».

<sup>39</sup> L'outil d'évaluation de la CEPEJ identifie comme risque n° 10 : « *Risque de discrimination ou d'amplification de la discrimination* ». Le processus d'engagement des parties prenantes de HUDERIA permet d'identifier les impacts différenciés sur les groupes vulnérables.

simulations de conformité superficielles qui masquent leur logique réelle<sup>40</sup>. Les systèmes devraient être capables d'indiquer les sources spécifiques de chaque déclaration générée. Seuls les modèles interprétables qui révèlent leurs processus décisionnels internes devraient être utilisés.

#### **2.2.6.2. Proportionnalité et prudence dans le déploiement de l'IA générative**

75. L'introduction de l'IA dans les fonctions judiciaires devrait suivre une approche progressive et incrémentale, initialement limitée aux tâches administratives ou aux procédures standardisées dont les résultats sont prévisibles. La précision, l'efficacité et la sécurité de l'IA générative devraient être rigoureusement évaluées avant d'être étendues à des fonctions judiciaires plus sensibles.

#### **2.2.6.3. Responsabilité de l'État pour les dommages causés par l'utilisation de l'IA générative**

76. L'État devrait assumer la responsabilité de tout préjudice causé par des erreurs judiciaires ou des défaillances des systèmes d'IA générative, conformément aux principes existants en matière de responsabilité gouvernementale. Ni l'autonomie ni la complexité technologique de l'IA générative ne devraient être invoquées pour exempter l'État de son obligation de réparation.

#### **2.2.6.4. Principe procédural de d'équité procédurales et d'efficacité**

77. L'introduction de l'IA dans l'administration de la justice ne devrait pas entraîner une détérioration des garanties procédurales ou une atteinte à la protection judiciaire effective. Au contraire, elle devrait renforcer les droits procéduraux et l'efficacité judiciaire. Par exemple, cela devrait améliorer l'accès à la cour, offrir un meilleur accès aux preuves ou améliorer les moyens de mieux présenter son dossier. En outre, la mise en œuvre de l'IA générative devrait se traduire par des améliorations mesurables en termes de temps de réponse, de clarté de la communication, de cohérence des décisions et d'efficacité de la gestion des affaires. Les tribunaux devraient conserver la possibilité de revenir aux systèmes précédents si l'IA générative n'améliore pas objectivement les services judiciaires. Des évaluations régulières devraient permettre de vérifier le respect de ce principe<sup>41</sup>.

### **3. GLOSSAIRE**

#### *Principe de neutralité technologique*

Elle est généralement définie comme la liberté dont jouissent les individus et les organisations de choisir la technologie la plus appropriée et la mieux adaptée à leurs besoins et à leurs exigences en matière de développement, d'acquisition, d'utilisation ou de commercialisation, sans dépendre des connaissances impliquées sous forme d'informations ou de données<sup>42</sup>.

#### *Régulateurs d'équité pendant l'entraînement du modèle*

---

<sup>40</sup> Greenblatt, R., Denison, C., Wright, B., Roger, F., MacDiarmid, M., Marks, S., Treutlein, J., Belonax, T., Chen, J., Duvenaud, D., Khan, A., Michael, J., Mindermann, S., Perez, E., Petri, L., Uesato, J., Kaplan, J., Shlegeris, B., Bowman, S. R., & Hubinger, E. (2024). Alignment faking in large language models [Prépublication arXiv arXiv:2412.14093v2]. <https://doi.org/10.48550/arXiv.2412.14093>

<sup>41</sup> Les indicateurs d'évaluation devraient inclure des enquêtes de satisfaction auprès des utilisateurs, une comparaison des mesures de performance avec les systèmes pré-IA, les commentaires des parties prenantes (juges, personnel judiciaire, avocats et citoyens) et une évaluation des améliorations apportées au calendrier procédural.

<sup>42</sup> Voir : [https://icannwiki.org/Technology\\_neutrality](https://icannwiki.org/Technology_neutrality)

Techniques mathématiques spécifiques telles que :

- Régulateurs de parité démographique : pénalisent le modèle lorsque les taux de prédiction diffèrent considérablement entre les groupes démographiques
- Régulateurs de chances égalisées : garantissent des taux de vrais positifs et de faux positifs similaires entre les groupes protégés
- Contraintes d'équité individuelle : garantissent que des individus similaires reçoivent un traitement similaire
- Débiaisement antagoniste : utilisation de réseaux antagonistes pour supprimer les informations relatives aux caractéristiques protégées des représentations apprises