



LIGNES DIRECTRICES POUR
**L'UTILISATION DES SYSTÈMES
D'IA DANS LES TRIBUNAUX**

Publié en 2026 par l'Organisation des Nations Unies pour l'éducation, la science et la culture
7, place de Fontenoy, 75352 Paris 07 SP, France

© UNESCO 2026

ISBN 9789232003737

DOI : 10.58338/IIZT8466



(Euvre publiée en libre accès sous la licence Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Les utilisateurs du contenu de la présente publication acceptent les termes d'utilisation de l'Archive ouverte de libre accès UNESCO (<https://www.unesco.org/fr/open-access/cc-sa>).

Titre original : Guidelines for the Use of AI Systems in Courts and Tribunals

Publié en 2025 par l'Organisation des Nations Unies pour l'éducation, la science et la culture

Les désignations employées dans cette publication et la présentation des données qui y figurent n'impliquent de la part de l'UNESCO aucune prise de position quant au statut juridique des pays, territoires, villes ou zones, ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

Les idées et les opinions exprimées dans cette publication sont celles des auteurs ; elles ne reflètent pas nécessairement les points de vue de l'UNESCO et n'engagent en aucune façon l'Organisation.

Les Lignes directrices ont été préparées par :

Juan David Gutiérrez, PhD, Professeur associé, Université des Andes, Colombie.

Évaluateurs de l'UNESCO :

Cedric Wachholz, Elvis Michel Kenmoe, Guilherme Canela De Souza, Ikran Abdirahman, Jaco Du Toit, Mehdi Benchelah, Prateek Sibal, Kamel El Hilali.

Une liste des évaluateurs externes et des consultations est disponible en annexe (p.31).

Gestion de projet et coordination :

Prateek Sibal, spécialiste de programme, Politiques numériques et transformation numérique, UNESCO.

Ikran Abdirahman, consultant, Politiques numériques et transformation numérique, UNESCO.

Gustavo Fonseca Ribeiro, consultant, Politiques numériques et transformation numérique, UNESCO.

Mei Jie Chen, stagiaire, Politiques numériques et transformation numérique, UNESCO.

Dr. Kamel El Hilali, spécialiste IA et État de droit, Politiques numériques et transformation numérique, UNESCO.

Traduction : Dr. Kamel El Hilali

Photo de couverture – Création graphique – Graphisme de la couverture – Illustrations – Mise en pages :

Abracadabra Estudio de Diseño, México

Impression : UNESCO

Imprimé en France

Citation : UNESCO, Lignes directrices pour l'utilisation des systèmes d'IA dans les tribunaux, 2026



Financé par
l'Union européenne

Les Lignes directrices ont été élaborées dans le cadre du projet financé par la Commission européenne « Soutenir les États membres dans la mise en œuvre de la Recommandation de l'UNESCO sur l'éthique de l'IA grâce à des outils innovants. »

RÉSUMÉ COURT

L'INTELLIGENCE ARTIFICIELLE DANS LES TRIBUNAUX : D'UNE SIMPLE IDÉE À LA RÉALITÉ

L'intelligence artificielle transforme les systèmes judiciaires à travers le monde, offrant de nouvelles possibilités pour améliorer l'accès à la justice, rationaliser la gestion des dossiers et faciliter la prise de décision judiciaire. Cependant, ces innovations soulèvent également des défis complexes en matière d'éthique et de droits humains. Les Lignes directrices de l'UNESCO pour l'utilisation de l'IA dans les tribunaux présentent le premier cadre éthique et opérationnel mondial visant à garantir que l'IA serve la justice tout en respectant l'État de droit et les droits fondamentaux.

Articulées autour de quinze principes universels – notamment la transparence, la responsabilité et le contrôle humain, la protection des droits fondamentaux, et la gouvernance multipartite – ces Lignes directrices fournissent une orientation pratique aux juges, chefs de juridictions et aux décideurs qui envisagent le déploiement de l'IA dans la justice. Elles prônent l'utilisation de l'IA comme un outil d'assistance, et non de substitution, utilisé de manière responsable et toujours sous la supervision des humains.

Complétées par un manuel de renforcement des compétences, les Lignes directrices s'appuient sur des consultations avec des experts de plus de 160 pays, impliquant plus de 36 000 acteurs judiciaires. À mesure que l'IA devient partie intégrante des systèmes judiciaires, cette publication offre des conseils opportuns pour garantir que la technologie renforce la justice, la transparence et la confiance, bâtissant ainsi les défenses de la paix dans l'esprit des hommes et des femmes.

15

principes universels
pour guider l'adoption, la
gouvernance et la supervision
de l'IA dans la justice.



LIGNES DIRECTRICES POUR

L'UTILISATION DES SYSTÈMES D'IA DANS LES TRIBUNAUX

Table des matières

Les lignes directrices en une page	6
Introduction	7
Objectifs des Lignes directrices	12
Méthodologie.....	15
Lignes directrices pour l'utilisation des systèmes d'IA dans les tribunaux	18
1. Principes	18
2. Recommandations spécifiques aux organisations judiciaires ..	22
2.1. Sur le développement, l'acquisition, l'utilisation et l'évaluation des systèmes d'IA	22
2.2. Talent humain	26
2.3. Utilisation des systèmes d'IA générative	27
3. Recommandations spécifiques aux membres individuels du pouvoir judiciaire	28
3.1. Avant d'utiliser les systèmes d'IA.....	28
3.2. Au cours de l'utilisation des systèmes d'IA.....	28
3.3. Au cours de l'utilisation des systèmes d'IA générative	29
Conclusion.....	33
Annexe	34
Notes de fin	36

Les lignes directrices en une page

« Les Lignes directrices énoncent 15 principes universels destinés à guider l'adoption, la gouvernance et le contrôle de l'IA dans le système judiciaire. »

L'intelligence artificielle (IA) est de plus en plus intégrée dans les systèmes judiciaires du monde entier, créant des opportunités pour améliorer l'accès à la justice, la gestion des affaires et la prise de décision. Son utilisation soulève toutefois d'importants défis sur le plan éthique, juridique et des droits humains.

Pour y répondre, l'UNESCO a élaboré les *Lignes directrices pour l'utilisation de l'IA dans les tribunaux*, le premier cadre éthique et opérationnel à vocation mondiale pour l'IA en contexte judiciaire. Ces Lignes directrices visent à garantir que les applications de l'IA servent la cause de la justice tout en protégeant l'État de droit et les droits fondamentaux.

Elles reposent sur quinze principes universellement applicables qui guident l'adoption, la gouvernance et le contrôle de l'IA dans la magistrature : la protection des droits humains, la proportionnalité, la faisabilité des avantages, la sécurité, la sécurité de l'information, l'exactitude et la fiabilité, l'explicabilité, l'auditabilité, la transparence, la sensibilisation et l'utilisation éclairée, la responsabilité et la contestabilité, le contrôle et la prise de décision humaine, la conception centrée sur l'humain et participative, et la gouvernance multipartite. Ensemble, ces principes fournissent des orientations pratiques aux chefs de juridictions, aux juges et aux personnels judiciaires qui envisagent le recours aux solutions d'IA. L'intelligence artificielle n'est pas envisagée comme un substitut au raisonnement judiciaire, mais comme un outil permettant d'élargir l'accès à l'information, d'améliorer l'efficacité et de soutenir une justice équitable dans le cadre d'un contrôle humain et d'un examen éthique rigoureux.

Les Lignes directrices sont complétées par une boîte à outils complète pour le renforcement des capacités. Grâce au programme de l'UNESCO destiné au pouvoir judiciaire sur la liberté d'expression et la sécurité des journalistes, ainsi qu'au programme « IA et État de droit », plus de 36 000 acteurs judiciaires dans plus de 160 pays ont été mobilisés, renforçant ainsi la prise de conscience sur les avantages potentiels et les risques de l'IA. L'élaboration des Lignes directrices a été informée par des consultations avec les parties prenantes à l'échelle mondiale afin de garantir que leurs recommandations reflètent la diversité des traditions juridiques, des capacités technologiques et des besoins sociétaux.

À mesure que l'IA prend de plus en plus d'importance dans les tribunaux, les systèmes juridiques se trouvent à un tournant charnière. Les décisions prises aujourd'hui concernant son déploiement façonneront l'administration de la justice pour les décennies à venir. Que vous soyez un professionnel de la justice, un juriste, un développeur ou un citoyen engagé, les *Lignes directrices* constituent une ressource précieuse pour informer et soutenir l'intégration responsable de l'IA dans les systèmes judiciaires du monde entier.

Introduction

La transformation numérique est en train de remodeler le fonctionnement de toutes les branches du gouvernement. Les gouvernements du monde entier adoptent l'intelligence artificielle (IA) avec des degrés d'intensité variables.¹ Par exemple, des organismes publics en Amérique latine et dans les Caraïbes mettent en œuvre ou testent environ 700 outils d'IA pour remplir diverses fonctions gouvernementales.² Le déploiement de systèmes d'IA peut aider à améliorer la prestation des services publics et à automatiser les processus administratifs, à renforcer l'engagement des citoyens, à surveiller les infrastructures publiques, à optimiser l'allocation des ressources et à soutenir des décisions politiques fondées sur des données probantes dans les secteurs de l'éducation, de la santé, de la mobilité, de l'agriculture, de la sécurité et de l'environnement.

L'adoption d'outils d'IA dans le secteur public exige des cadres de gouvernance qui concilient innovation et protection des droits humains, afin de garantir que le progrès technologique serve les valeurs démocratiques et contribue au développement durable.³

Les tribunaux du monde entier s'intéressent de plus en plus aux outils d'IA (*Boîte n° 1*)⁴ pour accomplir diverses tâches judiciaires, telles que la recherche, l'examen et la découverte de documents à grande échelle ; le règlement automatisé des litiges en ligne ; la prévision d'événements futurs ; les processus décisionnels et judiciaires ; l'anonymisation et la pseudonymisation des décisions de justice ; le classement des documents déposés ; la transcription (voix-texte) et la traduction (voix et texte).⁵

Boîte n° 1 .

CAS D'USAGE DE L'IA PAR LES TRIBUNAUX ET LES CABINETS D'AVOCATS

- La Cour suprême brésilienne a mis en place VICTOR, un système qui traite des milliers de recours introduits devant la Cour et facilite l'identification des affaires qui remplissent la condition préalable de « répercussion générale ».
- La Cour suprême indienne a déployé SUVAS, un logiciel qui traduit des milliers de documents de l'anglais vers dix langues locales et vice versa.
- Au sein de l'Union européenne (UE), les cabinets d'avocats utilisent des outils de traitement du langage naturel pour prédire les décisions judiciaires (par exemple, celles de la Cour de justice de l'UE).
- En Afrique du Sud et au Zimbabwe, les cabinets d'avocats ont adopté des outils d'IA pour l'examen et la gestion des contrats et la recherche de jurisprudence.

Sources : Aletras et al. (2016), Aneja et Mathew (2023), CEPEJ (2024), Kufakwababa (2021), Medvedeva et al. (2020), Ministro do Superior Tribunal de Justiça (2020) et UNESCO (2023).

Ces applications des systèmes d'IA dans le domaine judiciaire ont le potentiel d'améliorer l'efficacité, l'efficience et l'accès à la justice. De plus, l'utilisation d'outils d'IA fiables pourrait réduire le « bruit » dans le système judiciaire. Ce terme désigne la variabilité indésirable des jugements qui devraient être identiques, comme l'ont défini feu le prix Nobel Daniel Kahneman et ses collaborateurs Olivier Sibony et Cass R. Sunstein.⁶

Plus récemment, des juges, des assistants judiciaires, des procureurs et des avocats du monde entier ont commencé à utiliser des chatbots alimentés par des modèles de langage à grande échelle (LLM) pour rédiger des documents juridiques, des décisions judiciaires et élaborer des arguments lors d'audiences (*Boîte n° 2*).⁷ De plus, ces chatbots ont été utilisés par des juges et des tribunaux dans divers domaines du droit : notamment en droit administratif, constitutionnel, de la famille, pénal.⁸

Boîte n° 2

CAS D'USAGE DE L'IA GÉNÉRATIVE PAR LES JUGES ET LES AVOCATS

- En Argentine, les juges ont utilisé des LLM pour résumer leurs décisions dans un langage clair et accessible.
- Des cabinets d'avocats, des sociétés de services juridiques et des universités ont développé, de manière indépendante ou en collaboration avec des entreprises technologiques, des systèmes d'IA générative basés sur des LLM afin de mener des recherches juridiques et des travaux relatifs aux litiges, d'ajouter du contexte à une affaire, de résumer des textes juridiques et de rédiger des courriels et contrats. Ces outils ne se limitent pas aux LLM commerciaux à usage général ; les juges et les avocats peuvent également utiliser des LLM conçus exclusivement pour mener des activités juridiques et des LLM accessibles en open source.
- Néanmoins, des rapports ont fait état, en Australie, au Brésil, au Canada, en Afrique du Sud, en Espagne et aux États-Unis, de juges et d'avocats ayant rendu des décisions judiciaires ou soumis des documents juridiques contenant des références à des décisions inexistantes en raison de l'utilisation inappropriée des chatbots IA.

Sources : Adams (2023), Ambrogi (2023), Bennetts et al. (2023), CEPEJ (2024), Gutiérrez (2024), Harden (2024), LexisNexis (2023), Merken (2025), Morcillo (2025), Taylor (2025), Terzidou (2025), Weiser (2023), Weiser et Bromwich (2023), et Witten (2023).

Cependant, les recommandations officielles destinées aux particuliers ou aux organisations sur l'utilisation adéquate des outils d'IA dans le secteur de la justice sont rares. Seules quelques autorités judiciaires ont publié des principes, des règles ou des lignes directrices sur l'utilisation éthique et responsable des outils d'IA.

Parmi ceux-ci figurent l'Argentine (2024), l'Australie (2024), le Brésil (2020, 2025), le Canada (2023, 2024), la Colombie (2024), la Nouvelle-Zélande (2023), Singapour (2024), le Royaume-Uni de Grande-Bretagne et d'Irlande du Nord (2023), et les États-Unis (2024).⁹ En outre, plusieurs tribunaux ont rendu des décisions établissant des principes et des règles pour l'utilisation des outils d'IA par la justice (*Boîte n° 3*).

Boîte n° 3

DÉCISION DE LA COUR CONSTITUTIONNELLE COLOMBIENNE SUR L'UTILISATION DES SYSTÈMES D'IA PAR LA MAGISTRATURE

- En août 2024, la Cour constitutionnelle de Colombie s'est prononcée sur la constitutionnalité d'une décision rendue en deuxième instance concernant le droit fondamental à la santé d'un mineur. La Cour a examiné si un juge de deuxième instance avait méconnu le droit au procès équitable en intégrant dans sa décision un texte produit à l'aide de ChatGPT. Le contenu créé à l'aide du chatbot consistait en des réponses à quatre questions juridiques que le juge avait transcrites dans le cadre de la motivation de sa décision.
- Dans sa décision, la Cour a cité la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle, en a décrit les valeurs et les principes, et a reconnu que son cadre éthique est « l'un des plus consultés et des plus largement diffusés au monde au sujet de l'IA ». La décision a également cité le manuel mondial de l'UNESCO sur l'IA et l'État de droit pour le

pouvoir judiciaire et a reconnu que ce document « fournit des informations précieuses sur les normes, les principes et la jurisprudence relative au droit international des droits de l'Homme, pertinents pour l'utilisation éthique de l'IA, afin d'atténuer ses risques potentiels pour les droits de l'Homme. »

- La Cour a décidé que l'utilisation des systèmes d'IA par le pouvoir judiciaire est constitutionnelle tant que l'utilisation de la machine ne remplace pas le jugement humain, en particulier en ce qui concerne la prise de décision judiciaire. Les systèmes d'IA peuvent être utilisés par la justice pour soutenir des tâches administratives (par ex., la prise de notes et la rédaction de rapports), des tâches documentaires (par ex., le classement des documents), des fonctions judiciaires (par ex., l'identification de la jurisprudence pertinente), et pour corriger et synthétiser des textes sous la supervision de l'utilisateur.

- De plus, la Cour a ordonné que les membres de la magistrature qui souhaitent utiliser des systèmes d'IA pour soutenir leurs activités judiciaires doivent suivre « les principes de (i) transparence, (ii) de responsabilité, (iii) de confidentialité, (iv) de non-substitution de la rationalité humaine, (v) de sérieux et de vérification, (vi) de prévention des risques, (vii) d'égalité et d'équité, (viii) de contrôle humain, (ix) de réglementation éthique, (x) de conformité aux bonnes pratiques et aux normes collectives, (xi) de surveillance et d'adaptation continues, et (xii) d'adéquation. » (Source : Colombie, Cour constitutionnelle, décision T – 323 / 2024).

- À la suite de cette décision, la Colombie a décidé de mettre en œuvre le projet de lignes directrices de l'UNESCO sur l'utilisation de l'IA dans les tribunaux. Les lignes directrices colombiennes pour l'utilisation responsable et sûre de l'IA générative dans le secteur judiciaire ont été adoptées en décembre 2024.

Les nouvelles lois sur l'IA renforcent le besoin d'orientation. Par exemple, le règlement sur l'IA adopté par l'Union européenne classe les systèmes d'IA « destinés à être utilisés par les autorités judiciaires ou en leur nom, pour les aider à rechercher et à interpréter les faits ou la loi, et à appliquer la loi à un ensemble concret de faits, ou à être utilisés de manière similaire lors du règlement extrajudiciaire d'un litige » comme « à haut risque ». ¹⁰ Cette classification emporte des obligations pour les déployeurs de ces systèmes d'IA, telles que la mise en place d'un système de gestion des risques et la garantie d'une supervision humaine des outils. ¹¹

De plus, les systèmes d'IA intégrés dans les processus décisionnels peuvent avoir des effets en cascade sur l'ensemble du système judiciaire. D'une part, de tels impacts peuvent constituer des externalités négatives et des violations des droits humains. Par exemple, leur déploiement pourrait contribuer à une discrimination systématique à l'égard des personnes ou des groupes vulnérables ou marginalisés, ou à une restriction de l'accès à l'écosystème numérique et aux appareils électroniques. ¹² D'autre part, les systèmes d'IA destinés à la gestion des tribunaux peuvent bénéficier aux justiciables en garantissant que les procès se déroulent dans un délai raisonnable. Ces systèmes d'IA peuvent faciliter les activités préalables au procès (par ex., l'automatisation du système de dépôt et d'archivage des tribunaux), les audiences (par ex., la traduction automatique) et les procédures postérieures au jugement (par ex., l'anonymisation des décisions judiciaires). ¹³ Par conséquent, de nouvelles mesures sont nécessaires aux niveaux organisationnel et individuel pour prévenir et atténuer les effets négatifs et améliorer l'administration de la justice.

En outre, les systèmes d'IA constituent un ensemble de technologies susceptibles de transformer et de bouleverser la pratique du droit et l'élaboration des décisions de justice.¹⁴ Les changements qui en résultent dans le secteur de la justice ne doivent pas compromettre les droits humains. La résolution 78/265 adoptée par l'Assemblée générale des Nations Unies (ONU) le 21 mars 2024 a souligné qu'il faut « respecter, protéger et promouvoir les droits humains et les libertés fondamentales tout au long du cycle de vie des systèmes d'intelligence artificielle ». Par conséquent, le pouvoir judiciaire, et plus généralement les États membres de l'ONU, devraient « s'abstenir ou (...) cesser de se servir des systèmes

d'intelligence artificielle qu'il est impossible d'utiliser dans le respect du droit international des droits humains ou qui présentent des risques excessifs pour l'exercice des droits humains, en particulier pour les personnes en situation de vulnérabilité ».¹⁵

Enfin, l'utilisation des systèmes d'IA ne devrait pas entraver la compétence, l'indépendance et l'impartialité de la justice car, comme l'indiquent les principes de Bangalore sur la déontologie judiciaire, ces attributs sont essentiels à la protection des droits humains et parce que « la mise en œuvre de tous les autres droits dépend en dernier ressort d'une bonne administration de la justice ».¹⁶



Les Lignes directrices proposent des principes et des recommandations pour adopter, déployer et utiliser des outils d'IA afin de soutenir l'administration de la justice tout en respectant les droits humains.

Objectifs des Lignes directrices

Les *Lignes directrices de l'UNESCO pour l'utilisation des systèmes d'IA dans les tribunaux* (ci-après dénommées « *Lignes directrices* ») fournissent des orientations sur les mesures que le secteur judiciaire pourrait envisager pour renforcer ses capacités dans le contexte de la transformation numérique. Les *Lignes directrices* proposent des principes et des recommandations pour adopter, déployer et utiliser des outils d'IA afin de soutenir l'administration de la justice. Les *Lignes directrices* fournissent des orientations tant au niveau organisationnel (tribunal) qu'individuel (juges, magistrats) (NB : ci-après, nous désignerons collectivement ces organisations et individus sous le terme « pouvoir judiciaire »).

Bien que ces *lignes directrices* soient spécifiquement conçues pour guider le pouvoir judiciaire, elles pourraient être pertinentes à l'ensemble de la profession juridique. En ce sens, la plupart des principes et recommandations proposés pourraient être adaptés aux procureurs, arbitres,¹⁷ avocats, associations du barreau, fonctionnaires et chercheurs. Dans ce contexte, des organisations de la société civile telles que les associations professionnelles et les universités ont récemment publié des directives et recommandations similaires.¹⁸

Boîte n° 4

DÉFINITIONS CLÉS

- Système d'IA : systèmes informatiques capables de « traiter les données et l'information par un processus s'apparentant à un comportement intelligent, et comportant généralement des fonctions de raisonnement, d'apprentissage, de perception, d'anticipation, de planification ou de contrôle » (UNESCO 2022, 10).
- Système d'IA générative : systèmes informatiques « qui communiquent en langage naturel, sont capables de répondre à des questions relativement complexes et peuvent créer du contenu (fournir un texte, une image ou un son) à la suite d'une question ou d'instructions formulées (invites) » (CEPEJ-GT-CYBERJUST 2024, 2).
- Membres individuels de la magistrature : magistrats, juges, officiers judiciaires et personnel de soutien judiciaire.
- Opérateur judiciaire : professionnel du droit impliqué dans l'administration de la justice ou qui joue un rôle juridique actif dans une procédure judiciaire, tel que les juges, le personnel de soutien judiciaire, les procureurs et les avocats.
- Modèles linguistiques (LMs) : modèles d'IA générative « qui traitent des entrées textuelles, appelées invites, et génèrent des sorties textuelles à partir de celles-ci. Leurs entrées, ainsi que leurs sorties, peuvent se présenter sous différents formats textuels, tels que le langage naturel, le texte tabulé, voire le code de programme » (BSI 2024, 7).

- Cycle de vie des systèmes d'IA : cycle qui comprend les étapes de pré-conception, de conception, de développement, d'évaluation, de test, de déploiement, d'utilisation, de vente, d'approvisionnement, d'exploitation et de mise hors service (Assemblée générale des Nations Unies, résolution 78/265).
- Apprentissage automatique (ML) : « ensemble de techniques qui permettent aux machines d'apprendre automatiquement à partir de modèles et de déductions plutôt que d'instructions directes données par une personne » (UNESCO 2023, 21).
- Traitement du langage naturel (NLP) : « technique d'apprentissage automatique qui analyse de grandes quantités de données textuelles ou vocales humaines (transcrites ou acoustiques) à la recherche de propriétés spécifiques, telles que le sens, le contenu, l'intention, l'attitude et le contexte » (UNESCO 2023, 21).
- Organisations judiciaires : organismes qui régissent le pouvoir judiciaire, les cours et les tribunaux.
- Systèmes d'IA sûrs, sécurisés et fiables : systèmes d'IA « centrés sur l'humain, fiables, explicables, éthiques et inclusifs et pleinement ancrés dans le respect, la promotion et la protection des droits humains et du droit international, garantissent le respect de la vie privée, sont axés sur le développement durable et sont responsables » (Assemblée générale des Nations Unies, résolution 78/265).

De plus, les *Lignes directrices* comprennent des recommandations spécifiques pour les systèmes d'IA générative. Ces outils peuvent contribuer à améliorer la qualité du travail du pouvoir judiciaire en facilitant la gestion des tribunaux, la rédaction de documents, l'automatisation des tâches et le soutien aux processus décisionnels. En outre, les LLM peuvent faciliter l'interaction entre les ordinateurs et les utilisateurs grâce au langage naturel.

Cependant, l'utilisation abusive des systèmes d'IA pourrait nuire à la confiance de la société dans le système judiciaire. Les impacts négatifs potentiels exigent que le pouvoir judiciaire mette en place des cadres de responsabilité avant et après le déploiement des systèmes d'IA.

Par conséquent, les *Lignes directrices* proposent des orientations spécifiques pour garantir l'intégration par la justice de systèmes d'IA éthiques et respectueux des droits humains.



Les Lignes directrices constituent un document vivant qui devra être mis à jour périodiquement afin de tenir compte des progrès rapides de la technologie, en particulier de l'IA, et de ses applications croissantes dans les systèmes judiciaires.

Méthodologie

Les *Lignes directrices* s'appuient sur le travail approfondi mené par l'UNESCO auprès des acteurs judiciaires au cours de la dernière décennie. Elles s'appuient sur des documents de référence produits par l'UNESCO, un processus de consultation publique, puis un processus d'évaluation par les pairs du projet de lignes directrices mené par des experts juridiques internationaux en IA, des sources secondaires et de la littérature spécialisée.

Premièrement, depuis 2013, l'UNESCO a formé plus de 36 000 acteurs judiciaires dans 160 pays sur des sujets divers tels que la liberté d'expression, l'accès à l'information, l'application et l'impact des systèmes judiciaires basés sur l'IA. Ces acteurs judiciaires ont reçu une formation à travers une « série de cours en ligne ouverts à tous (MOOC), des formations et ateliers destinés aux juges de la Cour suprême, ainsi que la publication de plusieurs boîtes à outils et lignes directrices ». ¹⁹

De plus, plus de 5 000 personnes ont participé au MOOC mondial de l'UNESCO sur l'IA et l'État de droit, proposé afin de renforcer « les capacités des opérateurs judiciaires du monde entier à protéger les droits humains, les principes démocratiques et l'État de droit en relation avec l'utilisation de l'IA. » ²⁰ En outre, une enquête mondiale sur l'utilisation des systèmes d'IA par les opérateurs judiciaires a été menée fin 2023. Les résultats ont permis de mieux comprendre comment les opérateurs judiciaires accèdent aux systèmes d'IA et les utilisent, et ont mis en évidence leurs besoins en matière de formation et d'orientation. ²¹

Dans l'ensemble, cette expérience pratique a révélé les attentes et perceptions des juges concernant la gouvernance des systèmes d'IA dans les salles d'audience, fournissant ainsi une contribution substantielle à la rédaction des Lignes directrices.

Deuxièmement, les documents publiés par l'UNESCO sur l'éthique de l'IA ont également joué un rôle important dans la définition des principes, valeurs et pratiques essentielles inclus dans les Lignes directrices. Il en va ainsi des publications suivantes :

« Recommandation sur l'éthique de l'intelligence artificielle » ²²

« Évaluation de l'impact éthique : un outil de la recommandation sur l'éthique de l'intelligence artificielle » ²³ ; et

« Manuel de formation mondial : l'IA et l'état de droit pour le pouvoir judiciaire » ²⁴, qui fournit un cadre sur les applications et les risques des systèmes d'IA en contexte judiciaire.

Troisièmement, la consultation publique relative au projet de lignes directrices menée en août et septembre 2024 a permis à l'UNESCO de recueillir plus d'une centaine de contributions émises par des particuliers et des organisations de 41 pays, ainsi que des commentaires collectifs issus d'événements organisés avec des acteurs judiciaires dans différentes régions.²⁵

Quatrièmement, vingt évaluateurs externes ont participé à la révision du projet de lignes directrices. Ce panel mondial diversifié comprenait des juges, des juristes, des avocats et des spécialistes de la technologie venus d'Afrique, d'Afrique du Nord et du Moyen-Orient, d'Amérique du Nord et d'Amérique latine, d'Asie, et d'Europe. Leur expertise complémentaire, couvrant la pratique judiciaire, la recherche en technologie juridique, la gouvernance de l'IA et la transformation numérique des systèmes juridiques,

a permis d'évaluer de manière critique les lignes directrices sous plusieurs angles, et de renforcer leur précision et leur caractère pratique.

Cinquièmement, les lignes directrices ont pris en considération les Principes de Bangalore sur la déontologie judiciaire, diverses sources secondaires et la littérature spécialisée, citées dans les notes de fin, y compris les principes et lignes directrices publiés par les autorités judiciaires nationales, les tribunaux et les organisations de la société civile de diverses juridictions, y compris les barreaux et les associations juridiques.

Enfin, les Lignes directrices sont un document vivant qui devra être mis à jour régulièrement afin de tenir compte des progrès rapides de la technologie, en particulier de l'IA, et de ses applications croissantes dans les systèmes judiciaires.



Les Lignes directrices visent à promouvoir le respect universel de la justice et de l'État de droit tout en assurant la protection et la promotion des droits humains.

Lignes directrices pour l'utilisation des systèmes d'IA dans les tribunaux

Les *Lignes directrices* visent à promouvoir le respect universel de la justice et de l'État de droit tout en assurant la protection et la promotion des droits humains.

1. Principes

Les Lignes directrices proposent des principes et des recommandations spécifiques à l'intention des organisations judiciaires et des personnes qui mettent en œuvre des systèmes d'IA, y compris l'IA générative. À cet égard, les « principes » désignent des prescriptions générales destinées à orienter les organisations et les individus dans le développement, l'acquisition et l'utilisation des systèmes d'IA de manière éthique et respectueuse des droits humains. Les Lignes directrices développent 15 principes pour atteindre cet objectif (Figure n° 1).



Figure n° 1. Principes pour le développement, l'acquisition, l'adoption, le déploiement et l'utilisation des systèmes d'IA par le pouvoir judiciaire

Les organisations et les membres individuels du pouvoir judiciaire doivent respecter ces principes tout au long du cycle de vie des systèmes d'IA :



1.1. Protection des droits humains : Prendre les mesures nécessaires pour respecter, protéger et promouvoir les droits humains et l'État de droit tout au long du cycle de vie des systèmes d'IA.²⁶ Il doit exister des garanties efficaces concernant les systèmes d'IA qui pourraient représenter un risque élevé pour les droits humains, telles que la mise en œuvre des évaluations de responsabilité décrites à l'article 2 des Lignes directrices. Le déploiement des systèmes d'IA doit protéger particulièrement les droits des femmes et des filles, des personnes en situation de handicap, des enfants, des réfugiés, des migrants, des apatrides, des minorités et d'autres populations en situation de marginalisation et/ou de vulnérabilité. Bien que tous les droits humains puissent potentiellement être affectés par certains systèmes d'IA tout au long du cycle de vie de ces outils, les quatre suivants sont particulièrement importants dans le cadre des procédures judiciaires :

a. Non-discrimination : Adopter, déployer et utiliser des systèmes d'IA visant à atteindre des objectifs grâce à des processus garantissant l'équité et un accès inclusif à la technologie²⁷. Empêcher le développement et les applications biaisés des systèmes et des résultats d'IA qui pourraient aggraver, reproduire, renforcer ou perpétuer la discrimination fondée sur la race, la couleur, l'ascendance, le genre, l'âge, la langue, la religion, les opinions politiques, la situation nationale, ethnique, sociale ou économique, ou le handicap et tout autre motif.²⁸

b. Égalité : Garantir l'égalité d'accès et de traitement pour tous devant les tribunaux, indépendamment des capacités numériques.²⁹ Par exemple, si les outils d'IA facilitent l'accès à la justice, des obstacles supplémentaires peuvent apparaître pour ceux qui n'ont pas de ressources ou de compétences numériques. Par conséquent, la mise en œuvre par la justice d'outils d'IA devrait prendre en compte les personnes qui ne disposent pas des moyens de se procurer ces technologies, qui ne maîtrisent pas le numérique ou qui sont confrontées à d'autres obstacles technologiques.³⁰ À cette fin, il convient de suivre le principe « Ne laisser personne de côté » (Leave No One Behind, LNOB), élaboré dans le cadre du Programme de développement durable des Nations Unies³¹ pour lutter contre l'exclusion numérique.

c. Équité procédurale : Évaluer les implications des systèmes d'IA sur l'équité procédurale tout au long du cycle de vie du système d'IA et empêcher les déploiements qui enfreignent ce droit.

d. Droit à la vie privée et à la protection des données personnelles : Adopter, déployer et utiliser des systèmes d'IA qui protègent la vie privée et les données à caractère personnel traitées dans le cadre de l'administration de la justice. Le pouvoir judiciaire devrait éviter d'utiliser des outils d'IA qui présentent des risques de divulgation des données confidentielles ou à permettre un accès non autorisé par des tiers.³²

e. Liberté et sécurité : Veiller à ce que les individus ne soient pas détenus sur la base de décisions assistées par l'IA dont ils n'ont pas connaissance, dont ils ne comprennent pas le processus décisionnel ou qu'ils ne peuvent contester.



1.2. Proportionnalité : Adopter, déployer et utiliser des systèmes d'IA visant à atteindre des objectifs légitimes et proportionnels au regard du contexte de leur utilisation.³³



1.3. Faisabilité des bénéfices : Avant d'adopter les systèmes d'IA, la magistrature devrait évaluer à la fois les avantages potentiels spécifiques (pour la justice, les utilisateurs et le public) et sa capacité institutionnelle à réaliser ces avantages.³⁴



1.4. Sécurité : Adopter, déployer et utiliser des systèmes d'IA qui évitent, traitent, préviennent, atténuent ou éliminent les préjudices indésirables pour les parties, les juges individuels, les tribunaux auxquels ils appartiennent, la réputation et l'autorité de la justice, ainsi que pour le public.³⁵



1.5. Sécurité de l'information : Adopter, déployer et utiliser des systèmes d'IA qui protègent les informations confidentielles conformément aux normes internationales d'accès à l'information, mettre en place des mesures de protection contre les cybermenaces et garantir le contrôle continu de la justice sur ses données.³⁶



1.6. Précision et fiabilité : Adopter, déployer et utiliser des systèmes d'IA précis, c'est-à-dire des systèmes d'IA capables de fournir des informations utiles et pertinentes et de produire des résultats et prévisions corrects, ainsi que³⁷ des systèmes d'IA fiables, c'est-à-dire des systèmes d'IA qui fonctionnent « correctement avec une gamme d'entrées et dans diverses situations ».³⁸



1.7. Explicabilité : Adopter, déployer et utiliser des systèmes d'IA capables d'expliquer le raisonnement qui sous-tend leurs résultats et décisions, ainsi que la manière dont ils utilisent les données d'entrée à ces fins. L'explicabilité consiste à rendre intelligibles et à fournir des informations – aux déployeurs et aux utilisateurs – sur les données d'entrée, le raisonnement, les résultats et le fonctionnement des systèmes d'IA, en rendant ces éléments compréhensibles, interprétables et traçables pour les humains.³⁹



1.8. Auditabilité : Prendre des mesures administratives, juridiques et humaines pour garantir et permettre l'audit des systèmes d'IA pendant leur déploiement.⁴⁰



1.9. Justice transparente et ouverte : Adopter, déployer et utiliser des systèmes d'IA transparents en termes de développement, de fonctionnement, de données d'entraînement, de limites (y compris leur marge d'erreur), de capacités et d'objectifs des systèmes.⁴¹ Informer de manière appropriée et en temps utile quand et comment les systèmes d'IA sont acquis, déployés et utilisés, et comment ces outils fonctionnent, en particulier lorsque les décisions prises à l'aide ou sur la base de ces outils peuvent affecter les droits et libertés des individus ou des communautés.⁴² Veiller à informer les individus de leur droit à contester ces décisions.



1.10. Sensibilisation et utilisation éclairée : Comprendre les fonctionnalités, les types d'utilisation, les impacts potentiels, les limites et les risques des systèmes d'IA disponibles afin de prendre des décisions éclairées quant à leur mise en œuvre, connaître l'objectif visé par l'utilisation d'un système d'IA spécifique pour mener à bien des activités judiciaires et se tenir informé des avancées technologiques.⁴³



1.11. Responsabilité : Les organisations et les personnes qui développent et utilisent des systèmes d'IA doivent assumer la responsabilité des décisions et actions prises avec l'aide des outils d'IA, sans préjudice de la responsabilité potentielle du fournisseur en cas de défaillance du système d'IA.⁴⁴



1.12. Responsabilité et contestabilité : Garantir la responsabilité en informant et en expliquant comment et pourquoi certains outils d'IA ont été adoptés par la justice et les résultats obtenus (ou non) lors de leur déploiement. En outre, mettre en place des mécanismes permettant aux parties concernées de contester et de contre-interroger tout résultat produit par un système d'IA susceptible d'influencer l'issue de l'affaire en question, ainsi que de contester les décisions et de tenir les opérateurs judiciaires responsables de leurs erreurs.⁴⁵



1.13. Contrôle et prise de décision humains : Les juges ne doivent déléguer aucune partie de leur mandat ni s'appuyer exclusivement sur les systèmes d'IA pour adopter des décisions ou automatiser des processus entiers susceptibles de nuire aux droits des individus ou des communautés. Ils doivent préserver la possibilité d'une intervention humaine dans tous les processus décisionnels où des systèmes d'IA sont déployés.⁴⁶ Les outils d'IA ne doivent pas remplacer l'analyse indépendante des faits, du droit et des preuves par le décideur.



1.14. Conception centrée sur l'humain et la participation : Le développement, le déploiement et l'utilisation des systèmes d'IA doivent suivre les principes de conception centrée sur l'humain afin de compléter et de renforcer les capacités de la magistrature et de respecter la dignité et l'autonomie humaines. Le processus de conception doit inclure les membres de la magistrature et du public afin de garantir son caractère inclusif.⁴⁷



1.15. Gouvernance et collaboration multipartites : Les organisations judiciaires devraient consulter diverses parties prenantes, organiser des consultations publiques et mettre en œuvre des approches de participation citoyenne tout au long du cycle de vie du système d'IA, en associant en particulier les parties qui peuvent être directement ou indirectement affectées par son déploiement, telles que les femmes, les personnes en situation de handicap, les mineurs et les enfants. De même, il est pertinent de prendre en considération leurs commentaires dans le développement et l'utilisation d'outils d'IA utilisés pour rendre des jugements ou susceptibles d'impacter toute autre question juridique importante.⁴⁸

2. Recommandations spécifiques aux organisations judiciaires

Ces Lignes directrices définissent les « recommandations » comme des voies d'action suggérées qui découlent des « principes » qui aident à la fois les organisations à acquérir et développer des systèmes d'IA sûrs, sécurisés et fiables, et les individus à utiliser les systèmes d'IA de manière éthique, responsable et respectueuse des droits humains.

Les recommandations de cette section sont divisées en trois catégories (Figure n° 2) et s'adressent aux autorités judiciaires, aux tribunaux qui souhaitent évaluer, développer, adopter, et utiliser les systèmes d'IA.

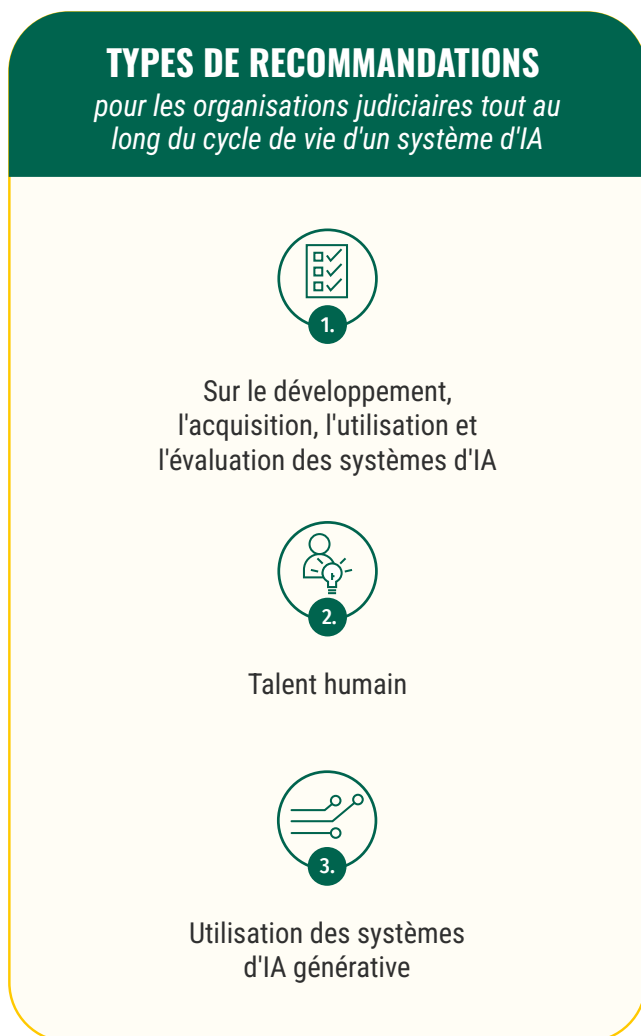


Figure n° 2. Types de recommandations pour les organisations judiciaires tout au long du cycle de vie d'un système d'IA

2.1. Sur le développement, l'acquisition, l'utilisation et l'évaluation des systèmes d'IA

Les recommandations suivantes fournissent des indications sur les actions que les organisations et autorités judiciaires peuvent entreprendre à différentes étapes du cycle de vie de l'IA.

2.1.1. Protéger l'indépendance et l'impartialité du pouvoir judiciaire. Veiller à ce que la justice ait un rôle de supervision et une responsabilité dans le développement, le déploiement et l'utilisation des outils d'IA, et à ce qu'il existe des réglementations protéger son indépendance.⁴⁹ Les outils d'IA ne devraient pas modifier, altérer ou neutraliser les caractéristiques fondamentales de l'impartialité des juges et des tribunaux.⁵⁰

2.1.2. Tenir compte des spécificités culturelles et contextuelles de la population affectée par le déploiement des systèmes d'IA et des obstacles à l'accès à la justice pour les personnes qui ne disposent pas des connaissances, compétences ou ressources nécessaires pour interagir avec le système judiciaire, par exemple les populations âgées, sous-représentées et les populations défavorisées.

2.1.3. Établir une gouvernance multipartite des outils d'IA. Élaborer une politique ou une stratégie en matière d'IA qui identifie les principaux obstacles à l'adoption efficace des systèmes d'IA, établisse des objectifs et des indicateurs, et formule des plans d'action. De plus, envisager la création d'un comité consultatif multipartite chargé de réviser et de mettre à jour en permanence la politique ou de la stratégie en matière d'IA. Les membres individuels de la magistrature devraient être invités à participer à l'amélioration continue de la politique d'IA et de tout autre instrument adopté par l'organisation, par exemple en fournissant des commentaires basés sur leur expérience des systèmes d'IA et des règles et orientations établies pour leur utilisation. En outre, avant d'adopter des systèmes d'IA utilisés pour la prise de décision ou des outils qui soutiennent les processus de gestion et qui peuvent avoir une incidence significative sur les droits fondamentaux, tels que l'accès à la justice, l'organisation devrait consulter les

parties prenantes concernées sur les opportunités, les risques et les impacts de l'introduction de tels outils. Le pouvoir judiciaire devrait impliquer de manière proactive les communautés marginalisées dans les processus de développement, en veillant à ce que leurs points de vue soient pris en compte.

2.1.4. Adopter dès le départ des mesures algorithmiques de transparence et d'explicabilité.

Commencez par déterminer si vous devez opter pour des systèmes open source ou propriétaires. Avant de s'engager à adopter un outil d'IA, l'organisation judiciaire doit évaluer les conséquences de l'adoption de l'une ou l'autre option en termes d'accès à l'information sur le système, d'explicabilité et d'auditabilité.⁵¹ Ensuite, envisagez les mesures de transparence dès les premières étapes du développement et de l'acquisition d'outils d'IA. Par exemple, développer, acquérir et obtenir^{52, 53} des informations auprès des développeurs et fournisseurs du système d'IA sur la manière dont le modèle a été développé (y compris les données utilisées pour l'entraîner) et son fonctionnement, les utilisations appropriées de l'outil, ainsi que les limites et les risques liés à son adoption.^{54, 55}

2.1.5. Nécessité, proportionnalité, adéquation et alignement. La nécessité, la proportionnalité et l'adéquation de l'utilisation d'un système d'IA pour effectuer des tâches doivent être évaluées et établies dès le départ, et alignées sur les objectifs de l'organisation qui utilise le système d'IA (par exemple, protéger les droits humains et assurer l'indépendance et l'autonomie de la justice).

2.1.6. L'équité dès la conception. Lorsque l'organisation judiciaire commande le développement ou l'acquisition d'outils d'IA, des mesures doivent être prises pour s'assurer que les développeurs ou les fournisseurs évitent les biais dans les données et dans la conception des algorithmes. Cela implique que les équipes de développement et de test soient multidisciplinaires et diversifiées.

2.1.7. Mise en œuvre progressive des outils d'IA. Envisager la possibilité de tester de nouveaux outils d'IA avant leur déploiement à grande

échelle et planifier une mise en œuvre progressive. Les tests permettent d'évaluer les fonctionnalités, l'impact et les risques potentiels de l'IA dans un environnement contrôlé. Une stratégie de mise en œuvre progressive permet d'apporter des ajustements en fonction des premiers résultats et des commentaires des parties prenantes, favorisant ainsi un processus de déploiement plus éclairé, responsable et éthique.

2.1.8. Durabilité et impact environnemental des systèmes d'IA. Le développement, le déploiement et l'utilisation de certains systèmes d'IA peuvent avoir une empreinte environnementale importante⁵⁶ que le pouvoir judiciaire devrait prendre en compte.

2.1.9. Établir des règles et des lignes directrices internes pour le développement, le déploiement et l'utilisation d'outils d'IA fondés sur le cadre des droits humains.⁵⁷ Déterminer quels systèmes d'IA peuvent être utilisés, les utilisations acceptées et interdites pour des processus et des tâches spécifiques. De plus, les directives devraient inclure des protocoles de signalement des incidents et des informations claires sur la manière dont chaque acteur est responsable au sein de l'organisation. Les directives devraient insister sur le fait que les membres de la magistrature assument la responsabilité de tout matériel qu'ils produisent grâce à l'IA. À cette fin, les lignes directrices pourraient inclure une évaluation des responsabilités juridiques potentielles et des mécanismes correspondants pour les prévenir. En outre, les lignes directrices devraient proposer des scénarios et des exemples détaillés afin d'orienter les membres du pouvoir judiciaire sur ce qui constitue une utilisation négligente et les mesures à prendre lorsqu'ils ne respectent pas les protocoles d'utilisation de l'IA établis.

2.1.10. Garantir l'intervention humaine. Veiller à ce que le système permette toujours une intervention humaine. En d'autres termes, le contrôle et la supervision humains, ou au moins la surveillance, devraient avoir lieu pendant toutes les étapes de mise en œuvre et d'utilisation du système d'IA.

2.1.11. Mettre en place des cadres de responsabilité avant de déployer des systèmes d'IA.

Avant de déployer un système d'IA pour soutenir les processus décisionnels, définir un cadre de responsabilité⁵⁸ afin de prévenir et d'atténuer les impacts et les risques négatifs et renforcer les effets positifs. Le pouvoir judiciaire peut mettre en œuvre une « évaluation de l'impact algorithmique »⁵⁹ pour identifier les effets probables (positifs et négatifs), une « évaluation des risques algorithmiques »⁶⁰ afin de cartographier les risques potentiels (tels que des défaillances du système), ou une « évaluation de l'impact sur les droits fondamentaux ».⁶¹ Veiller à ce que le développeur ou le fournisseur du système d'IA accepte d'autoriser et de collaborer dans le cadre d'audits algorithmiques réalisés par l'organisation ou par des parties indépendantes et externes.⁶² Si le système d'IA a été développé en interne, faciliter le développement de compétences internes pour surveiller et auditer l'outil (avec des conseils externes si nécessaire), par exemple en établissant « des mécanismes qui facilitent la pertinence du système, tels que la traçabilité et la consignation des processus et résultats du système d'IA ».⁶³ Le pouvoir judiciaire devrait tenir compte du rôle de la société civile et des auditeurs indépendants, en particulier dans les cas où les outils d'IA peuvent présenter des risques élevés d'atteinte aux droits fondamentaux. Ces rapports peuvent déterminer si les systèmes doivent être déployés et recommander des mesures de prévention, d'atténuation, de recours et de surveillance. Les évaluations algorithmiques devraient être accessibles au public. L'« Évaluation de l'impact éthique : un outil de la recommandation sur l'éthique de l'intelligence artificielle » publié par l'UNESCO peut être utilisé pour examiner les impacts potentiels sur les droits humains, tels que le droit à un procès équitable et à une procédure régulière, l'accès à la justice et à un recours effectif, et le droit à la vie privée⁶⁴. Ces évaluations devraient être réalisées en consultation avec les parties prenantes concernées.

2.1.12. Réaliser des audits algorithmiques et des évaluations d'impact après le déploiement des systèmes d'IA.⁶⁵ À cette fin, le pouvoir judiciaire devrait envisager la création de lignes

directrices pour le suivi, l'audit et l'évaluation des systèmes d'IA dans le secteur judiciaire.⁶⁶ Les évaluations devraient porter sur la conformité de ces systèmes avec la réglementation ou sur des questions plus ciblées telles que les biais, l'équité, la transparence, l'explicabilité, la sécurité ou les performances.⁶⁷ Elles peuvent également évaluer les effets des systèmes d'IA sur les utilisateurs, les populations concernées et la société. Ces dernières devraient inclure l'évaluation de l'impact sur les droits humains et l'identification des principaux effets (intentionnels et non intentionnels) du système sur divers groupes d'utilisateurs et populations. Les évaluations complètes de l'impact sur les droits humains devraient être publiques.⁶⁸ Publier des rapports périodiques sur les impacts et les performances (par exemple, efficacité et efficience) des systèmes d'IA adoptés pour atteindre les objectifs de l'organisation. De plus, il convient d'établir des processus permettant aux tiers (par exemple, les fournisseurs, les consommateurs, les distributeurs/fournisseurs) ou aux travailleurs de signaler les vulnérabilités et les risques potentiels.⁶⁹ En outre, les autorités judiciaires devraient mettre en place un système de gestion des risques qui lui permette d'identifier, de détecter, de surveiller, de classer, de diagnostiquer, de contrôler, de prévenir les risques et atténuer les dommages. Le système de gestion devrait attribuer clairement des rôles, responsabilités et procédures tout au long du cycle de vie de l'IA.⁷⁰

2.1.13. Cesser l'utilisation des systèmes d'IA qui affectent négativement les droits humains.

Mettre en œuvre des systèmes développés et proposés conformément au droit international des droits humains, à la constitution nationale et aux lois locales, telles que les lois relatives à la protection des données et la vie privée, la non-discrimination, l'équité procédurale, les droits de la consommation, de la concurrence et de la propriété intellectuelle. De plus, lorsque le système d'IA provient d'un fournisseur externe, il convient de s'assurer que les conditions d'utilisation de l'outil sont compatibles avec les lois locales et les normes organisationnelles des autorités judiciaires. La magistrature devrait s'abstenir d'adopter des

systèmes d'IA susceptibles d'avoir un impact négatif sur les droits humains et/ou civils, par exemple en perpétuant des résultats discriminatoires et d'autres injustices.⁷¹ Dans le même ordre d'idées, comme le recommande le rapport du Secrétaire général de l'ONU « Les droits humains dans l'administration de la justice », le pouvoir judiciaire devrait s'abstenir ou cesser d'utiliser des outils d'IA « qui ne peuvent fonctionner dans le respect du droit international des droits humains », ce qui inclut des applications prédictives utilisés « dans la prise de décisions judiciaires dans les affaires pénales et dans les évaluations du risque de récidive pour les décisions de mise en liberté sous caution et de libération conditionnelle, à moins et jusqu'à ce que les autorités compétentes puissent démontrer que ces applications sont conformes au droit à un procès équitable, y compris à l'indépendance du pouvoir judiciaire, et au droit à la liberté et à la sécurité, au droit à la non-discrimination, au droit de ne pas être soumis à la torture et à des mauvais traitements et au droit au respect de la vie privée, ainsi qu'aux autres droits humains concernés, et que les autorités se sont attaquées aux incidences disproportionnées que l'utilisation de ces technologies peut avoir sur des groupes particuliers.⁷² » Lorsque la justice dispose de preuves préliminaires que l'utilisation d'un outil d'IA peut produire des effets négatifs importants ou de donner lieu à des violations des droits humains, le déploiement et l'utilisation du système doivent être suspendus jusqu'à ce qu'une enquête approfondie soit menée. Lorsque des évaluations algorithmiques ou d'impact concluent que des systèmes d'IA violent les droits humains, le pouvoir judiciaire cesse leur utilisation et met l'outil hors service.

2.1.14. Accessibilité des outils d'IA. Veiller à ce que les membres de la magistrature qui pourraient bénéficier des outils d'IA aient un accès adéquat aux systèmes. Cela implique notamment de s'assurer que les individus peuvent disposer du matériel nécessaire pour faire fonctionner les outils et des compétences nécessaires pour les utiliser efficacement. À cette fin, il convient de diagnostiquer, de surveiller et de lever périodiquement les obstacles à l'adoption des systèmes d'IA, en particulier

ceux auxquels sont confrontés les personnes appartenant au pouvoir judiciaire.

2.1.15. Mesures visant à renforcer la sécurité et la cybersécurité. Les systèmes d'IA sont vulnérables aux attaques malveillantes, qui peuvent modifier leurs résultats. Par exemple, les systèmes d'IA générative peuvent être vulnérables aux injections de prompt, à l'extraction de données, à l'évasion et aux attaques par empoisonnement.⁷³ Par conséquent, avant de déployer un outil, le pouvoir judiciaire doit tester sa résilience, le sécuriser et le maintenir tout au long de son cycle de vie. Adoptez des mesures techniques, managériales et humaines pour prévenir, contrôler et atténuer les risques et incidents en cybersécurité. Lorsque le déploiement et l'utilisation du système d'IA impliquent l'accès aux services cloud, assurez-vous que le niveau de protection offert par le fournisseur correspond aux normes juridiques locales et aux normes de sécurité organisationnelles.

2.1.16. Renforcer la protection des données personnelles et de la vie privée. Établir un cadre et une infrastructure robustes de gouvernance des données pour protéger les données personnelles, promouvoir des pratiques responsables de partage des données personnelles et non personnelles et veiller à ce que ces données soient inclusives. L'organisation devrait également établir des « protocoles de données régissant l'accès aux données. Ces protocoles devraient préciser qui peut accéder aux données et dans quelles circonstances. Seul le personnel dûment qualifié, compétent et ayant besoin d'accéder aux données individuelles devrait être autorisé à le faire ». ⁷⁴ Compte tenu de la nature sensible des données personnelles et juridiques traitées dans le cadre des procédures judiciaires, il est essentiel de mettre en place des mesures strictes de protection de la vie privée. (i) Minimisation des données. Pour atténuer le risque de violations de données, il convient d'utiliser des systèmes d'IA nécessitant un minimum de données à caractère personnel pour fonctionner, en particulier dans les cas impliquant des informations personnelles sensibles. Par conséquent, les organisations judiciaires devraient déployer des systèmes d'IA basés sur

des processus renforçant la protection de la vie privée. (ii) Protocoles de consentement. Conformément aux lois locales sur la protection des données et la vie privée, rédiger et mettre en œuvre des protocoles de consentement pour le traitement des données personnelles à l'aide d'outils d'IA au sein du système judiciaire. Cela permettra de garantir que les parties sont informées et peuvent contrôler la manière dont leurs données sont utilisées. (iii) Des techniques d'anonymisation des données devraient être employées lorsque les systèmes d'IA traitent des données personnelles, notamment lors de la création de bases de données utilisées pour l'analyse juridique ou l'étude de la jurisprudence (par exemple, anonymisation des données personnelles des parties au litige mentionnées dans les décisions de justice avant la publication des décisions).⁷⁵ Toutefois, cette approche doit être mise en balance avec la nécessité de disposer de jeux de données ouverts ainsi qu'avec le droit à la liberté d'expression et à l'accès à l'information, afin de ne pas entraver le développement des systèmes de justice numériques. Le point d'équilibre entre ces droits dépend de la manière dont chaque juridiction régit le droit d'accès aux documents judiciaires et des règles d'anonymisation issues des lois sur la protection des données et la vie privée.

2.1.17. Divulguer de manière proactive les informations clés sur les systèmes d'IA utilisés par la justice. Par exemple, les autorités judiciaires peuvent publier en ligne un répertoire contenant des informations clés sur les systèmes d'IA adoptés pour administrer la justice.⁷⁶ Ce répertoire peut contenir des informations sur les données d'entraînement, le modèle (par exemple variables, pondérations), s'il s'agit d'un modèle propriétaire ou open source, le fonctionnement des systèmes d'IA, les utilisateurs de ces outils, leur utilisation, les implications pour les processus décisionnels judiciaires, et la possibilité de contester les décisions prises avec le soutien du système d'IA.⁷⁷ Ces informations doivent être mises à jour périodiquement, au moins une fois par an, et indiquer la date de la dernière mise à jour. En outre, la création de référentiels publics des systèmes d'IA utilisés par les tribunaux

pourrait contribuer à l'ouverture des données. À cette fin, il convient d'évaluer dans quelle mesure les instruments de transparence peuvent satisfaire aux principes dits « FAIR » en matière de données ouvertes : faciles à trouver, accessibles, interopérables et réutilisables par tout le monde et à toutes fins. Les lignes directrices de l'UNESCO pour l'établissement et la mise en œuvre de politiques de données ouvertes peuvent être prises en compte à cette fin.⁷⁸ Parallèlement à cet effort, le pouvoir judiciaire pourrait également mener des campagnes de sensibilisation du public sur l'utilisation des systèmes d'IA dans les processus judiciaires.

2.1.18. Garantir la contestabilité des décisions produites ou influencées par les systèmes d'IA.

Mettre en œuvre les mesures nécessaires pour garantir le droit de s'opposer aux décisions produites ou influencées par les outils d'IA. Cela nécessite une compréhension approfondie de la manière dont les résultats ont été produits et le droit pour les parties de contrôler et d'examiner le modèle d'IA sous-jacent afin d'identifier les lacunes susceptibles d'affecter leur dossier.

2.2. Talent humain

Les recommandations suivantes fournissent des indications sur les actions à entreprendre et les mesures à adopter par le pouvoir judiciaire pour soutenir les individus qui déploieront et utiliseront des systèmes d'IA.

2.2.1. Réviser le programme des écoles judiciaires.

Envisager d'intégrer des sujets relatifs à l'IA dans le programme permanent des écoles judiciaires, tels que les principes fondamentaux de l'IA, l'application des technologies d'IA dans les systèmes judiciaires, et l'impact de l'IA sur l'indépendance des juges, l'administration de la justice, notamment en ce qui concerne les droits humains, l'éthique et la gouvernance.⁷⁹

2.2.2. Donner accès à la formation. Offrir régulièrement des formations aux membres de la magistrature qui utiliseront les outils d'IA et veiller à développer les compétences fondamentales en IA, notamment pour identifier l'adéquation entre les outils d'IA et les tâches

à effectuer, comprendre comment utiliser ces outils et leurs impacts et risques potentiels conformément au droit international, régional et national des droits humains, et évaluer leurs résultats de manière critique.⁸⁰ Les organisations judiciaires devraient examiner s'il est possible et pertinent de mettre en place des programmes de formation obligatoires pour le personnel judiciaire afin de maintenir à jour ses connaissances et compétences.

2.2.3. Recherche sur l'utilisation et les effets des outils d'IA. Les autorités judiciaires devraient étudier périodiquement la façon dont leurs membres utilisent les systèmes d'IA. La recherche pourrait documenter des études de cas afin de mettre en lumière les bonnes pratiques et les écueils à éviter. Ces dernières aideraient les magistrats à partager leurs expériences et les défis avec leurs pairs. Un tel exercice devrait être réalisé dans des conditions propres à garantir le plein respect de l'indépendance du pouvoir judiciaire (voir supra, 2.1.1).

2.2.4. Évaluer la nécessité de mettre à jour les politiques et pratiques en ressources humaines. En fonction des outils d'IA utilisés, le pouvoir judiciaire pourrait être amené à revoir la gestion de ses effectifs, soit en renforçant les compétences du personnel en place et/ou en recrutant de nouveaux collaborateurs. Cela peut nécessiter des mises à jour des politiques et des processus de recrutement pour l'évaluation des performances afin de garantir que l'utilisation appropriée des outils d'IA est évaluée et encouragée.

2.3. Utilisation des systèmes d'IA générative

Les recommandations suivantes fournissent des orientations sur les mesures que les autorités judiciaires pourraient prendre en ce qui concerne l'adoption, le déploiement et l'utilisation des systèmes d'IA générative.

2.3.1. Sensibilisation aux fonctions et aux limites. Veiller à ce que les membres de la magistrature qui utiliseront les outils d'IA générative soient conscients des usages adéquats de l'outil, des limites et des risques associés (par exemple, résultats biaisés ou incorrects) à leur adoption pour la rédaction de documents juridiques et l'aide à la partici-

pation aux procédures judiciaires, ainsi que pour les activités de gestion des tribunaux (par exemple, la traduction).

2.3.2. Garantir l'authenticité et l'intégrité du contenu. Pour y parvenir, plusieurs mesures doivent être prises. (i) Tous les documents juridiques, les présentations de preuves ou les avis judiciaires générés par l'IA doivent être clairement étiquetés comme ayant été (co-)produits par l'IA afin que les juges, avocats et parties concernées soient conscients de la nature du contenu. (ii) Des systèmes robustes doivent être mis en place pour suivre l'évolution et les modifications du contenu juridique généré par l'IA. Cela est important à des fins de preuve, car cela garantit que l'authenticité de tous les documents utilisés au tribunal peut être vérifiée. Par exemple, si un outil d'IA a été utilisé pour créer ou modifier un document juridique, il devrait y avoir un enregistrement clair du moment et de la manière dont il a été utilisé. (iii) Des protocoles de certification devraient être élaborés pour les outils d'IA utilisés dans le système judiciaire, afin de vérifier qu'ils respectent les recommandations éthiques internationales et locales ainsi que les normes juridiques de chaque juridiction en matière de précision et de fiabilité. Un exemple pourrait être un processus de certification pour un outil d'IA utilisé pour analyser des preuves, garantissant qu'il est suffisamment précis et fiable pour être utilisé devant un tribunal.

2.3.3. Restrictions d'utilisation. Des directives spécifiques devraient être publiées pour régir l'utilisation de l'IA générative dans le système judiciaire afin de prévenir les abus et de protéger l'intégrité du processus judiciaire, conformément aux normes internationales. Certaines applications spécifiques de l'IA devraient être interdites ou limitées, compte tenu de leur impact sur les droits humains, en particulier la vie privée et la protection des données. Par exemple, lorsque les conditions d'utilisation d'un outil d'IA générative indiquent que le fournisseur utilisera les invites (requêtes ou *prompts*) de l'utilisateur pour entraîner ses modèles ou que des tiers peuvent y accéder, alors l'utilisation de cet outil devrait être interdite ou restreinte afin d'éviter que la justice ne perde le contrôle de l'accès aux informations

confidentielles ou aux données à des données personnel. De plus, l'utilisation de l'IA dans certains domaines sensibles, comme la génération unilatérale de décisions contraignantes ou la création de preuves fabriquées, devrait être interdite. Une attention particulière doit être accordée au respect du droit de la propriété intellectuelle en veillant à ce que tout contenu généré par l'IA respecte la propriété intellectuelle du texte original.

3. Recommandations spécifiques aux membres individuels du pouvoir judiciaire

Les recommandations de cette section s'adressent aux personnes faisant partie du pouvoir judiciaire, y compris les magistrats, procureurs, officiers et assistants de justice qui souhaitent déployer et utiliser des systèmes d'IA. Les recommandations sont divisées en trois catégories (Figure n° 3).



Figure n° 3. Types de recommandations pour le déploiement et l'utilisation des systèmes d'IA par les membres individuels du pouvoir judiciaire

3.1. Avant d'utiliser les systèmes d'IA

Les recommandations suivantes fournissent des suggestions sur les actions que les membres individuels du pouvoir judiciaire peuvent prendre avant de déployer et d'utiliser des systèmes d'IA.

3.1.1. Sensibilisation à l'IA, connaissance en la matière et renforcement des compétences.

Les personnes doivent être conscientes des fonctionnalités, des atouts et de la précision des systèmes d'IA ainsi que de leurs limites, biais et risques dans le cadre des activités juridiques.⁸¹ Ce dernier point inclut également la prise de conscience des responsabilités découlant de l'utilisation négligente de ces outils. Les organisations judiciaires doivent veiller à ce que les personnes aient accès à des programmes d'éducation et à une formation continue afin d'évaluer de manière critique les résultats et les performances des outils.⁸²

3.1.2. Utiliser des outils d'IA ayant fait l'objet d'une analyse d'impact et dont l'usage est conforme aux recommandations issues des rapports d'évaluation en matière de droits humains.

Comme expliqué précédemment, les organisations judiciaires devraient mettre en place des cadres de responsabilité avant et après le déploiement des systèmes d'IA afin d'identifier les implications et les risques pour l'accès à la justice et aux droits humains. Cependant, si votre organisation n'évalue pas ses systèmes d'IA, veillez à utiliser des outils testés (par le prestataire ou des tiers experts) et sélectionnez ceux qui ont fait l'objet d'évaluations d'impact sur les droits humains.⁸³ De plus, suivez les recommandations des rapports de diligence raisonnable (due diligence) en matière de droits humains concernant les systèmes d'IA, qui doivent être réalisés tout au long du cycle de vie des systèmes d'IA par les autorités judiciaires.

3.2. Au cours de l'utilisation des systèmes d'IA

Les recommandations suivantes fournissent des indications sur les mesures que les membres du pouvoir judiciaire peuvent prendre lors du déploiement et de l'utilisation des systèmes d'IA.

- 3.2.1. Évitez d'accorder une confiance excessive aux systèmes d'IA pour prendre des décisions importantes.** Les juges ne devraient pas se fier entièrement aux systèmes d'IA pour statuer sur le fond d'une affaire ni pour trancher sur des questions de procédure pouvant affecter les droits humains ; ils doivent plutôt utiliser les résultats des outils d'IA pour compléter l'analyse juridique effectuée à l'aide d'autres méthodes et sources d'information. La décision judiciaire n'est pas un simple résultat, mais une décision prise à l'issue d'un processus.
- 3.2.2. Respectez les conditions d'utilisation.** Lisez et respectez les conditions d'utilisation des systèmes d'IA (publiées par le fournisseur externe de l'outil ou fournies par votre organisation lorsque le système a été développé en interne). Ces conditions indiquent généralement les utilisations appropriées et interdites des outils, ainsi que les risques que les déployeurs et les utilisateurs doivent éviter. Cependant, consultez votre organisation au sujet des conditions d'utilisation des systèmes d'IA, en particulier si vous estimez que certaines conditions peuvent restreindre indûment les droits des utilisateurs, nuire à la transparence ou dégager le développeur de toute responsabilité.
- 3.2.3. Transparence proactive.** Fournir des informations pertinentes sur le moment où l'outil d'IA est utilisé et sur la manière dont son utilisation pourrait affecter les personnes impliquées dans une procédure judiciaire. Indiquer quels outils ont été utilisés et leurs versions. Veiller à ce que les documents produits ou publiés à l'aide de l'IA sont reconnus et distingués à l'aide de guillemets ou de citations.
- 3.2.4. Responsabilité et obligation de rendre des comptes.** Respectez les normes établies par votre organisation. Cette recommandation implique à la fois une obligation *ex-ante* de divulguer l'utilisation d'un système d'IA et une obligation *ex-post* de fournir des informations supplémentaires lorsque cela est nécessaire pour déterminer la responsabilité.
- 3.2.5. Possibilité de revoir les décisions et la contestabilité.** Informer les parties de l'utilisation des systèmes d'IA et offrir des opportunités effectives de contester les décisions prises avec ou à l'aide de systèmes d'IA. Fournir des informations de base sur le fonctionnement du système d'IA, son développement, les données utilisées pour le faire fonctionner et la manière dont les résultats produits par l'outil d'IA ont influencé la décision. Veiller à ce que chacun ait droit à un recours effectif dans les cas où les systèmes d'IA ont un impact négatif sur leurs droits humains, y compris pendant et après la fin des procédures pénales à leur encontre. Le pouvoir judiciaire doit veiller à ce que toute personne souhaitant contester ces décisions ait accès à des services d'aide juridique, à des sources d'information ou à une assistance pour s'orienter dans le processus d'appel.
- 3.2.6. Rapports proactifs pour prévenir les dommages.** Informez l'organisation en cas de suspicion de dysfonctionnement ou d'impacts négatifs potentiels ou probables, et cessez d'utiliser le système d'IA si vous constatez qu'il peut nuire aux droits humains.
- 3.3. Au cours de l'utilisation des systèmes d'IA générative**
- Les recommandations suivantes fournissent des orientations sur les actions à entreprendre par les membres individuels du pouvoir judiciaire en ce qui concerne l'adoption, le déploiement et l'utilisation des systèmes d'IA générative.
- 3.3.1. Protégez les données personnelles et confidentielles.** N'incluez pas de données personnelles ou d'informations confidentielles dans les requêtes (prompts) lorsque vous utilisez des outils d'IA générative externes.⁸⁴ Sachez que toute information saisie dans un chatbot public basé sur l'IA (par exemple, dans le cadre d'une invite ou prompt) doit être considérée comme accessible à tous, notamment parce que les conditions d'utilisation de la plupart des entreprises d'IA qui offrent un accès gratuit à ces systèmes indiquent que les données saisies seront utilisées pour entraîner de futurs modèles.⁸⁵

3.3.2. Principales utilisations des LLM. Les LLM peuvent être utilisés pour différentes tâches, notamment la rédaction de documents juridiques de base, la rédaction de discours et de présentations, la synthèse, la traduction, la correction grammaticale, la modification du ton d'un texte (par exemple, informel à formel), l'amélioration de sa lisibilité, l'exploration de sujets spécifiques et l'exécution de tâches administratives (par exemple, la rédaction d'e-mails).⁸⁶ Cependant, toutes ces tâches impliquent de vérifier les résultats du système d'IA et la vérification avec des sources fiables. Les outils d'IA peuvent vous aider à trouver des documents que vous connaissez et que vous pouvez évaluer vous-même, mais qui sont actuellement indisponibles. En revanche, ces outils sont inefficaces pour mener des recherches sur des informations qui ne peuvent être évaluées ou vérifiées de manière indépendante.⁸⁷ Voir les utilisations suggérées des outils d'IA générative à la fin de cette section.

3.3.3. Manque de fiabilité en tant que moteurs de recherche et pour l'analyse juridique. Les LLM commerciaux à usage général ne sont pas des sources d'information fiables ni des moyens adéquats pour effectuer des analyses juridiques ou des tâches mathématiques. « Même avec les meilleurs prompts, le résultat peut être inexact, incomplet, trompeur ou biaisé. »⁸⁸

3.3.4. Ne remplacez pas les témoignages d'experts par des LLM. Les témoignages d'experts englobent les processus et les méthodologies utilisés pour produire les rapports d'experts, ainsi que les descriptions, les raisons, les arguments, les analyses, les évaluations et les conclusions présentés dans le texte. Les LLM ne doivent pas être utilisés pour produire le type de résultat attendus des experts, non seulement en raison du manque de fiabilité des résultats, mais également de l'opacité avec laquelle le système tire ses conclusions à partir de ses données d'entraînement.⁸⁹

3.3.5. Conscience des limites et des risques des LLM. Sachez que les résultats générés par les LLM actuels peuvent contenir des informations incorrectes, imprécises ou fictives

concernant des questions factuelles, juridiques (lois et jurisprudence) et techniques. Les LLM peuvent produire des réponses sans aucun rapport avec le monde réel et des textes absurdes.⁹⁰ De plus, vous devez être conscient des biais potentiels dans les textes produits par les LLM : « Les déséquilibres dans les données d'entraînement peuvent également entraîner des biais dans le modèle. »⁹¹ De plus, certains LLM peuvent avoir été entraînés sur des données jusqu'à un certain moment (par exemple, les données disponibles jusqu'au moment de leur dernier entraînement). Par conséquent, les informations relatives aux affaires, aux lois et aux faits récents n'ont peut-être pas été incluses dans les données d'entraînement. Ainsi, les résultats produits par le chatbot alimenté par LLM peuvent être obsolètes ou inexacts.

3.3.6. Vérifiez les résultats avant de les utiliser. La structure convaincante du texte produit par un LLM ne doit pas conduire à une confiance excessive dans la factualité et la véracité du résultat.⁹² « Pour diverses raisons, les LLM n'offrent aucune garantie quant à la factualité, la qualité et le format souhaité (par exemple, le format de code spécifique) de leurs résultats. »⁹³ Par conséquent, évitez une confiance excessive dans les outils d'IA en veillant toujours à ce que les résultats des systèmes d'IA générative soient vérifiés et comparés à des sources d'information fiables avant leur utilisation dans les documents et supports judiciaires. La pensée critique humaine doit être au cœur de toutes les interactions avec les systèmes d'IA générative.

3.3.7. Utilisation transparente. Déclarez l'utilisation de systèmes d'IA générative pour la rédaction de textes – décisions, avis et autres documents pouvant avoir des conséquences juridiques – ou lorsqu'ils sont explicitement utilisés lors des audiences. À cette fin, distinguez le texte produit par le chatbot IA utilisé dans une décision en utilisant des guillemets et un système de citation.⁹⁴ Lorsque le texte produit par le chatbot est inclus dans une décision de justice, l'ensemble de l'interaction ayant conduit au résultat final – à la fois les requêtes (prompts) de l'utilisateur et les réponses du chatbot – doit être documenté. En

fonction de la longueur et de la complexité de l'interaction, cette documentation peut être incluse dans le corps principal du texte, dans une note de bas de page ou dans une annexe.

3.3.8. Intégrité. Ne présentez pas un texte synthétique comme étant le vôtre. De plus, évitez toute violation potentielle des droits d'auteur et de propriété intellectuelle associée à l'utilisation de contenus produits par des systèmes d'IA générative.

3.3.9. Responsabilité. Assumez la responsabilité de l'utilisation des résultats produits par les systèmes d'IA générative pour rédiger les jugements et décisions judiciaires et pour vous éclairer lors des audiences. Assurez-vous de l'exactitude des résultats avant de les utiliser.⁹⁵

ENFIN, VOICI DIX EXEMPLES D'UTILISATIONS DES SYSTÈMES D'IA GÉNÉRATIVE :

1. Corriger ou améliorer les premières ébauches de rapports administratifs, procès-verbaux de réunions, d'e-mails, de notes, d'allocation, de décisions et jugements. Exemples d'utilisation :
 - Améliorer la grammaire, l'orthographe et la structure des textes.
 - Modifier le ton du texte, par exemple, en le rendant plus accessible à différents publics cibles ou en utilisant un langage simple.
 - Préparer des projets de traduction.
2. Réorganiser les références selon un système de citation de sources donné.
3. Comparer des textes et des sources.
4. Concevoir des présentations et créer des visuels.
5. Produire et modifier du code informatique.
6. Effectuer des explorations sur des sujets déterminés et poursuivre les étapes de recherche ultérieures à partir de sources fiables.
7. Réfléchir, concevoir, et tester des idées.
8. Structurer des arguments.
9. Explorer des explications alternatives ou critiques sur des points de droit.
10. Produire des résumés : toujours vérifier l'exactitude des résultats.



*Les outils d'IA ne peuvent se substituer
à un raisonnement juridique qualifié,
au jugement humain ou aux conseils
juridiques personnalisés.*

Conclusion

Les outils d'IA peuvent soutenir les objectifs fondamentaux du secteur judiciaire. Cependant, l'adoption d'instruments défectueux et l'utilisation négligente des systèmes d'IA par la justice peuvent également porter atteinte aux droits humains, tels que le droit à un procès équitable, l'accès à la justice et à un recours effectif, la protection de la vie privée et des données, l'égalité devant la loi et la non-discrimination, ainsi qu'aux valeurs judiciaires telles que l'impartialité, l'indépendance et la responsabilité.

Les *Lignes directrices* proposent des principes et des recommandations visant à maximiser les avantages des outils d'IA dans le système judiciaire et à prévenir les préjudices potentiels, en particulier les violations des droits humains. Si le pouvoir judiciaire peut appliquer directement les principes et recommandations pour les organisations et les individus, ceux-ci peuvent également servir **de modèle pour l'élaboration de lignes directrices nationales et infranationales adaptées au contexte.**

Toutefois, la mise en œuvre efficace des outils d'IA dans la justice nécessite davantage que l'adoption de règles et de lignes directrices. Il est également essentiel de veiller à ce que les professionnels soient préparés à utiliser les outils d'IA de manière éthique et responsable, et que ces outils soient accessibles, fiables et conformes aux droits humains.

Les outils d'IA ne peuvent se substituer à un raisonnement juridique qualifié, au jugement humain ou à des conseils juridiques personnalisés. Par exemple, les chatbots IA alimentés par des LLM génèrent du texte en assemblant des séquences de formes linguistiques détectées dans les données d'entraînement « selon un calcul probabiliste sur la manière dont elles s'associent, mais sans aucune référence au sens ».⁹⁶ Ces outils d'IA ne peuvent remplacer un raisonnement juridique qualifié car ils n'ont aucune compréhension rationnelle ou contextuelle d'un problème juridique. L'utilisation des LLM à des fins juridiques par des non-juristes nécessite la plus grande prudence, car les résultats fournis par ces systèmes peuvent sembler fiables et cohérents alors qu'ils produisent des faits inexacts et des réponses non conformes au droit.⁹⁷

En résumé, la justice devrait avoir pour objectif d'utiliser des outils d'IA pour améliorer, plutôt que remplacer, le jugement humain.

Annexe

Évaluateurs externes (par ordre alphabétique) :

Karine Caunes (Global Program Director, Droit de l'UE, Center for AI and Digital Policy, États-Unis), Giulia Gentile (Assistant Professor, Essex Law School, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord), Felipe Guzmán (Consultant sur les instruments réglementaires pour la transformation numérique de la justice, Conseil supérieur de la magistrature de Colombie), Ameen Jauhar (Chef d'équipe, Centre for Applied Law & Tech, Vidhi Centre for Legal Policy), Amanda Leal (spécialiste de la gouvernance et des politiques de l'IA chez HealthAI), Martin Okumu Masiga (Secrétaire général, Africa Jurists and Judges Forum, Johannesburg, Afrique du Sud), Grace Mutung'u (Chercheuse et Avocate à la Haute Cour du Kenya), Jean Aloïse Ndiaye (Juge à la Cour suprême du Sénégal), Alfonso Peralta Gutiérrez (Juge de première instance et d'enquête pénale, Grenade, Espagne), Matthieu Quinou (Docteur en droit privé et avocat au barreau de Paris, France), Anca Radu (Chercheuse, Institut universitaire européen, Département de droit, Florence, Italie), Dr. Krishna Ravi Srinivas (Professeur adjoint à l'Université de droit NALSAR, Inde), Dr. Catherine Regis (Professeure, Faculté de droit, Université de Montréal, Canada), Dr. Dory Reiling (Juge senior à la retraite, Expert en Cour internationale et TIC, Amsterdam, Royaume des Pays-Bas), Agneris Sampieri (Avocate mexicaine), Dr. Miriam Stankovich (Spécialiste en politique numérique, Center for Digital Acceleration, États-Unis), Kalliopi Terzidou (doctorante, Faculté de droit, Université du Luxembourg), Colin van Noordt (Chercheur, Royaume des Pays-Bas), Carolina Villadiego (Conseillère juridique pour l'Amérique latine, Commission internationale des juristes, Genève, Suisse), Anthony Wong (Président de l'IFIP, Fédération internationale pour le traitement de l'information, Autriche).

Au-delà des consultations d'experts, des consultations publiques sur les projets de lignes directrices ont été organisées entre août et septembre 2024. Les lignes directrices finales ont été enrichies par plus d'une centaine de contributions envoyées par des particuliers et des organisations dans 41 pays, ainsi que par des retours collectifs issus d'événements organisés par l'UNESCO avec des acteurs judiciaires dans différentes régions du monde.

ONU HCDH, Organisation Mondiale de la Santé, Access Now (États-Unis), Access to Justice Centre for Excellence, ACT (Canada) : The App Association (États-Unis), Advocacy and Awareness Centre Africa (AAC Africa) Initiative (Kenya), African IGF & African Digital Rights Network (Éthiopie, contributeur depuis l'Égypte), Anticorruption Research Centre (Kenya), ARTICLE 19 (Brésil, et Office for Mexico and Central America), Auticon (France), Cambridge Existential Risk Initiative (Royaume-Uni de Grande-Bretagne et Irlande du Nord), Center for AI and Digital

Policy (États-Unis), Centre for Health Equity, Law & Policy (C-HELP, Inde), Centre for Internet and Society (CEI, Inde), Centre for Space, Cyberspace & Data Law (Bond University, Australie), Centre for the Study of Law and Governance (Jawaharlal Nehru University, Inde), Centro de Capacitación y Perfeccionamiento 'Ricardo C. Núñez' del Poder Judicial de la Provincia de Córdoba (Argentine), Centro Federal de Conciliación y Registro Laboral (Mexique), Centro para la Investigación del Cibercrimen y la Ciberseguridad (Centre du CIC, Colombie), Centre for Intellectual Property and Information Technology Law (CIPIT, Kenya), Consejo de la Judicatura Federal (Mexique), Council for Research on International Economic Relations (ICRIER, Inde), Cour d'appel du Caire (Égypte), Cuevas Mella (République dominicaine), Société Daksh (Inde), Rights Nepal, El Veinte (Colombie), Laboratoire Empatia (Pérou), Escuela Judicial Electoral (École judiciaire électorale, Mexique), Institut européen d'administration publique (Luxembourg) en collaboration avec le Maastricht Law & Tech Lab de l'Université de Maastricht (Royaume des Pays-Bas), et la Faculdade de Tecnologia de Teresina (Brésil), Faculté de droit, Université de technologie de Sydney (Australie), Université fédérale balte Immanuel Kant (Fédération de Russie), Institut hellénique de droit international et étranger (Grèce), Ikigai Law (Inde), Institut indien de commerce extérieur de Kolkata (Inde), Inno Canyon Consulting (Nigeria), Association Internationale des Procureurs (Autriche), Kictanet (Kenya), Laboratorio de Innovación Tecnológica de la Facultad de Derecho de la Universidad Nacional de Lomas de Zamora UNLZ-LAB (Argentine), Avocats sans frontières (Canada, El Salvador), Learned Hand (États-Unis), Commission nationale lituanienne pour l'UNESCO (Lituanie), Maseru Magistrate Court

(Lesotho), Ministerio Publico (Panama), Ministère des Affaires étrangères d'Espagne, Ministère de la Justice, des Affaires parlementaires (Népal), Académie nationale d'études juridiques et de recherche (NALSAR, Inde), Académie judiciaire nationale (Népal), Association du barreau du Népal, New Zealand Law Society, Osh Technology & Artificial Intelligence Industries (Égypte), Pan African Lawyers Union (Tanzanie), Poder Judicial Asunción (Paraguay), Priva Sapien (Inde), Rama Judicial (Colombie), Rencontre Africaine pour la Défense des Droits de l'Homme (RADDHO, Sénégal), cabinet S.B Law (Islamabad, Pakistan), Saffi technologies/Kenya internet action network (Kenya), Faculté de droit, Trinity College Dublin (Irlande), Tribunal régional de Sofia (Bulgarie), Cour suprême du Paraguay, Tribunal suprême de justice de l'État d'Aguascalientes (Mexique), Université Taiz (Yémen), TeachSomebody (Royaume des Pays-Bas), L'équipe de droit civil de l'Association des juges polonais IUSTITIA (Pologne), Cour suprême du Népal, École d'économie de Varsovie SGH (Pologne), UBA – IALAB (Argentine), Universidad Americana (Paraguay), Universidad Galileo (Guatemala), Université de Bologne (Italie), Université du Minas Gerais (UFMG, Brésil), Université du Nouveau-Brunswick, Faculté de droit (Canada), Université de Strasbourg (France), Université Western Sydney, IFIP, ACS (Australie).

Contributions individuelles de juges (Argentine, Costa Rica, Égypte, Guatemala, Paraguay, Tanzanie), d'agents des tribunaux et assistants de justice (Népal), d'avocats et de professionnels du droit (Cameroun, Colombie, Guatemala, Kenya, Mexique), d'organisations de la société civile (Népal), d'universitaires (Belgique, Costa Rica, Philippines, Suisse) et de journalistes (El Salvador).

Notes de fin

- 1 GPAI, *Algorithmic Transparency: A State-of-the-Art Report* (Partenariat mondial sur l'intelligence artificielle (GPAI), novembre 2024) [🔗](#); Nestor Maslej et d'autres, *The AI Index 2025 Annual Report* (AI Index Steering Committee, Institute for Human-Centered AI, Université Stanford, avril 2025), p. 456 [🔗](#); Sarah Muñoz-Cadena et Juan David Gutiérrez, « Repositorios de Algoritmos Públicos En El Mundo » (Sistemas de Algoritmos Públicos, Universidad de los Andes, mars 2025) [🔗](#).
- 2 Sarah Muñoz-Cadena et d'autres, 'Sistemas de IA En El Sector Público de América Latina y El Caribe' (Universidad de los Andes, mars 2025) [🔗](#).
- 3 Yannis Charalabidis, Rony Medaglia et Colin van Noordt, *Research Handbook on Public Management and Artificial Intelligence* (Edward Elgar Publishing, 2024), doi :10.4337/9781802207347 ; M. Manzoni et d'autres, *AI Watch Road to the Adoption of Artificial Intelligence by the Public Sector : A Handbook for Policymakers, Public Administrations and Relevant Stakeholders* (Bureau des publications de l'Union européenne, 2022), doi :10.2760/288757 (en ligne), 10.2760/693531 (imprimé) ; Regine Paul, Emma Carmel et Jennifer Cobbe, *Handbook on Public Policy and Artificial Intelligence* (Edward Elgar Publishing, 2024), doi :10.4337/9781803922171.
- 4 Pour un aperçu des différentes tâches réalisées avec les systèmes d'IA dans le secteur de la justice, voir F. Bell et autres, *AI Decision-Making and the Courts : A Guide for Judges, Tribunal Members and Court Administrators* (Australasian Institute of Judicial Administration, 2022) ; UNESCO, *Manuel de formation mondiale: l'IA & l'Etat de droit pour le pouvoir judiciaire* (Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), 2023) [🔗](#).

Pour des rapports et documents sur des outils d'IA spécifiques utilisés par les opérateurs judiciaires, voir Nikolaos Aletras et d'autres, « Predicting Judicial Decisions of the European Human Rights : A Natural Language Processing Perspective », éd. par Lexing Xie, *PeerJ Computer Science*, 2 (2016), pp. 1–19, doi :10.7717/peerj-cs.93 ; Urvashi Aneja et Dona Mathew, *Artificial Intelligence in India's Judicial System: A Case of Organised Irresponsibility?* (Digital Futures Lab, mars 2023) [🔗](#); Elsa Estevez, Sebastián Linares et Pablo Fillotrani, *PROMETEA : Transformando La Administración de Justicia Con Herramientas de Inteligencia Artificial* (Banco Interamericano de Desarrollo, 2020) [🔗](#); Collen Zvandasara Kufakwababa, « Artificial Intelligence Tools in Legal Work Automation: The Use and Perception of Tools for Document Discovery and Privilege Classification Processes in Southern African Legal Firms » (Université de Stellenbosch, 2021) [🔗](#); ; Goretty Carolina Martinez, « La Inteligencia Artificial y Su Aplicación al Campo Del Derecho », *Alegatos*, n° 82 (2012), pp. 827–46 ; Masha Medvedeva, Michel Vols, et Martijn Wieling, « Using Machine Learning to Predict Decisions

of the European Court of Human Rights », *Artificial Intelligence and Law*, 28.2 (2020), pp. 237–66, doi :10.1007/s10506-019-09255-y ; Ministro do Superior Tribunal de Justiça, *Artificial Intelligence: Technology Applied to Conflict Resolution in the Brazilian Judiciary* (Fundação Getulio Vargas, 2020) [🔗](#) ; Marion Oswald et autres, « Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and “Experimental” Proportionality », *Information & Communications Technology Law*, 27.2 (2018), pp. 223–50, doi :10.1080/13600834.2018.1458455 ; Víctor Saavedra et Juan Carlos Upegui, *Pretoria y La Automatización Del Procesamiento de Causas de Derechos Humanos* (Derechos Digitales América Latina y Dejusticia, mars 2021) [🔗](#).

Pour un référentiel d’outils et de projets d’IA développés par le CNJ brésilien, voir CNJ, *Repositório Nacional de Projetos de Software e Versionamento de Arquivos do – Git.jus* (Conselho Nacional de Justiça (CNJ)) [🔗](#) [consulté le 27 mai 2024].

- 5 Par exemple, la Commission européenne pour l’efficacité de la justice du Conseil de l’Europe (CEPEJ) a recensé plus d’une centaine d’outils d’IA qui contribuent à la gestion des processus judiciaires et aux tâches décisionnelles des juges. Voir CEPEJ, *Centre de ressources Cyberjustice et IA* (Commission européenne pour l’efficacité de la justice (CEPEJ), 6 mai 2024) [🔗](#).
- 6 Kahneman et ses collaborateurs qualifient le « bruit » d’erreur de jugement humain – distincte du biais – qui consiste en une variabilité indésirable dans la prise de décision qui aurait dû être identique, provoquant des incohérences aléatoires dans les décisions. En ce sens, « les jugements sont bruyants, car les personnes dont on attend qu’elles soient d’accord finissent à des moments très différents ». De plus, leur livre a analysé comment les algorithmes d’apprentissage automatique pouvaient réduire le bruit dans les scénarios de prise de décision judiciaire. Daniel Kahneman, Olivier Sibony et Cass R. Sunstein, *Noise. Pourquoi nous faisons des erreurs de jugement et comment les éviter* 1ère éd. (Little, Brown and Company, 2021), p. 4.
- 7 Nathan-Ross Adams, « Parker v Forsyth NO: Lessons for Using AI for Legal Research », *Michalsons*, 11 juillet 2023 [🔗](#) ; Robert J. Ambrogi, « Four Months After Launching Its “Homegrown” GenAI Tool, Law Firm Gunderson Dettmer Reports On Results So Far, New Features, And A Surprise on

Cost », *Lawsites*, 20 décembre 2023 [🔗](#) ; Francesca Bennetts et d’autres, « Responsible AI: Navigating the Risks and Embracing the Possibilities », *Allen Overy*, 12 novembre 2023 [🔗](#) ;

Juan David Gutiérrez, « Judges and Magistrates in Peru and Mexico Have ChatGPT Fever », *Tech Policy Press*, 19 avril 2023 [🔗](#) ; Juan David Gutiérrez, « ChatGPT in Colombian Courts: Why We Need to Have a Conversation about the Digital Literacy of the Judiciary », *VerfBlog*, 23 février 2023 [🔗](#) ; Juan David Gutiérrez, « AI Technologies in the Judiciary: Critical Appraisal of Large Language Models in Judicial Decision-Making », in *Handbook on Public Policy and AI*, éd. par Regine Paul, Emma Carmel et Jennifer Cobbe (Edward Elgar Publishing, 2014) [🔗](#) ; LexisNexis, « LexisNexis Launches Lexis+ AI, a Generative AI Solution with Linked Hallucination-Free Legal Citations », *LexisNexis*, 25 octobre 2023 [🔗](#) ;

Benjamin Weiser, « Here’s What Happens When Your Lawyer Uses ChatGPT », *The New York Times*, 27 mai 2023 [🔗](#) ; Benjamin Weiser et Jonah E. Bromwich, « Michael Cohen Used Fake Cases Cited by A.I. to Seek an End to Court Supervision », *The New York Times*, 29 décembre 2023 [🔗](#) ;

Mark Witten, « Applying Generative AI to Law: Opportunities and Risks », *Queen’s Law, Queen’s University*, 23 octobre 2023 [🔗](#) ;

Kieran Woboditsch-Velasco, « The Quiet Revolution: Generative AI’s Rise in Canadian Law Firms », *MyOpenCourt*, 25 novembre 2023 [🔗](#) ;

Nuria Morcillo, « Sentencias falsas, leyes extranjeras y filtrado de datos : los riesgos de usar ChatGPT se cuelan en los despachos de abogados », *El País*, 18 février 2025, section Economía [🔗](#) ;

Sara Merken, « Lawyers in Walmart Lawsuit Admit AI “hallucinated” Case Citations », *Reuters*, 10 février 2025, section Legal Industry [🔗](#) ;

Sam Harden, « You’d Think Lawyers Would Know by Now », *Team Do Something*, 14 février 2025 [🔗](#) ;

Kalliopi Terzidou, « Generative AI Systems in Legal Practice Offering Quality Legal Services While Upholding Legal Ethics », *International Journal of Law in Context*, 2025, pp. 1–22, doi :10.1017/S1744552325000047.

- 8 Juan David Gutiérrez, « Chapter 24: Critical Appraisal of Large Language Models in Judicial Decision-Making », in *Handbook on Public Policy and Artificial Intelligence* (Edward Elgar Publishing,

2024), pp. 323–38, doi :10.4337/9781803922171.00033.

- 9 Pour les directives et politiques publiées par les autorités judiciaires, voir :

UK Courts and Tribunals Judiciary, *Artificial Intelligence (AI) - Guidance for Judicial Office Holders* (Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, 12 décembre 2023) [🔗](#) ;

Conselho Nacional de Justiça, *Dispõe Sobre a Ética, a Transparência e a Governança Na Produção e No Uso de Inteligência Artificial No Poder Judiciário e Dá Outras Providências* (Conselho Nacional de Justiça (CNJ), 21 août 2020) [🔗](#) ;

Tribunaux de Nouvelle-Zélande, *Guidelines for Use of Generative Artificial Intelligence in Courts and Tribunals* (Nouvelle-Zélande, 7 décembre 2023) [🔗](#) ;

Federal Court, *Interim Principles and Guidelines on the Court's Use of Artificial Intelligence* (Cour fédérale, 20 décembre 2023) [🔗](#) ;

Tribunal fédéral, *Notice to the Parties and the Profession: The Use of Artificial Intelligence in Court Proceedings* (Cour fédérale, 7 mai 2024) [🔗](#) ;

Queensland Courts, *The Use of Generative Artificial Intelligence (AI) Guidelines for Responsible Use by Non-Lawyers*, 14 mai 2024 [🔗](#) ;

Cour suprême de Victoria, *Guidelines for Litigants: Responsible Use of Artificial Intelligence in Litigation* (Cour suprême de Victoria, mai 2024) [🔗](#) ;

Conseil judiciaire du Canada, *Guidelines for the Use of Artificial Intelligence in Canadian Courts* (Conseil judiciaire du Canada, septembre 2024) [🔗](#) ;

Cour suprême de l'État du Delaware, *Interim Policy on the Use of GenAI by Judicial Officers and Court Personnel* (Cour suprême de l'État du Delaware, 21 octobre 2024) [🔗](#) ;

Cour suprême de la République de Singapour, *Guide on the Use of Generative Artificial Intelligence Tools by Court Users* (Cour suprême de la République de Singapour, 23 septembre 2024) [🔗](#) ;

Consejo Superior de la Judicatura, *Lineamientos Para El Uso y Aprovechamiento Respetuoso, Responsable, Seguro y Ético de La Inteligencia Artificial En La Rama Judicial* (Consejo Superior de la Judicatura, de diciembre de 2024) [🔗](#) ;

Conselho Nacional de Justiça (CNJ), *Resolução N° 615, de 11 de Março de 2025. Estabelece Diretrizes para o Desenvolvimento, Utilização e Governança de Soluções Desenvolvidas Com Recursos de*

Intelligence Artificial no Poder Judiciário. (Conselho Nacional de Justiça (CNJ), 11 mars 2025) [🔗](#).

Il convient de noter que les Tribunaux de Nouvelle-Zélande ont publié trois lignes directrices distinctes pour les juges, les magistrats, les membres des tribunaux et le personnel de soutien judiciaire ; avocats ; et non juristes, respectivement.

De même, la Cour fédérale du Canada a publié des « Principles and Guidelines on the Court's Use of Artificial Intelligence » et un avis aux parties, aux justiciables qui se représentent seuls et aux autres intervenants concernant l'utilisation de l'IA dans les procédures judiciaires.

De même, la Cour suprême de Singapour, en septembre 2024, a publié son « Guide on the Use of Generative Artificial Intelligence Tools by Court Users » qui s'applique aux utilisateurs des tribunaux (« toute personne impliquée dans une affaire, y compris les procureurs, avocats, personnes auto-représentées ou témoins ») dans « toutes les affaires devant la Cour suprême, les tribunaux d'État et les tribunaux de justice familiale ».

En octobre 2024, la Cour supérieure de justice de Rio Negro en Argentine a approuvé son « Protocole de bonnes pratiques pour l'utilisation de l'intelligence artificielle générative (IAGen) » destiné aux magistrats, fonctionnaires et employés.

Pour les directives publiées par d'autres organismes publics, voir : CEPEJ, *Charte européenne d'éthique sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement* (Commission européenne pour l'efficacité de la justice (CEPEJ), 2019) ; CEPEJ-GT-CYBERJUST, *Utilisation de l'intelligence artificielle générative (IA) par les professionnels judiciaires dans un contexte professionnel* (Commission européenne pour l'efficacité de la justice (CEPEJ). Groupe de travail CEPEJ sur la cyberjustice et l'intelligence artificielle (CEPEJ-GT-CYBERJUST), 12 février 2024) [🔗](#).

- 10 Point 8(a) de l'annexe III du règlement sur l'IA de l'Union européenne.

- 11 Voir les articles 2 à 5 du chapitre III du règlement sur l'IA de l'Union européenne

- 12 Bell et d'autres, *AI Decision-Making and the Courts : Guide for Judges, Courts Members and Court Administrators* ; Juan David Gutiérrez, « Retos Éticos de La Inteligencia Artificial En El Proceso Judicial »,

- in *Derecho Procesal. #NuevasTendencias. XLI Congreso Colombiano de Derecho Procesal*, éd. par l'ICDP (Instituto Colombiano de Derecho Procesal (ICDP) y Universidad Libre., 2020) [🔗](#); MSI-NET, *Algorithms and Human Rights - Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications* (Conseil de l'Europe, 2018) [🔗](#); HCDH, *Taxonomy of Human Rights Risks Connected to Generative AI - Supplement to B-Tech's Foundational Paper on the Responsible Development and Deployment of Generative AI* (Bureau des droits humains des Nations Unies pour le Haut-Commissariat (HCDH), 2023) [🔗](#); Kalliopi Terzidou, « The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial », *Journal of Judicial Administration*, 31.3 (2022), pp. 154–68 ; Giulia Gentile, « AI in the Courtroom and Judicial Independence: An EU Perspective », *EUIdeas*, 22 août 2022 [🔗](#) .
- 13** Terzidou, « The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial » ; Kalliopi Terzidou, « Automated Anonymization of Court Decisions: Facilitating the Publication of Court Decisions through Algorithmic Systems », in *Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law, ICAIL '23* (Association for Computing Machinery, 2023), pp. 297–305, doi :10.1145/3594536.3595151.
- 14** Pour un aperçu des différents types et classifications d'outils d'IA, voir OCDE, « Cadre OCDE pour la classification des systèmes d'IA » (OECD Publishing, 2022) [🔗](#) .
- 15** La résolution 78/265 de l'Assemblée générale de l'ONU est disponible ici : [🔗](#)
- 16** Les Principes de Bangalore établissent des standards de déontologie judiciaire. « En juillet 2006, le Conseil économique et social des Nations Unies (ECOSOC) a adopté une résolution reconnaissant les Principes de Bangalore comme représentant un développement supplémentaire et complémentaires aux Principes fondamentaux des Nations Unies sur l'indépendance de la magistrature de 1985. ECOSOC a invité les États à encourager leurs magistratures à prendre en compte les Principes lors de la révision ou de l'élaboration de règles relatives à la déontologie judiciaire. » Les Principes de Bangalore sont disponibles ici : [🔗](#)
- 17** Sur l'utilité et l'acceptabilité de l'utilisation d'outils d'IA générative par les arbitres et les risques potentiels liés à l'adoption de ces technologies, voir João Ilhão Moreira et Jiawei Zhang, « ChatGPT as a Fourth Arbitrator? The Ethics and Risks of Using Large Language Models in Arbitration », *Arbitration International*, 2024, pp. 1–14, doi :10.1093/arbint/aiae031 ; David M. den Bakker, « Arbitration and AI: The Role of LLMs in Drafting Awards » (mémoire de master non publié, Université Radboud de Nimègue, 2024), doi :10.5281/zenodo.14601385.
- 18** Pour les lignes directrices et recommandations émises par les universités, les barreaux et les associations professionnelles, voir :
- Bell et autres, *AI Decision-Making and the Courts : Guide for Judges, Tribunal Members and Court Administrators* ;
- James E. Baker, Laurie Hobart et Mathew Mittels-teadt, *AI for Judges : A Framework* (Center for Security and Emerging Technology, décembre 2021) [🔗](#) ;
- Juan G. Corvalán et Mariana Sánchez Caparrós, *Guía de Directrices Para El Uso de ChatGPT e IA Generativa de Texto En La Justicia* (UBA IALAB, 24 novembre 2023) [🔗](#) ;
- Fédération des Barreaux d'Europe, *European Lawyers in the Era of ChatGPT : Guidelines on Lawyers Should Advantage the Opportunity Offers par les grands modèles de langage et l'IA générative* (Commission des nouvelles technologies de l'Association européenne des barreaux, 25 juillet 2023) [🔗](#) ;
- The State Bar of California, *Recommendations from Committee on Professional Responsibility and Conduct on Regulation of Use of Generative AI by Licensees*, 16 novembre 2023 ;
- The Bar Council, *Considerations When Using ChatGPT and Generative Artificial Intelligence Software Based on Large Language Models*, 30 janvier 2024 [🔗](#) ;
- Queensland Law Society, *Guidance Statement No 37 Artificial Intelligence in Legal Practice*, 6 juin 2024 [🔗](#) ;
- New South Wales Bar Association, *Issues Arising from the Use of AI Language Models (Including ChatGPT) in Legal Practice*, 12 juillet 2023 [🔗](#) ;

Guideline on the Use of AI in Arbitration (The Chartered Institute of Arbitrators (CIARB), mars 2025) [🔗](#).

L'American Association for the Advancement of Science (AAAS) a développé des documents pour les juges (articles et podcasts) sur l'IA dans les tribunaux, notamment des recommandations pour l'utilisation de l'IA dans le contentieux et la recherche juridique, entre autres. Les documents sont disponibles ici : [🔗](#).

- 19 UNESCO, « Formation des opérateurs judiciaires », 15 octobre 2024 [🔗](#).
- 20 UNESCO, « Le MOOC mondial de l'UNESCO sur l'IA et l'État de droit a impliqué des milliers d'opérateurs judiciaires », 20 avril 2023 [🔗](#).
- 21 UNESCO, *UNESCO Global Judges' Initiative: Survey on the Use of AI Systems by Judicial Operators* (Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), 2024) [🔗](#).
- 22 UNESCO, « Recommandation sur l'éthique de l'intelligence artificielle » (2022) [🔗](#). Pour une évaluation des outils d'IA générative à travers le prisme de la recommandation de l'UNESCO, voir : UNESCO, *Foundation Models Such as ChatGPT through the Prism of the UNESCO Recommendation's on the Ethics of Artificial Intelligence* (2023) [🔗](#).
- 23 UNESCO, *Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence* (2023) [🔗](#).
- 24 UNESCO, *Manuel de formation mondial: l'IA et l'État de droit pour le pouvoir judiciaire*.
- 25 Lors du processus de consultation publique des projets de lignes directrices, entre août et septembre 2024, l'UNESCO a reçu des retours individuels de 99 individus et organisations basés en Argentine, Australie, Belgique, Brésil, Bulgarie, Cameroun, Canada, Colombie, Costa Rica, Danemark, République dominicaine, Égypte, El Salvador, France, Guatemala, Inde, Irlande, Italie, Kenya, Lesotho, Libéria, Lituanie, Mexique, Népal, Royaume des Pays-Bas, Nouvelle-Zélande, Nigeria, Pakistan, Panama, Paraguay, Pérou, Philippines, Pologne, Sénégal, Espagne, Suisse, Tanzanie, Royaume des Pays-Bas, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, États-Unis d'Amérique et Yémen. De plus, l'UNESCO a reçu des contributions de membres de la magistrature

lors d'événements organisés en Inde, au Kenya, à Madagascar et au Népal.


- 26 Pour une explication de ce que signifie le « respect, la protection et la promotion des droits humains, des libertés fondamentales et de la dignité humaine » dans le contexte des technologies d'IA, consultez le Préambule et les sections III.1 et III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Pour des recommandations sur l'évaluation de l'impact éthique, voir « Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence » de l'UNESCO. Pour un aperçu de la manière dont les droits humains peuvent être affectés par l'IA générative, voir HCDH, *Taxonomy of Human Rights Risks Connected to Generative AI - Supplement to B-Tech's Foundational Paper on the Responsible Development and Deployment of Generative AI*.
- 27 Pour une explication du principe d'« équité et de non-discrimination », se référer à la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle.
- 28 *Ibid.* Voir aussi, ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général* (Nations Unies (ONU), 7 août 2024) [🔗](#) ; sur les biais de genre, voir UNESCO, *Challenging Systematic Prejudices: An Investigation into Gender Bias in Large Language Models*, 2024, [🔗](#).
- 29 Les Principes de Bangalore proposent les applications suivantes du principe d'« égalité » pertinentes pour ces lignes directrices : « 5.1. Un juge doit être conscient et comprendre la diversité de la société ainsi que les différences issues de diverses sources, y compris, mais sans s'y limiter, la race, la couleur, le sexe, la religion, l'origine nationale, la caste, le handicap, l'âge, l'état civil, l'orientation sexuelle, le statut social et économique ainsi que d'autres causes similaires (« motifs non pertinents »). 5.2. Un juge ne doit pas, dans l'exercice de ses fonctions judiciaires, manifester par des paroles ou un comportement de partialité ou de préjugé envers une personne ou un groupe pour des motifs non pertinents. 5.3. Un juge doit exercer ses fonctions judiciaires avec la considération appropriée pour toutes les personnes, telles que les parties, témoins, avocats, personnel judiciaire et collègues judiciaires, sans distinction fondée sur des motifs non pertinents, sans importance pour l'exercice approprié de ces fonctions. *Les Principes de Bangalore sur la*




déontologie judiciaire (Nations Unies, 2018), pp. 14–15 [🔗](#).

- 30 UNSDG, « Ne laisser personne de côté », *Groupe des Nations Unies pour le développement durable (GNUDS)* [🔗](#) [consulté le 26 mars 2025].
- 31 *Ibid.*
- 32 Pour une explication du principe de « vie privée et protection des données », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Voir aussi, ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*.
- 33 Pour une explication du principe de « proportionnalité et absence de préjudice », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle.
- 34 Un principe analogue appelé « principe de justification » a été proposé par Dan Jerker B. Svantesson, « Cybercrime and the Adoption of Artificial Intelligence Systems for Judicial Decision-Making in Criminal Justice Systems », *Commonwealth Cybercrime Journal*, 1.1 (2023), pp. 152–79.
- 35 Pour une explication du principe de « sécurité et sûreté », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle.
- 36 Pour une analyse des risques publiée pour les autorités visant à intégrer des modèles d'IA générative dans leurs flux de travail, ainsi que des contre-mesures possibles, voir *Generative AI Models. Opportunities and Risks for Industry and Authorities* (Federal Office for Information Security - Bundesamt für Sicherheit in der Informationstechnik (Office fédéral de la sécurité de l'information - Bundesamt für Sicherheit in der Informationstechnik, BSI), 4 avril 2024, [🔗](#)). Pour plus d'informations sur les lois sur l'accès à l'information, consultez la page web de l'UNESCO : [🔗](#).
- 37 Voir AI HLEG, *Ethics Guidelines for Trustworthy AI* (Commission européenne, 8 avril 2019), p. 17 [🔗](#).
- 38 IA HLEG, *Ethics Guidelines for Trustworthy AI*.
- 39 Pour une explication du principe de « transparence et d'explicabilité », se référer à la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. De même, comme l'exposent les lignes directrices publiées par le Conseil judiciaire canadien : « L'explicabilité désigne la nécessité pour les outils d'IA de fournir des explications claires et compréhensibles au sujet de leurs résultats, afin de faciliter l'interprétation, la confiance, la contestation ou l'acceptation des résultats IA dans les flux de travail critiques. L'explicabilité est comparable à l'exigence pour les juges de fournir des explications raisonnées à leurs décisions ». Voir, Canadian Judicial Council, *Guidelines for the Use of Artificial Intelligence in Canadian Courts*, p. 9.
- 40 Voir IA HLEG, *Ethics Guidelines for Trustworthy AI*. Pour un aperçu de la portée et du processus des audits algorithmiques, voir Ada Lovelace Institute et DataKind UK, *Examining the Black Box : Tools for Evaluating Algorithmic Systems*, 29 avril 2020 [🔗](#). ONU, *Droits humains dans l'administration de la justice : rapport du Secrétaire général*.
- 41 ONU, *Droits humains dans l'administration de la justice : rapport du Secrétaire général*.
- 42 Pour une explication du principe de « transparence et d'explicabilité », se référer à la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Pour un aperçu des divers instruments de transparence algorithmique adoptés par le secteur public dans le monde, voir GPAI, *Algorithmic Transparency : A State-of-the-Art Report*. Pour un aperçu du principe de « justice ouverte » dans le contexte des systèmes d'IA, voir Brian M Barry, « AI for Assisting Judicial Decision-Making: Implications for the Future of Open Justice », *Australian Law Journal*, 98.9 (2024), pp. 656–69.
- 43 Pour une explication du principe « Sensibilisation et alphabétisation », se référer à la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Voir aussi, ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*.
- 44 Pour une explication du principe de « responsabilité et de responsabilité », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle.
- 45 Pour une discussion approfondie sur ce que signifie la « responsabilité » dans le contexte de la gouvernance de l'IA, voir Claudio Novelli, Mariarosaria Taddeo et Luciano Floridi, « Accountability in Artificial Intelligence : What It Is and How It Works », *AI & SOCIETY*, 2023, doi :10.1007/s00146-023-01635-y.

- 46 Pour une explication du principe de « supervision humaine », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Voir aussi IA HLEG, *Ethics Guidelines for Trustworthy AI*.
- 47 Voir IA HLEG, *Ethics Guidelines for Trustworthy AI*.
- 48 Pour une explication du principe de « gouvernance et collaboration multipartites et adaptatives », voir la section III.2 de la Recommandation de l'UNESCO sur l'éthique de l'intelligence artificielle. Voir aussi IA HLEG, *Ethics Guidelines for Trustworthy AI* ; UNESCO et i4Policy, *Multistakeholder AI Development: 10 Building Blocks for Inclusive Policy Design* (Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), Innovation for Policy Foundation (i4Policy), 2022) [🔗](#) ; Anri Van der Spuy, *What If We All Governed the Internet? Advancing Multistakeholder Participation in Internet Governance* (UNESCO, 2017) [🔗](#).
- 49 Les Principes de Bangalore proposent les applications suivantes du principe d'« indépendance judiciaire » pertinentes pour ces Lignes directrices : « 1.1. Un juge exerce la fonction judiciaire de manière indépendante sur la base de son évaluation des faits par le juge et conformément à une compréhension consciente de la loi, libre de toute influence, incitation, pression, menace ou ingérence extérieure, directe ou indirecte, de quelque part que ce soit ou pour quelque raison que ce soit. 1.2. Un juge doit être indépendant vis-à-vis de la société en général et des parties à un litige que le juge doit trancher. 1.3. Un juge doit non seulement être exempt de liens inappropriés avec les branches exécutive et législative du gouvernement, mais doit aussi apparaître aux yeux d'un observateur raisonnable exempt de ces liaisons. *Les Principes de Bangalore sur la déontologie judiciaire*, pp. 8–9.
- 50 Les Principes de Bangalore proposent les applications suivantes du principe d'« impartialité » pertinentes pour ces Lignes directrices : « 2.1. Un juge exerce ses fonctions judiciaires sans faveur, 2.2. Un juge doit veiller à ce que sa conduite, tant en justice qu'en dehors, maintienne et renforce la confiance du public, de la profession juridique et des justiciables dans l'impartialité du juge et de la magistrature. *Les Principes de Bangalore sur la déontologie judiciaire*, p. 10.
- 51 La justice devrait envisager les compromis entre l'adoption d'outils sur mesure et d'outils prêts à l'emploi. Ces derniers peuvent être moins coûteux et plus rapides à mettre en œuvre, mais moins transparents et avec une personnalisation limitée. La première option peut être plus coûteuse (bien que les coûts puissent être réduits lorsqu'elle est construite sur du code open source) et nécessiter plus de temps pour sa mise en œuvre (y compris le développement et les tests), mais elle est plus transparente et en adéquation avec les besoins et objectifs de l'organisation.
- 52 IA HLEG, *Ethics Guidelines for Trustworthy AI*.
- 53 Par exemple, les erreurs induites par l'humain lors de la sélection ou de l'étiquetage des données dans les outils d'apprentissage automatique par apprentissage supervisé ; un manque statistique d'entrée, de représentation ou de données brutes pouvant renforcer le biais dans les outils d'apprentissage automatique non supervisé ; Erreurs d'environnement dans les outils d'apprentissage automatique par renforcement, etc.
- 54 « Les cartes modèles sont de courts documents accompagnant des modèles entraînés en apprentissage automatique qui fournissent une évaluation comparative dans diverses conditions, telles que selon différents groupes culturels, démographiques ou phénotypiques (par exemple, race, localisation géographique, sexe, [...] Type de peau [...] et des groupes intersectionnels (par exemple, âge et race, ou sexe et type de peau de Fitzpatrick) pertinents pour les domaines d'application visés. Les cartes modèles révèlent également le contexte dans lequel les modèles sont destinés à être utilisés, les détails des procédures d'évaluation des performances, ainsi que d'autres informations pertinentes. » Margaret Mitchell et d'autres, « Model Cards for Model Reporting » (présenté à la Conference on Fairness, Accountability, and Transparency, Atlanta, GA, États-Unis, 2019), pp. 220–29, doi :10.1145/3287560.3287596. Voir aussi : GPAI, *Algorithmic Transparency: A State-of-the-Art Report* ; Matías Valderrama, María Paz Herмосilla et Romina Garrido, *State of the Evidence : Algorithmic Transparency* (Open Government Partnership et GobLab (Universidad Adolfo Ibáñez), mai 2023) [🔗](#) ; Open Government Partnership, « CHAPTER 8: Algorithmic Transparency », dans *The Skeptic's Guide to Open Government*, 2022.

- 55 UNESCO, *Manuel de formation mondial: l'IA et l'État de droit pour le pouvoir judiciaire*, p. 42.
- 56 Voyez-vous, Brian Calvert, « AI Already Uses as Much Energy as a Small Country. It's Only the Beginning », *Vox*, 28 mars 2024 [🔗](#) ; Alexandra Sasha Luccioni, Sylvain Viguier et Anne-Laure Ligozat, « Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model », *Journal of Machine Learning Research*, 24.253 (2023), pp. 1–15 ; Sasha Luccioni, Yacine Jernite et Emma Strubell, « Power Hungry Processing: Watts Driving the Cost of AI Deployment? », dans *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24 (Association for Computing Machinery, 2024), pp. 85–99, doi :10.1145/3630106.3658542.
- 57 ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*.
- 58 Pour un aperçu des différents cadres de responsabilité pouvant être mis en œuvre avant le déploiement des systèmes d'IA, voir Dillon Reisman et d'autres, *Algorithmic Impact Assessments Report : A Practical Framework for Public Agency Accountability* (AI Now Institute, 9 avril 2018) [🔗](#) ; Ada Lovelace Institute et DataKind UK, *Examining the Black Box : Outils pour l'évaluation des systèmes algorithmiques* ; Alessandro Mantelero, « The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template », *Computer Law & Security Review*, 54 (2024), p. 106020, doi :10.1016/j.clsr.2024.106020 ; UNESCO, *Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence* ; Gouvernement des Pays-Bas, *Fundamental Rights and Algorithms Impact Assessment (FRAIA)* (Royaume des Pays-Bas, 31 juillet 2021) [🔗](#) .
- 59 Une « évaluation algorithmique de l'impact » est un processus qui évalue les impacts futurs – au niveau individuel, communautaire et sociétal – pouvant être associés au déploiement et à l'utilisation d'un système d'IA, et qui identifie comment gérer les conséquences négatives potentielles. L'évaluation peut porter sur les implications sociales et éthiques, les effets économiques, l'impact sur les droits humains et les conséquences environnementales. Reisman et autres, *Algorithmic Impact Assessments Report: A Practical Framework for Public Agency Accountability*.
- 60 Une « évaluation algorithmique des risques » est un processus qui cartographie les risques spécifiques pouvant être générés par le déploiement et l'utilisation d'un système d'IA, analyse les vulnérabilités et les risques potentiels, et identifie les mesures nécessaires pour prévenir et traiter les risques. Les évaluations portent sur les erreurs ou défaillances potentielles du système, les vulnérabilités de sécurité et de sûreté, ainsi que les violations réglementaires. Ada Lovelace Institute et DataKind UK, *Examining the Black Box: Tools for Assessing Algorithmic Systems*.
- 61 Une « évaluation de l'impact des droits fondamentaux » est un processus qui identifie, comprend et évalue les impacts négatifs potentiels des systèmes d'IA concernant les droits fondamentaux et prescrit comment traiter et prévenir les impacts négatifs. Mantelero, « The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template' ».
- 62 ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*, p. 19.
- 63 IA HLEG, *Ethics Guidelines for Trustworthy AI*, p. 31.
- 64 De plus, les implications potentielles des outils d'IA pour l'écosystème de l'information pourraient nécessiter une évaluation des risques plus spécifique fondée sur *les Guidelines for the Governance of Digital Platforms: Safeguarding Freedom of Expression and Access to Information through a Multistakeholder Approach* (UNESCO, 2023).
- 65 Ada Lovelace Institute et DataKind UK, *Examining the Black Box: Tools for Assessing Algorithmic Systems*.
- 66 Alessandro Mantelero, « Human Rights Impact Assessment and AI », dans *Beyond Data : Human Rights, Ethical and Social Impact Assessment in AI*, éd. par Alessandro Mantelero (T.M.C. Asser Press, 2022), pp. 45–91, doi :10.1007/978-94-6265-531-7_2 ; Alessandro Mantelero et Maria Samantha Esposito, « An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems », *Computer Law & Security Review*, 41 (2021), p. 105561, doi :10.1016/j.clsr.2021.105561 ; Gouvernement des Pays-Bas, *Fundamental Rights and Algorithms*

Impact Assessment (FRAIA) ; Gouvernement du Canada, *Algorithmic Impact Assessment Tool* (Canada, 12 avril 2023) .

- 67 Ada Lovelace Institute et DataKind UK, *Examining the Black Box: Tools for Assessing Algorithmic Systems*.
- 68 Une « évaluation de l'impact des droits humains » est un processus qui identifie l'impact des systèmes d'IA sur les droits humains et les titulaires de droits sociaux et fournit des recommandations sur la manière de les prévenir ou de les atténuer. Il est généralement réalisé en réponse à des situations critiques. Mantelero, « The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template ». Voir aussi, ONU, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*.
- 69 IA HLEG, *Ethics Guidelines for Trustworthy AI*, p. 31.
- 70 Alessandro Mantelero, « Human Rights Impact Assessment and AI », dans *Beyond Data : Human Rights, Ethical and Social Impact Assessment in AI*, éd. par Alessandro Mantelero (T.M.C. Asser Press, 2022), pp. 45–91, doi :10.1007/978-94-6265-531-7_2 ; Alessandro Mantelero et Maria Samantha Esposito, « An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems », *Computer Law & Security Review*, 41 (2021), p. 105561, doi :10.1016/j.clsr.2021.105561 ; Gouvernement des Pays-Bas, *Fundamental Rights and Algorithms Impact Assessment (FRAIA)* ; Gouvernement du Canada, *Algorithmic Impact Assessment Tool* (Canada, 12 avril 2023) .
- 71 Par exemple, le 21 juin 2022, la Cour de justice de l'Union européenne (CJUE) a rendu un jugement interdisant certains instruments spécifiques de « détection prédictive automatique des menaces » (outils qui « passent automatiquement au crible d'énormes quantités de données afin de prédire les menaces potentielles à la sécurité publique »). La CJUE a conclu que les outils d'IA basés sur l'apprentissage automatique « capables de modifier sans intervention humaine ni revoir le processus d'évaluation » étaient incompatibles avec les droits à la vie privée et à la protection des données personnelles. La CJUE a soutenu que « l'opacité qui caractérise la manière dont fonctionne la technologie de l'intelligence artificielle ... peuvent également priver les personnes concernées de leur droit à un recours judiciaire effectif ». Le jugement de la CJUE est disponible ici :  Pour un résumé et une analyse de l'affaire, voir Christian Thönnès et Niovi Vavoula, « Automated Predictive Threat Detection after Ligue Des Droits Humains : Implications for ETIAS and CSAM (Part I) », *VerfBlog*, 8 mai 2023 .
- 72 Nations Unies, *Droits humains dans l'administration de la justice : Rapport du Secrétaire général*, pp. 18–19.
- 73 Pour un compte rendu détaillé des vulnérabilités des outils d'IA générative, voir BSI, *Generative AI Models. Opportunités et risques pour l'industrie et les autorités*.
- 74 IA HLEG, *Ethics Guidelines for Trustworthy AI*, p. 17.
- 75 Voir Terzidou, « Automated Anonymization of Court Decisions: Facilitating the Publication of Court Decisions through Algorithmic Systems ».
- 76 Par exemple, la Commission européenne pour l'efficacité de la justice du Conseil de l'Europe (CEPEJ) et la magistrature brésilienne ont publié des dépôts documentant les systèmes d'IA adoptés dans leurs juridictions.
- 77 Pour une revue de divers instruments de transparence algorithmique, avec un accent sur les dépôts ou registres d'algorithmes publics, visant à divulguer des informations sur les systèmes d'IA adoptés dans le secteur public, voir GPAI, *Algorithmic Transparency : A State-of-the-Art Report*.
- 78 UNESCO, *Open Data for AI: What Now?* (2023) .
- 79 Voir par exemple le programme du MOOC mondial de l'UNESCO sur l'IA et l'État de droit. UNESCO, « UNESCO Global MOOC on AI and the Rule of Law Engaged Thousands of Judicial Operators ».
- 80 Des instruments d'évaluation d'impact, tels que « Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence » de l'UNESCO et le cadre d'évaluation de l'impact des droits humains proposé par Mantelero, peuvent être utilisés pour évaluer les impacts sur les droits humains des outils utilisés par les acteurs judiciaires. Voir Mantelero, « Human Rights Impact Assessment and AI ».
- 81 Dans ce contexte, les Principes de Bangalore proposent les applications suivantes du principe de

- « compétence et de diligence » pertinentes pour ces Lignes directrices : « 6.3. Un juge doit prendre des mesures raisonnables pour maintenir et améliorer ses connaissances, compétences et qualités personnelles nécessaires à l'exercice approprié des fonctions judiciaires, en profitant à cette fin de la formation et des autres installations qui devraient être mises à la disposition des juges, sous contrôle judiciaire. *Les Principes de Bangalore sur la déontologie judiciaire*, p. 16.
- 82** Juan David Gutiérrez, « ¿Están los Jueces En Capacidad de Usar Modelos de Lenguaje a Gran Escala (LLMs) ? », *Revista EXCEJLENCIA*, 7 (2023), pp. 10–15 ; David Bawden, « Origins and Concepts of Digital Literacy », in *Digital Literacies : Concepts, Politiques et Pratiques*, édité par Colin. Lankshear et Michele. Knobel (Peter Lang, 2008), pp. 17–32 [🔗](#).
- 83** Mantelero, « Human Rights Impact Assessment and AI » ; Mantelero et Esposito, « An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems ».
- 84** Pour des recommandations générales sur la manière dont les opérateurs judiciaires peuvent évaluer les questions de vie privée et de protection des données, voir les *Guidelines for Judicial Actors on Privacy and Data Protection* de l'UNESCO (2022) [🔗](#).
- 85** UK Courts and Tribunals Judiciary, *Artificial Intelligence (AI) - Guidance for Judicial Office Holders*, p. 3.
- 86** Voir Courts of New Zealand, *Guidelines for Use of Generative Artificial Intelligence in Courts and Tribunals*; UK Courts and Tribunals Judiciary, *Artificial Intelligence (AI) - Guidance for Judicial Office Holders*, p. 6.
- 87** UK Courts and Tribunals Judiciary, *Artificial Intelligence (AI) - Guidance for Judicial Office Holders*, p. 3.
- 88** Courts of New Zealand, *Guidelines for Use of Generative Artificial Intelligence in Courts and Tribunals*, p. 2.
- 89** Pour des directives similaires concernant la production de preuves destinées aux utilisateurs des tribunaux, voir la Cour suprême de la République de Singapour, *Guide on the Use of Generative Artificial Intelligence Tools by Court Users*.
- 90** Les *Lignes directrices* ne font pas allusion au terme « hallucination » car cette utilisation contribue à l'anthropomorphisation des systèmes d'IA, en particulier des chatbots.
- 91** BSI, *Generative AI Models. Opportunities and Risks for Industry and Authorities*, p. 9 ; Voir aussi UNESCO, *Challenging Systematic Prejudices: An Investigation into Gender Bias in Large Language Models*, 2024, [🔗](#).
- 92** « Les LLM génèrent généralement des textes linguistiquement corrects et convaincants et sont capables de faire des déclarations sur une grande variété de sujets. Cela peut donner l'impression d'une performance humaine, conduisant à une confiance excessive dans les énoncés et la performance des modèles (ce qu'on appelle biais d'automatisation). Pour les utilisateurs, cela peut les conduire à tirer de mauvaises conclusions à partir des textes générés ou à accepter des affirmations sans les remettre en question » BSI, *Generative AI Models. Opportunities and Risks for Industry and Authorities*, p. 10.
- 93** BSI, *Generative AI Models. Opportunities and Risks for Industry and Authorities*, p. 9.
- 94** Par exemple, pour des recommandations sur la manière de citer les résultats des chatbots IA avec le système de citation APA, voir Timothy McAdoo, « How to Cite ChatGPT », *APA Style*, 2024.
- 95** Pour des recommandations similaires, voir Cour suprême de l'État du Delaware, *Interim Policy on the Use of GenAI by Judicial Officers and Court Personnel*.
- 96** Emily M. Bender et al., « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? », in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21* (Association for Computing Machinery, 2021), pp. 610–23 (p. 616), doi :10.1145/3442188.3445922.
- 97** Matthew Dahl et d'autres, « Large Legal Fictions : Profiling Legal Hallucinations in Large Language Models », [🔗](#); 2024 ; Varun Magesh et d'autres, « Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools », 2024.

L'IA est entrée dans la salle d'audience. Les juges et les professionnels du droit ont commencé à utiliser des systèmes d'IA pour gagner du temps et se concentrer sur des tâches importantes. Cette évolution technologique pourrait améliorer significativement les droits humains et l'administration de la justice.

Comment garantir que les avantages de l'IA l'emportent sur ses risques (biais, discrimination, erreurs et hallucinations) ?

Les tribunaux ont le pouvoir de contrôler et de façonner la manière dont les systèmes d'IA sont développés et utilisés afin de protéger l'indépendance judiciaire, l'État de droit et les droits humains.

Les acteurs du monde judiciaire ont appelé l'UNESCO à élaborer ces Lignes directrices afin de fournir des recommandations et des principes sur les mesures à prendre avant, pendant et après le déploiement de systèmes d'IA.

Ces lignes directrices aideront les chefs de juridictions à concevoir des cadres politiques adaptés au contexte local. Elles aideront les juges, le personnel judiciaire et les professionnels du droit, quelles que soient leurs compétences numériques, à développer une utilisation professionnelle de l'IA.

Le programme « IA et État de droit » de l'UNESCO dote les juges et professionnels du droit du monde entier des compétences nécessaires pour faire respecter la justice, l'État de droit et les droits humains à l'ère de l'IA.

**LIGNES DIRECTRICES POUR L'UTILISATION
DES SYSTÈMES D'IA DANS LES TRIBUNAUX**



unesco