

4 Décembre 2023



CEPEJ(2023)16final

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

Outil d'évaluation pour l'opérationnalisation de la Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement

Document adopté par la CEPEJ lors de sa 41ème réunion plénière, les 4-5 décembre 2023¹

© CEPEJ, Conseil de l'Europe, 2023. Tous droits réservés. Ce document ne peut être reproduit, traduit, distribué, vendu ou utilisé d'une autre manière sans autorisation écrite préalable. Pour soumettre une demande d'autorisation ou pour toute autre question, veuillez contacter la CEPEJ, Conseil de l'Europe via cepej@coe.int.

Introduction

En 2018, la Commission européenne pour l'efficacité de la justice (ci-après "CEPEJ") a adopté sa Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement (ci-après "la Charte de la CEPEJ").

La Charte de la CEPEJ énonce cinq principes clés qui devraient être respectés par le pouvoir judiciaire dans la conception et l'utilisation de l'intelligence artificielle (ci-après IA) :

- 1) principe de respect des droits fondamentaux : assurer une conception et une mise en œuvre des outils et services d'IA qui soient compatibles avec les droits fondamentaux ;
- 2) principe de non-discrimination : prévenir spécifiquement la création ou le renforcement de discriminations entre individus ou groupes d'individus ;
- 3) principe de qualité et de 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 de manière multi-disciplinaire, dans un environnement technologique sécurisé ;
- 4) principe de transparence, de neutralité et d'intégrité intellectuelle : rendre accessibles et compréhensibles les méthodes de traitement des données, autoriser des audits externes ;
- 5) 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.

La Charte de la CEPEJ est la première étape de l'action de la CEPEJ pour promouvoir une utilisation responsable de l'IA dans les systèmes judiciaires européens, conformément aux valeurs du Conseil de l'Europe.

Les systèmes d'IA destinés à l'administration de la justice peuvent avoir un impact significatif sur la démocratie, l'État de droit, les libertés individuelles ainsi que le droit à un recours effectif et à un procès équitable, et peuvent entraîner des biais, des erreurs et de l'opacité.

Si les systèmes d'IA qui n'affectent pas l'administration effective de la justice peuvent dans certains cas être considérés comme présentant un risque faible, ils nécessitent néanmoins des évaluations appropriées pour s'assurer que leur utilisation n'entraîne pas de dommages inattendus ou non désirés. Lorsque les systèmes d'IA sont destinés à être utilisés par une autorité judiciaire, ou en son nom, pour aider les autorités judiciaires à rechercher et à interpréter les faits et le droit et à appliquer le droit à un ensemble concret de faits, ils ne doivent pas affecter l'indépendance des juges dans leur processus de décision. La prise de décision finale doit rester une activité et une décision pilotées par l'homme. Dans tous ces cas, la conformité éthique de ces systèmes avec les principes de la Charte de la CEPEJ devrait être évaluée.

Conformément à ce qui précède, les discussions au sein de la CEPEJ et avec les partenaires externes ont montré que les décideurs au sein des systèmes judiciaires (c'est-à-dire les représentants des instances (de gestion) judiciaires, les chefs de projets, les responsables informatiques, etc.) pourraient tirer profit de directives plus pratiques quant à l'application des cinq principes énoncés dans la Charte de la CEPEJ. Dès lors, une opérationnalisation détaillée desdits principes a été jugée nécessaire².

² En 2022, le Groupe de travail sur la qualité de la justice (CEPEJ-GT-QUAL) a pris la décision d'avancer dans la conception d'un outil de soutien approprié pour mettre en pratique la Charte de la CEPEJ, en consultation avec le Groupe de travail sur la cyberjustice et l'intelligence artificielle (CEPEJ-GT-CYBERJUST) et dans le cadre du mandat du Conseil consultatif sur l'intelligence artificielle (AIAB) récemment mis en place pour soutenir la CEPEJ sur les

Par conséquent, le présent *Outil d'évaluation pour l'opérationnalisation de la Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement* (ci-après "l'Outil d'évaluation") vise à opérationnaliser ladite Charte de la CEPEJ en prévoyant un ensemble de vérifications, de mesures clés et de garanties que les décideurs au sein des systèmes judiciaires devraient suivre lorsqu'ils achètent, conçoivent, développent, mettent en œuvre et/ou utilisent l'IA dans les systèmes judiciaires et leur environnement, en conformité avec la Charte de la CEPEJ. L'outil d'évaluation aligne également sa logique sur celles des réglementations basées sur la conformité et le risque. En outre, il vise à compléter l'évaluation de l'impact sur les droits humains, la démocratie et l'État de droit (HUDERIA)³, qui est en cours de développement par le Comité sur l'intelligence artificielle (CAI) du Conseil de l'Europe, en ajoutant une série pratique de mesures axées sur le respect éthique appliquées au pouvoir judiciaire. Cette approche devrait garantir l'uniformité de l'identification, de l'analyse et de l'évaluation des niveaux significatifs de risques et de l'évaluation de l'impact des systèmes d'IA sur la jouissance des droits humains, le fonctionnement de la démocratie et le respect de l'État de droit.

Présentation générale de l'outil d'évaluation de la Charte de la CEPEJ

À qui cet outil d'évaluation est-il destiné et qui réalise l'évaluation ?

Cet outil d'évaluation est essentiellement destiné à deux types de décideurs au sein des systèmes judiciaires :

- les représentants des instances (de gestion) judiciaires chargés d'acquiescer des licences de logiciels, y compris pour les systèmes d'IA;
- les responsables informatiques, chefs de projet et/ou représentants occupant des postes de management au sein d'organes judiciaires (de gestion) et supervisant le développement de logiciels, y compris le développement de systèmes d'intelligence artificielle.

L'outil d'évaluation est conçu pour être complété par les décideurs énumérés ci-dessus.

Dans le cas d'un logiciel développé par des tiers (comme dans le cas de l'acquisition d'une licence), les décideurs peuvent demander aux représentants légaux du fournisseur du logiciel de compléter cet outil d'évaluation dans le cadre, par exemple, d'une réponse à un appel d'offres. Ce document, complété par le fournisseur, peut servir de base à une analyse (plus approfondie) par le décideur ou constituer pour le fournisseur une obligation de rendre des comptes.

Quand peut-on utiliser cet outil d'évaluation ?

Cet outil d'évaluation est conçu pour être utilisé avant l'acquisition ou le développement d'un système d'IA par une institution judiciaire. Il peut également être utilisé en cas de mise à jour du système d'IA une fois déployé. En outre, cet outil d'évaluation peut servir de point de référence pour l'idéation et la détermination des spécifications dans la conception d'un système d'IA dans le domaine judiciaire.

aspects techniques de l'intelligence artificielle. Le travail de développement de l'outil a été confié à Alexandra TSVETKOVA et Matthieu QUINIOU, membres de l'AIAB.

³ HUDERIA consiste essentiellement en une obligation de répondre à un certain nombre de questions concernant les contextes de conception, de développement, d'acquisition et d'utilisation ainsi que les impacts potentiels à court, moyen et long terme du système d'IA examiné.

Cet outil d'évaluation suit-il une approche fondée sur les risques ou sur les principes (en particulier en ce qui concerne les droits humains) ?

Cet outil d'évaluation suit une approche mixte. Il est structurellement basé sur les cinq principes définis dans la Charte de la CEPEJ. L'opérationnalisation des principes implique la prise en compte du contexte technique et d'utilisation ainsi que des risques identifiables à cet égard. De nouveaux risques sont susceptibles d'être identifiés avec l'évolution des techniques et des usages de l'IA, ce qui nécessitera une mise à jour de cet outil d'évaluation après son pilotage⁴.

Quels risques ont été pris en compte dans l'outil d'évaluation ?

Les exemples énumérés ci-dessous correspondent aux principaux risques envisagés dans le contexte des cinq principes de la Charte de la CEPEJ, des rapports et ouvrages universitaires qui s'y rapportent.

- 1) Risque de réutilisation des données et/ou d'utilisation d'un modèle d'IA entraîné à d'autres fins

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », en ce qui concerne le traitement proportionné des données à caractère personnel (vie privée) et la clarté des finalités. Il est dû à l'absence de prise en compte de la finalité du traitement des données et à l'inadaptation du modèle d'IA entraîné à d'autres fins ou dans un contexte différent (notamment dans un autre système juridique). Ce risque peut également être observé dans l'influence du premier contexte d'entraînement sur celui de réutilisation, ce qui peut impliquer, par exemple, l'influence d'une culture juridique sur une autre par l'intermédiaire du système d'IA.

- 2) Risque de divulgation de données personnelles ou de secrets commerciaux

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », en ce qui concerne le traitement proportionné des données à caractère personnel (vie privée) et la clarté des finalités. Le risque de divulgation de données à caractère personnel ou de secrets commerciaux peut être causé par des "backdoors", mais aussi par le renforcement d'un modèle d'IA avec des données à caractère personnel ou des secrets commerciaux partagés par l'utilisateur. Ce risque est amplifié par l'absence de précautions visant à informer l'utilisateur des risques liés au partage de données avec le système d'IA.

- 3) Risque de profilage des juges et de "forum shopping"

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant de l'indépendance des juges dans leur processus de prise de décision. Le risque de profilage des juges peut provenir d'une IA généraliste ou d'une IA spécifiquement dédiée au domaine judiciaire. Ce risque de profilage dépend essentiellement des données utilisées pour entraîner l'IA (décisions de justice non anonymisées, activités et publications sur les réseaux sociaux, etc.). Le profilage des juges peut conduire à la récusation d'un juge, à la recherche d'une juridiction statistiquement favorable à la demande judiciaire, ou à l'adaptation de l'argumentation au profil du juge. Les risques de profilage existent également dans une certaine mesure pour les jurés et les parties.

⁴ Voir le point 2 de la Feuille de route révisée pour assurer un suivi approprié de la Charte éthique de la CEPEJ sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement, [CEPEJ\(2021\)16](#).

4) Risque de résultats trompeurs générés par l'IA

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant de l'indépendance des juges dans leur processus de prise de décision. Le risque de résultats trompeurs peut être causé par l'IA qui présente à tort des résultats comme certains (ou avec des pourcentages élevés de fiabilité), ainsi que par les « hallucinations » de l'IA générative. Les conséquences de ce risque peuvent être importantes pour les systèmes d'IA utilisés pour la prise de décision ou le règlement automatisé des litiges en lignes.

5) Risque de manque de clarté des critères et de pondération inadéquate des critères de traitement de l'IA

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant de l'indépendance des juges dans leur processus de prise de décision. Ce risque est lié à une méthodologie inadéquate suivie lors de l'entraînement de l'IA, et à l'introduction de biais difficiles à détecter dans le modèle d'IA. Ce risque est amplifié dans le cas d'une IA non supervisée. Ce risque peut se traduire, par exemple, par des erreurs ou une discrimination amplifiée.

6) Risque de remplacement de l'accès au juge par l'IA

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant du droit d'accès au juge et du droit à un procès équitable. Ce risque peut résulter d'un choix institutionnel, mais aussi d'une interface qui dissuade le justiciable de recourir au juge. Ce risque est consubstantiel aux IA dans le domaine du règlement automatisé des litiges en ligne, et existe également dans une certaine mesure pour les IA d'aide à la décision, si les juges sont incités à ne pas remettre en cause le résultat de l'IA.

7) Risque de fondement flou ou non justifié des décisions

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant du droit d'accès au juge et du droit à un procès équitable. Les causes de ce risque sont essentiellement liées à la méthode d'entraînement de l'IA, inadaptée à l'usage qui en est fait, ainsi qu'au caractère non explicable du système d'IA.

8) Risque d'avantage déloyal pour l'une des parties au procès

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant du droit d'accès au juge et du droit à un procès équitable. Ce risque peut être causé par le coût de la licence et/ou par les compétences requises pour accéder et utiliser un système d'IA qui donne un avantage dans un procès (par exemple, un système d'aide à l'argumentation en temps réel).

9) Risque de violation des droits fondamentaux ou d'arbitrage inapproprié entre deux droits fondamentaux

Ce risque se réfère au premier principe de la Charte de la CEPEJ, à savoir le « Principe de respect des droits fondamentaux », s'agissant des garanties attendues en matière d'éthique et de droits humains dès la conception. Ce risque est transversal et risque d'être particulièrement

important pour les IA de règlement automatisé des litiges en ligne, qui sont appelées à arbitrer entre deux droits fondamentaux.

10) Risque de discrimination ou d'amplification de la discrimination

Ce risque se réfère au deuxième principe de la Charte de la CEPEJ, à savoir le « Principe de non-discrimination », qui vise à éviter les biais et la discrimination fondés sur des données sensibles. Le risque de discrimination est principalement causé par des biais dans le système d'IA. Ces biais peuvent être techniques ou cognitifs et provenir de la sélection des données, de l'annotation ou de l'entraînement de l'IA.

11) Risque de génération et d'utilisation de dispositions légales inexistantes par l'IA générative

Ce risque se réfère au troisième principe de la Charte de la CEPEJ, à savoir le « Principe de qualité et de sécurité ». Ce risque est spécifique aux IA génératives, et il est consubstantiel à leur fonctionnement, même s'il peut être limité par l'utilisation de sources officielles pour renforcer le modèle. Ce phénomène est également décrit comme une « hallucination » de l'IA.

12) Risque de déresponsabilisation et de limitation de la responsabilité du juge par l'utilisation d'IA non explicables

Ce risque se réfère au quatrième principe de la Charte de la CEPEJ, à savoir le « Principe de transparence, de neutralité et d'intégrité intellectuelle », et plus précisément à la nécessité d'une IA qui soit juste, responsable et transparente. Ce risque est un risque institutionnel, qui peut être amplifié par l'utilisation de boîtes noires d'IA qui ne peuvent être expliquées, en particulier si elles sont protégées par des secrets commerciaux.

13) Risque d'utilisation abusive de l'IA

Ce risque se réfère au cinquième principe de la Charte de la CEPEJ, à savoir le « Principe de maîtrise par l'utilisateur ». Ce risque découle de l'utilisation de l'IA à des fins autres que celles envisagées lors de l'entraînement. Ce risque est amplifié par l'absence d'explication à l'utilisateur quant aux possibles utilisations du système d'IA.

14) Risque d'utilisation forcée de l'IA

Ce risque se réfère au cinquième principe de la Charte de la CEPEJ, à savoir le « Principe de maîtrise par l'utilisateur ». Ce risque d'utilisation forcée de l'IA est dû à la manière dont le système d'IA est intégré dans le processus judiciaire, et à un système d'"*opt-out*" inexistant ou difficile d'accès.

Cet outil peut-il être utilisé en complément d'outils d'analyse d'impact et d'autres outils d'évaluation ?

Cet outil d'évaluation aborde les défis liés à la conception et à l'utilisation de l'IA dans le domaine judiciaire, à l'intention des décideurs des institutions judiciaires. Il est par nature spécifique et son utilisation peut être complétée par d'autres outils applicables au déploiement de l'IA et/ou des outils dédiés aux défis liés à la transparence, à la sécurité et à la divulgation des données personnelles.

Un outil complémentaire d'évaluation de l'éthique et des droits humains qui n'est pas spécifique au domaine judiciaire est HUDERIA,⁵ développé par le Comité sur l'intelligence artificielle (CAI) du Conseil de l'Europe (encore en cours de développement à la date d'adoption du présent outil).

Quels sont les domaines d'application concernés ?

Les domaines d'application concernés par l'outil d'évaluation sont ceux répertoriés dans le centre de ressources sur la cyberjustice et l'IA de la CEPEJ⁶ :

- recherche et examen de documents et découverte à grande échelle : *ces solutions créent une collection consultable de descriptions de jurisprudence, de textes juridiques et d'autres informations à partager avec des experts juridiques pour une analyse plus approfondie et une découverte à grande échelle sur des volumes importants de documents électroniques ;*
- règlement automatisé des litiges en ligne : *ces solutions couvrent les technologies utilisées pour le règlement des litiges entre les parties avec une intervention humaine limitée, qui peut être réalisée par le biais de matériel et/ou de logiciels ;*
- prédiction de l'issue des litiges : *ces solutions apprennent, à partir de vastes ensembles de données, à identifier des modèles dans les données qui sont ensuite utilisés pour visualiser, simuler ou prédire l'issue de nouveaux litiges ;*
- aide à la décision et prise de décision : *ces solutions facilitent ou automatisent entièrement les processus de prise de décision dans les systèmes judiciaires ;*
- anonymisation : *ces solutions sont utilisées pour supprimer les informations d'identification telles que les données personnelles des utilisateurs du tribunal ;*
- dépôt électronique : *ces solutions technologiques facilitent l'accès à la justice en établissant un canal numérique qui permet l'interaction et l'échange de données et de documents électroniques entre les tribunaux et les usagers ;*
- triage, attribution et automatisation du flux de travail : *Ces solutions sont utilisées pour faciliter ou compléter certaines tâches et activités au cours du cycle de vie des procédures dans le système de gestion des affaires, en réduisant au minimum la nécessité d'une intervention humaine. Exemples : enregistrement et attribution des affaires judiciaires, attribution de niveaux de priorité aux tâches ou aux personnes afin de déterminer l'ordre le plus efficace pour les traiter ;*
- traitement du langage naturel : *ces solutions sont capables de reconnaître et d'analyser la parole, le texte écrit et de communiquer en retour. Elles sont principalement utilisées dans les tribunaux pour la reconnaissance vocale et la transcription (auditive) des procédures judiciaires ;*
- services d'information et d'assistance : *ces solutions fournissent aux individus des informations sur les services disponibles dans les systèmes judiciaires et les mettent en relation avec les services et les opportunités disponibles.*

Lorsqu'une question figurant dans l'outil d'évaluation ne s'applique qu'à un certain nombre de ces domaines d'application, il est fait explicitement référence aux domaines d'application spécifiques concernés.

Comment utiliser cet outil ?

L'outil d'évaluation se compose de 29 questions conçues pour sensibiliser et identifier les atteintes potentielles d'un système d'IA aux principes de la Charte de la CEPEJ, et en particulier celles relatives aux droits humains et à l'État de droit. Des questions subsidiaires sont également

⁵ <https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688>

⁶ [Centre de ressources sur la cyberjustice et l'IA - Commission européenne pour l'efficacité de la justice \(CEPEJ\) \(coe.int\)](#). Veuillez noter que les domaines d'application sont indicatifs et qu'un chevauchement est possible.

prévues pour affiner les réponses, ou pour vérifier/suggérer la mise en œuvre de mesures d'atténuation ou de gestion des risques.

En outre, dans le cas de réponses susceptibles de porter gravement atteinte aux principes protégés par la Charte de la CEPEJ, la mention « vous devriez reconsidérer le déploiement de ce système d'IA » est affichée pour informer la personne qui remplit le formulaire de l'importance de ce point spécifique.

Une fois l'outil d'évaluation complété par la personne compétente, il est recommandé d'en discuter au sein d'un groupe de supervision afin d'assurer un suivi approprié.

Outil d'évaluation de la Charte de la CEPEJ pour les systèmes d'IA dans le système judiciaire

Description et contexte du système d'IA

1. Qui a développé le système d'IA ?

- Développé en interne par l'institution judiciaire
- Développé par un prestataire de services sur la base du cahier des charges de l'institution judiciaire
- Développé par un tiers fournissant une licence d'utilisation à l'institution judiciaire
- Autres _____

2. Quels sont les domaines d'application concernés par le système d'IA ?

Plusieurs réponses sont possibles.

- Recherche et examen de documents et recherche probatoire à grande échelle (*discovery*)
- Règlement automatisé des litiges en ligne
- Prédiction de l'issue des litiges
- Aide à la décision et prise de décision
- Anonymisation
- Dépôt électronique
- Triage, attribution et automatisation des flux de travail
- Traitement du langage naturel
- Services d'information/assistance

3. Le problème que le système d'IA vise à résoudre (ou le(s) bénéfice(s) qu'il produit) dans le domaine judiciaire est-il clairement défini ?

- Oui Non

Si votre réponse est « non », vous devriez reconsidérer le déploiement de ce système d'IA.

Si votre réponse est « oui », veuillez décrire le problème que le système d'IA vise à résoudre dans l'espace ci-dessous.

4. Le contexte⁷ dans lequel le système d'IA doit être déployé et utilisé est-il clairement défini ?

- Oui Non

⁷ Les risques - et les avantages - de l'IA peuvent résulter de l'interaction entre des aspects techniques et des facteurs sociétaux liés à la manière dont un système est utilisé, à ses interactions avec d'autres systèmes d'IA, à la personne qui l'exploite et au contexte professionnel et/ou social dans lequel il est déployé. Dans le domaine judiciaire, ces facteurs pourraient s'appliquer à la couverture géographique, au type d'affaires et/ou de problèmes à résoudre couverts par l'objectif et la fonctionnalité du système, à la procédure judiciaire, etc.

*Si votre réponse est « non », expliquez pourquoi et vous devriez reconsidérer le déploiement de ce système d'IA.
Si vous avez répondu « oui », veuillez décrire le contexte d'utilisation dans l'espace ci-dessous.*

Principe du respect des droits fondamentaux

Le principe du respect des droits fondamentaux consiste à veiller à ce que la conception et la mise en œuvre des outils et services d'IA soient compatibles avec les droits fondamentaux.

5. Le système d'IA est-il susceptible de permettre ou de faciliter le profilage des juges et/ou des tribunaux ?

Oui Non

5.1. Dans l'affirmative, quelles sont les mesures d'atténuation ou de gestion des risques envisagées ?

Plusieurs réponses sont possibles.

- Anonymisation des données relatives à un juge
 Anonymisation des données relatives à une juridiction et à ses membres
 Exclusion des critères de recherche facilitant un tel profilage
 Autres _____

6. Le système d'IA est-il susceptible de permettre ou de faciliter le profilage des jurés ?

Oui Non

6.1. Dans l'affirmative, le système d'IA est-il accessible à toutes les parties au procès ?

Oui Non

6.2. Dans l'affirmative, quelles sont les mesures d'atténuation ou de gestion des risques envisagées ?

Plusieurs réponses sont possibles.

- Anonymisation des données relatives à un juré
 Exclusion des critères de recherche facilitant un tel profilage
 Autres _____

7. Le système d'IA est-il susceptible de permettre ou de faciliter le profilage des parties à un procès ?

Oui Non

7.1. Dans l'affirmative, quelles sont les mesures d'atténuation ou de gestion des risques envisagées ?

Plusieurs réponses sont possibles.

- Exclusion des données externes au dossier
- Exclusion des données non officielles
- Autres _____

8. (Applicable uniquement aux domaines d'application de l'aide à la décision et de la prise de décision, du règlement automatisé des litiges en ligne, du dépôt électronique et des services d'information/d'assistance) **Existe-t-il une alternative humaine au système d'IA ?**

Oui Non

8.1. Dans l'affirmative, l'utilisateur doit-il indiquer son refus ("*opt-out*") pour accéder à l'alternative humaine au système d'IA ?

Oui Non

8.1.1. Dans l'affirmative, des mesures ont-elles été mises en place pour faciliter l'*opt-out* ?

Oui Non

Si votre réponse est « oui », veuillez décrire les mesures.

9. (Applicable seulement aux domaines d'application du règlement automatisé des litiges en ligne, de la prédiction de l'issue des litiges, de l'aide à la décision et de la prise de décision) **Le système d'IA est-il capable de fournir un fondement juridique systématiquement justifié ?**

Oui Non

9.1. Dans la négative, le système d'IA est-il conçu pour aider les juges à prendre des décisions ?

Oui Non

Si votre réponse est « oui », veuillez vérifier s'il est clairement expliqué aux juges que le système d'IA n'est pas en mesure de fournir un fondement juridique justifié. Une attention particulière doit être accordée au respect du principe « sous le contrôle de l'utilisateur ».

9.2. Dans la négative, le système d'IA est-il conçu pour prendre des décisions ?

Oui Non

*Si votre réponse est « oui », vous devriez reconsidérer le déploiement de ce système d'IA.
Si vous répondez « non », l'objectif de l'IA doit être clairement expliqué à l'utilisateur et un effort particulier doit être fait pour garantir le principe « sous le contrôle de l'utilisateur ».*

10. (Applicable uniquement aux domaines d'application suivants : recherche et examen de documents et découverte à grande échelle, règlement automatisé des litiges en ligne, prédiction de l'issue des litiges, aide à la décision et prise de décision, dépôt électronique et traitement du langage naturel) **Le système d'IA est-il susceptible d'apporter un avantage (par exemple, traitement de données en temps réel lors d'entretiens avec des témoins) à son ou ses utilisateur(s) dans le cadre d'une procédure judiciaire ?**

Oui Non

10.1. Les conditions financières d'accès au système d'IA pourraient-elles potentiellement empêcher certains utilisateurs de l'utiliser ?

Oui Non

Si votre réponse est « oui », veuillez décrire si des mesures spécifiques ont été prises pour tenir compte du contexte et des ressources (par exemple, des réductions offertes à l'institution judiciaire publique, à l'association, à l'avocat commis d'office, etc.)

–

–

–

–

–

11. Le système d'IA est-il susceptible de collecter et/ou de traiter des données personnelles sensibles et/ou des secrets commerciaux ?

Oui Non

11.1. Dans l'affirmative, quelles sont les mesures d'atténuation ou de gestion des risques envisagées pour protéger ces données ou décourager les utilisateurs de les partager ?

Plusieurs réponses sont possibles.

Chiffrement des données

Avertissement contextuel invitant l'utilisateur à ne pas partager des données sensibles

Autres _____

12. (Applicable uniquement aux domaines d'application du règlement automatisé des litiges en ligne, de la prédiction de l'issue des litiges, de l'aide à la décision et de la prise de décision) **Le système d'IA est-il utilisé dans le cadre d'une ou de plusieurs décision(s) liée(s) à la privation de liberté ?**

Oui Non

12.1. Dans l'affirmative, le système d'IA encourage-t-il la libération et les mesures alternatives plutôt que l'incarcération ?

Oui Non

Si votre réponse est « oui », veuillez décrire les mesures d'étalonnage et d'évaluation mises en œuvre.

–

–

–

–

Principe de non-discrimination

Le principe de non-discrimination consiste à empêcher spécifiquement la création ou le renforcement de toute discrimination entre individus ou groupes d'individus.

13. Les données, le modèle et/ou les résultats fournis par le système d'IA ont-ils été audités afin d'identifier les biais ?

Oui Non

Si vous répondez « non », vous devriez envisager un audit dédié aux données, aux modèles et aux résultats avant de déployer le système d'IA.

13.1. Dans l'affirmative, par qui l'audit a-t-il été effectué ?

Plusieurs réponses sont possibles.

- Département informatique de l'institution judiciaire
- Équipe pluridisciplinaire de l'institution judiciaire (informatique, juristes, etc.)
- Tiers proposant le système d'IA
- Audit externe indépendant
- Autres _____

13.2. Dans l'affirmative, quel a été le type d'audit réalisé ?

Plusieurs réponses sont possibles.

- Audit du code source
- Audit d'explication du modèle
- Audit des données
- Test A/B des résultats (en modifiant légèrement le "prompt" (texte en entrée) ou le paramètre)
- Audit UX/UI
- Autres _____

13.3. Dans l'affirmative, des biais ont-ils été identifiés ?

Oui Non

Si votre réponse est « oui », veuillez décrire les biais qui ont été identifiés.

—
—
—
—
—

13.4. Dans l'affirmative, les biais identifiés peuvent-ils conduire à une forme de discrimination ?

Oui Non

13.4.1. Veuillez préciser les types de discrimination concernés:

Plusieurs réponses sont possibles.

- Identité de genre
- Couleur de peau
- Religion
- Opinions politiques
- Origine ethnique/nationale
- Age
- Handicap
- Langue
- Citoyenneté
- Orientation sexuelle
- Autres _____

Veuillez détailler les effets discriminatoires potentiels.

—
—
—
—
—
—
—
—

13.4.2. Dans l'affirmative, le type de biais ou la cause du biais a-t-il été identifié ?

Plusieurs réponses sont possibles.

- Code source, algorithme ou modèle biaisé

- Échantillonnage et données d'entraînement non représentatives
- Biais d'étiquetage des données et des jetons
- Biais de généralisation ou d'association
- Biais de mesure non normalisé
- Biais en faveur de l'expérience utilisateur (UX)
- Biais de l'interface utilisateur (IU)
- Autres _____

Si la cause du biais ne peut être identifiée, vous devriez reconsidérer le déploiement de ce système d'IA.

13.4.3. Dans l'affirmative, des mesures d'atténuation ou de gestion des risques de discrimination ont-elles été mises en œuvre ?

- Oui Non

Si votre réponse est « oui », veuillez décrire les mesures d'atténuation ou de gestion mises en œuvre dans l'espace ci-dessous.

Si votre réponse est « non », veuillez expliquer pourquoi. Dans tous les cas, vous devriez reconsidérer le déploiement de ce système d'IA si aucune autre mesure n'est prise.

Principe de qualité et de sécurité

Le principe de qualité et de sécurité concerne le traitement des décisions juridictionnelles et des données judiciaires, et plus particulièrement l'utilisation de sources certifiées et de données intangibles avec des modèles conçus de manière multi-disciplinaire dans un environnement technologique sécurisé.

14. Le système d'IA est-il basé sur un modèle pré-entraîné par des tiers ?

- Oui Non

14.1. Dans l'affirmative, les bases de données utilisées par ces modèles ont-elles été auditées ?

- Oui Non

Si votre réponse est négative, vous devriez envisager un audit des bases de données utilisées dans ces modèles avant de déployer le système d'IA.

15. **Le modèle d'IA est-il entraîné uniquement à partir de données officielles ou certifiées ?**

Oui Non

Si vous avez répondu « non », veuillez préciser les sources des données non officielles ou non certifiées utilisées pour entraîner le modèle d'IA (y compris les données des utilisateurs ou générées par les utilisateurs en cas de renforcement du modèle d'IA) et expliquer pourquoi ces données sont importantes.

Si l'utilisation de données non officielles ou non certifiées ne peut être justifiée, vous devez reconsidérer l'utilisation de ces données. Si cette dernière ne peut être évitée, vous devriez reconsidérer le déploiement de ce système d'IA.

16. **Existe-t-il des mécanismes de contrôle des procédures et des protocoles permettant de s'assurer que les données fondées sur des procédures judiciaires ne sont pas modifiées avant que le mécanisme d'entraînement du modèle du système d'IA soit alimenté ?**

Oui Non

Si votre réponse est « oui », veuillez décrire les mécanismes, procédures et protocoles de suivi.

Si votre réponse est « non », vous devriez reconsidérer le déploiement de ce système d'IA.

17. **Existe-t-il des mécanismes de contrôle des procédures et des protocoles permettant de garantir que les ensembles de données et les modèles sont stockés et exécutés dans des environnements sécurisés ?**

Oui Non

Si votre réponse est « oui », veuillez décrire les mécanismes, procédures et protocoles de suivi.

Si votre réponse est « non », vous devriez reconsidérer le déploiement de ce système d'IA.

18. **L'entraînement de l'IA (sélection des données, annotation des données, renforcement, etc.) a-t-il été réalisé par une équipe multi-disciplinaire ?**

Oui Non

19. **Les annotations des données et des jetons (par exemple, les mots clés décrivant une décision de justice, etc.) pour l'entraînement de l'IA ont-elles été effectuées avec des professionnels du droit et des experts en éthique ?**

Oui Non

Si votre réponse est négative, vous devriez envisager un audit des ensembles d'entraînement utilisés dans ces modèles avant de déployer le système d'IA.

20. **Un(e) délégué(e) à la protection des droits humains a-t-il/elle été nommé(e) au sein de l'institution judiciaire pour veiller au respect des droits fondamentaux dans le cadre du déploiement et de l'utilisation du système d'IA ?**

Oui Non

Si votre réponse est « non », vous devriez envisager de nommer un expert possédant les compétences nécessaires pour superviser l'utilisation du système d'IA au sein de l'institution judiciaire.

Principe de transparence, de neutralité et d'intégrité

Le principe de transparence, de neutralité et d'intégrité intellectuelle consiste à rendre accessibles et compréhensibles les méthodes de traitement des données et à autoriser des audits externes.

21. **Le code source est-il auditable ? Un secret commercial est-il susceptible d'entraver l'auditabilité du système d'IA ?**

Oui Non

Si votre réponse est « non », vous devriez reconsidérer le déploiement de ce système d'IA s'il est susceptible d'enfreindre le principe du procès équitable et/ou le principe de l'égalité des armes.

22. **Les données utilisées pour entraîner le modèle d'IA sont-elles vérifiables ?**

Oui Non

Si votre réponse est « non », vous devriez reconsidérer le déploiement de ce système d'IA s'il est susceptible d'enfreindre le principe du procès équitable et/ou le principe de l'égalité des armes.

23. **Le modèle d'IA, les critères et les pondérations sont-ils clairs et compréhensibles ?**

Oui Non

En cas de réponse négative, des mesures appropriées doivent être prises pour expliquer et clarifier le modèle d'IA et ses résultats.

24. **Les informations conservées à des fins d'audit, de contrôle et d'examen sont-elles rédigées dans un langage clair, compréhensible et cohérent ?**

Oui Non

Si votre réponse est « non », des mesures appropriées doivent être prises pour sécuriser cette documentation.

Principe de maîtrise par l'utilisateur

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

25. L'utilisateur est-il clairement informé de l'utilisation d'un système d'IA ?

Oui Non

Si votre réponse est « non », cette indication doit être fournie avant le déploiement du système d'IA.

26. Le système d'IA est-il susceptible de fournir des résultats erronés ou trompeurs ?

Oui Non

26.1. Dans l'affirmative, quelles sont les mesures d'atténuation ou de gestion des risques envisagées ?

Recommandations aux utilisateurs sur l'utilisation de l'IA

Explication des résultats fournis par l'IA

Autres _____

27. Existe-t-il une documentation facilement accessible pour former les utilisateurs ?

Oui Non

Si votre réponse est « non », cette documentation doit être mise à la disposition des utilisateurs avant et/ou pendant le déploiement du système d'IA.

27.1. Dans l'affirmative, de quel type de documentation s'agit-il ?

Plusieurs réponses possibles

FAQ

Tutoriel vidéo

Guide interactif

Chatbot

Guide d'utilisation détaillé

Test de maniabilité

Autres _____

28. Existe-t-il des supports de formation et/ou de sensibilisation facilement accessibles sur le système d'IA permettant aux utilisateurs d'acquérir les compétences nécessaires ?

Oui Non

Si votre réponse est « non », ces supports de formation et/ou de sensibilisation doivent être mis à la disposition des utilisateurs avant et/ou pendant le déploiement du système d'IA.

28.1. Dans l'affirmative, quels sont les thèmes abordés ?

Plusieurs réponses sont possibles.

Prise en main du système d'IA et présentation des utilisations possibles

- Aspects techniques de l'IA
- Risques de partialité et de discrimination dans les systèmes d'IA
- Explicabilité et audit de l'IA
- Aspects éthiques de l'IA
- Aspects de l'IA liés aux droits humains
- Cas d'utilisation de l'IA et défis dans le système judiciaire
- Exigences de conformité avec les cinq principes de la Charte de la CEPEJ
- Autres _____

29. Des documents ou des sessions de formation ont-ils été élaborés pour ce système d'IA spécifique ?

- Oui Non