

1425^e réunion, 16 février 2022

10 Questions juridiques

10.1 Comité Ad hoc sur l'intelligence artificielle (CAHAI)

b. Éléments potentiels d'un cadre juridique sur l'intelligence artificielle, fondés sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit

Pour examen par le GR-J lors de sa réunion du 3 février 2022

Partie I : Introduction

I Contexte

1. Le présent document rend compte des résultats des travaux du Comité ad hoc sur l'intelligence artificielle du Conseil de l'Europe (CAHAI) concernant les éléments potentiels d'un cadre juridique pour le développement, la conception et l'application de l'intelligence artificielle, fondés sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit.

2. Il a été rédigé par deux groupes de travail du CAHAI, le Groupe d'élaboration des politiques (CAHAI-PDG) et le Groupe sur les cadres juridiques (CAHAI-LFG), en tenant compte des résultats de la consultation multipartite menée au printemps 2021 par le troisième groupe de travail, le Groupe de consultation et de sensibilisation (CAHAI-COG). Il a été examiné et adopté par le CAHAI lors de sa sixième réunion plénière, qui s'est déroulée du 30 novembre au 2 décembre 2021, puis soumis à l'examen du Comité des Ministres, conformément au mandat du CAHAI.

II Considérations générales

3. Le CAHAI constate que l'application des systèmes d'intelligence artificielle (IA) est à même de favoriser la prospérité et le bien-être individuel et social en apportant progrès et innovation, mais que, dans le même temps, certaines applications des systèmes d'IA suscitent des inquiétudes en raison des risques qu'elles pourraient présenter pour les droits de l'homme, la démocratie et l'État de droit.

4. Pour prévenir et/ou atténuer ces risques de manière effective, le CAHAI considère qu'un cadre juridique sur l'IA approprié fondé sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit devrait être un instrument transversal juridiquement contraignant. Le CAHAI note qu'en plus du cadre juridiquement contraignant transversal proposé qui énonce des normes et principes généraux et des normes légales spécifiques, des instruments juridiquement contraignants et/ou non juridiquement contraignants existants ou futurs pourraient être nécessaires au niveau sectoriel, afin de fournir des orientations plus détaillées pour garantir que la conception, le développement et l'application de l'IA se déroulent conformément aux droits de l'homme, à la démocratie et à l'État de droit dans des domaines spécifiques.

¹ Ce document a été classé en diffusion restreinte jusqu'à la date de son examen par le Comité des Ministres.

5. L'instrument juridiquement contraignant transversal devrait mettre l'accent sur la prévention et/ou la réduction des risques induits par les applications des systèmes d'IA susceptibles d'entraver la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit, et ce tout en valorisant les applications bénéfiques pour la société. Il devrait être étayé par une approche fondée sur les risques : les exigences juridiques relatives à la conception, au développement et à l'utilisation des systèmes d'IA devraient être proportionnées à la nature du risque qu'ils représentent pour les droits de l'homme, la démocratie et l'État de droit. Les principes de base qui permettent de déterminer un tel risque (par exemple, les exigences de transparence) devraient être applicables à tous les systèmes d'IA.

6. Conformément à l'article 1 d du Statut du Conseil de l'Europe, les questions relatives à la défense nationale ne sont pas de la compétence du Conseil de l'Europe et ne sont donc pas couvertes par le champ d'application d'un instrument juridiquement contraignant (ou non contraignant) du Conseil de l'Europe. Le CAHAI est d'avis que la question de savoir si ce champ d'application pourrait couvrir le "double usage" et la sécurité nationale devrait être examinée plus en profondeur dans le contexte de l'élaboration d'un cadre juridique du Conseil de l'Europe sur l'IA, en tenant compte des difficultés éventuelles à cet égard.

7. Les diverses questions juridiques soulevées par l'application des systèmes d'IA ne sont pas spécifiques aux États membres du Conseil de l'Europe, mais transnationales par nature, en raison de l'implication de grand nombre d'acteurs mondiaux et des effets globaux qu'ils engendrent. Le CAHAI recommande donc qu'un instrument juridiquement contraignant transversal du Conseil de l'Europe, quoique d'évidence fondé sur les normes de l'Organisation, soit rédigé de manière à faciliter l'adhésion d'États qui n'en sont pas membres qui partagent ces normes. Cela permettra non seulement d'augmenter de manière significative l'impact et l'efficacité de l'instrument proposé, mais aussi d'établir des conditions équitables indispensables pour les acteurs concernés, y compris l'industrie et les chercheurs en IA, lesquels opèrent souvent par-delà les frontières nationales et les régions du monde. Les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit ont un caractère suffisamment universel pour que cette solution soit réaliste. D'autres traités du Conseil de l'Europe s'appliquent d'ores et déjà au-delà du continent européen, notamment la Convention de Budapest (sur la cybercriminalité) ainsi que la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE n°108), qui comptent actuellement respectivement 66 et 55 États parties, dont beaucoup ne sont pas membres du Conseil de l'Europe.

8. Il est aussi recommandé que, pour assurer une cohérence juridique tant au niveau mondial que régional, un instrument juridiquement contraignant transversal du Conseil de l'Europe devrait tenir compte des cadres juridiques et réglementaires existants et à venir des autres enceintes internationales et régionales, en particulier le système des Nations Unies (y compris l'UNESCO), l'Union européenne et l'Organisation de coopération et de développement économiques, qui s'investissent toutes actuellement dans l'élaboration de diverses formes de normes en lien avec les systèmes d'IA.

9. Le CAHAI note que la finalité d'un cadre juridique international ne devrait pas être de définir des paramètres techniques détaillés pour la conception, le développement et l'application des systèmes d'IA, mais d'établir des principes et des normes de base régissant le développement, la conception et l'application des systèmes d'IA, et de définir, de façon cohérente et délibérée, des règles permettant de savoir si et dans quelles conditions des systèmes d'IA susceptibles de présenter des risques pour la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit peuvent être développées, conçues et appliquées par tous types d'organisations, y compris les acteurs publics comme privés.

10. Dans la Partie II (*chapitres III à XI*, développés par le CAHAI-LFG), les éléments qui pourraient faire partie d'un instrument juridiquement contraignant transversal sont présentés. La Partie III (*chapitres XII et XIII*, développés par le CAHAI-PDG) présente les éléments qui pourraient faire partie d'éventuels instruments juridiques supplémentaires.

PARTIE II : Éléments pour un instrument juridiquement contraignant transversal

III Éléments concernant l'objet et le but, le champ d'application et les définitions

11. En ce qui concerne *l'objet et le but* de l'instrument juridiquement contraignant transversal, le CAHAI recommande, en particulier, de préciser que le but de l'instrument est de veiller à ce que le développement, la conception et l'application des systèmes d'IA, que ces activités soient entreprises par des acteurs publics ou privés, soient en totale adéquation avec le respect des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit. En outre, il convient d'indiquer que l'instrument doit faciliter la coopération de ses Parties à cette fin, tant au niveau international qu'au niveau national, et que les mécanismes de suivi nécessaires doivent être mis en place. Enfin, l'objet et le but devraient mettre en avant la nécessité de créer un cadre juridique commun contenant certaines normes minimales pour le développement, la conception et l'application de l'IA en lien avec les droits de l'homme, la démocratie et l'État de droit.

12. Le CAHAI considère que l'instrument juridiquement contraignant transversal devrait contenir une disposition délimitant son *champ d'application*. Cette disposition devrait préciser que l'instrument est applicable à tout le développement, la conception et l'application de systèmes d'IA, que ces activités soient entreprises par des acteurs publics ou privés, en mettant l'accent sur les systèmes dont on estime qu'ils présentent des risques potentiels pour la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit. Si nécessaire, de potentielles exceptions au champ d'application devraient également être abordées.

13. En ce qui concerne les *définitions*, le CAHAI considère que, au minimum, un instrument juridiquement contraignant transversal devrait définir les expressions et termes suivants : « *système d'intelligence artificielle* », « *cycle de vie* », « *fournisseur d'IA* », « *utilisateur d'IA* », « *sujet d'IA* » et « *préjudice illégal* ». Le CAHAI recommande que toutes les définitions utilisées soient, dans la mesure du possible, compatibles avec les définitions analogues utilisées dans d'autres instruments pertinents portant sur l'IA. De plus, les définitions devraient être rédigées avec soin pour garantir, d'une part, qu'elles respectent les exigences de précision juridique et, d'autre part, qu'elles sont suffisamment abstraites pour rester valides malgré les évolutions technologiques futures des systèmes d'IA.

IV Éléments concernant les principes fondamentaux de protection de la dignité humaine et du respect des droits de l'homme, de la démocratie et de l'État de droit

14. Le CAHAI considère qu'il est nécessaire qu'un instrument juridiquement contraignant transversal contienne certains *principes fondamentaux de protection de la dignité humaine et du respect des droits de l'homme, de la démocratie et de l'État de droit*, qui devraient s'appliquer à tout le développement, la conception et l'application de systèmes d'IA, que l'acteur soit public ou privé.

15. Dans le même temps, le CAHAI, reconnaissant les risques de duplication, voire de fragmentation des normes générales du droit international existantes, notamment le droit relatif aux droits de l'homme, recommande que ces principes fondamentaux soient rédigés de telle manière que les risques de duplication ou de fragmentation injustifiés soient dûment réduits au minimum. Il faudra donc, entre autres, adapter davantage les droits et obligations relatifs aux droits de l'homme, à la démocratie et à l'État de droit aux fins de cet instrument uniquement si, après un examen approfondi, on arrive à la conclusion que les normes existantes, dans leur forme actuelle, ne protègent pas suffisamment les droits des personnes dans le contexte spécifique du développement, de la conception et de l'application des systèmes d'IA.

16. En ce qui concerne le concept de « dignité humaine », le CAHAI note qu'il est universellement admis que la dignité de la personne humaine constitue le véritable fondement des droits de l'homme (voir également la place importante accordée à ce concept dans le préambule de la Déclaration universelle des droits de l'homme de 1948). Le CAHAI considère qu'il est particulièrement pertinent d'utiliser ce concept dans un instrument juridiquement contraignant transversal portant sur les effets potentiellement négatifs du développement, de la conception et de l'application des systèmes d'IA sur les droits fondamentaux des personnes.

17. Le CAHAI note par ailleurs que certaines des dispositions relatives à ces éléments particuliers peuvent être formulées comme des droits positifs directs des individus, ou bien comme des obligations, pour les Parties, de veiller à l'introduction dans leur droit et pratique interne de mesures visant à protéger les droits des individus relativement aux systèmes d'IA. Au vu de ses délibérations, le CAHAI serait plutôt favorable, si cela est possible et nécessaire, à une solution alliant à la fois la mise en place de certains *droits* directs, concrets et *positifs des individus relativement au développement, à la conception et à l'application des systèmes d'IA*, et la *mise en place de certaines obligations incombant aux Parties*, afin d'assurer une application plus uniforme de l'instrument juridiquement contraignant transversal parmi les Parties.

V Éléments concernant la classification des risques liés aux systèmes d'IA et les applications d'intelligence artificielle interdites

18. Le CAHAI considère qu'un instrument juridiquement contraignant transversal devrait prévoir la mise en place d'une méthode de *classification des risques* liés aux systèmes d'IA mettant l'accent sur les droits de l'homme, la démocratie et l'État de droit. Les critères utilisés pour évaluer l'impact de l'application des systèmes d'IA à cet égard devraient être concrets, clairs, et fondés sur une base objective et l'évaluation devrait être effectuée de manière équilibrée, afin d'assurer la sécurité juridique tout en permettant la nuance.

19. En particulier, le CAHAI considère que la classification des risques devrait comporter un certain nombre de catégories (par exemple, « risque faible », « risque élevé », « risque inacceptable ») basées sur une évaluation des risques relatifs à la jouissance des droits de l'homme, le fonctionnement de la démocratie et le respect de l'État de droit. La classification des risques sera basée sur un examen initial visant à déterminer si une analyse d'impact HUDERIA (« *Human Rights, Democracy and Rule of Law Impact Assessment* ») complète est nécessaire (cf. chapitre XII) tout comme l'analyse d'impact elle-même peut avoir une incidence sur le maintien ou la modification de la classification initiale des risques du système d'IA en question. Cette analyse d'impact est considérée comme un élément du cadre juridique global sur les systèmes d'IA proposé par le CAHAI. Cependant, le modèle spécifique d'HUDERIA ne doit pas nécessairement faire partie intégrante d'un éventuel instrument juridiquement contraignant.

20. En ce qui concerne les critères qui pourraient être considérés aux fins de l'analyse des risques, il est fait référence aux éléments énumérés au paragraphe 51 au chapitre XII ci-dessous. Certains de ces critères devront peut-être être inscrits dans l'instrument juridiquement contraignant, afin de garantir qu'ils sont dûment pris en compte et appliqués de manière cohérente.

21. En ce qui concerne les *applications interdites de l'IA* (ce que l'on appelle les « lignes rouges » ou « *risque inacceptable* »), le CAHAI considère qu'un instrument juridiquement contraignant transversal devrait prévoir la possibilité d'imposer un moratoire total ou partiel ou une interdiction visant l'application des systèmes d'IA qui, en vertu de la classification des risques susmentionnée, sont considérés comme présentant un risque inacceptable d'entrave à la jouissance des droits de l'homme, au bon fonctionnement de la démocratie et au respect de l'État de droit. Une telle possibilité devrait également être envisagée pour la recherche et le développement de certains systèmes d'IA qui présentent un risque inacceptable. Le CAHAI souhaite notamment attirer l'attention, par exemple, sur certains systèmes d'IA utilisant la biométrie pour identifier, catégoriser ou déduire les caractéristiques ou les émotions des individus, en particulier s'ils conduisent à une surveillance de masse, et sur les systèmes d'IA utilisés pour la notation sociale afin de déterminer l'accès aux services essentiels, en tant qu'applications pouvant nécessiter une attention particulière, en tenant compte des potentielles exceptions légitimes. Un moratoire ou une interdiction ne devraient toutefois être envisagés que lorsqu'un risque inacceptable pour les droits de l'homme, la démocratie et l'État de droit a été identifié sur une base objective et qu'après un examen minutieux, il n'existe pas d'autres mesures applicables et tout aussi efficaces pour atténuer ce risque et compte tenu du domaine d'application spécifique. Des procédures de révision devraient être mises en place pour permettre de lever une interdiction ou un moratoire si les risques sont suffisamment réduits ou si des mesures d'atténuation appropriées deviennent disponibles, sur une base objective, pour ne plus présenter de risque inacceptable.

VI Éléments concernant le développement, la conception et l'application des systèmes d'intelligence artificielle en général

22. Le CAHAI recommande qu'un instrument juridiquement contraignant transversal devrait contenir un certain nombre de dispositions applicables à tout le développement, la conception et l'application de systèmes d'IA, afin de permettre leur classification de manière appropriée quant au risque potentiel qu'ils font peser sur la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit, et d'assurer la conformité avec ceux-ci en fixant des garanties minimales. Celles-ci peuvent inclure, par exemple, des dispositions concernant la transparence des systèmes d'IA. Conformément à l'approche fondée sur les risques mentionnée ci-dessus, des dispositions supplémentaires devraient être rendues applicables aux systèmes d'IA sur la base de la classification de leurs risques et proportionnellement à celle-ci, afin de garantir que les risques qu'ils présentent pour les droits de l'homme, la démocratie et l'État de droit sont dûment atténués.

23. Un instrument juridiquement contraignant transversal devrait, de façon générale, stipuler que, sous réserve de certaines limitations, *le développement et la conception des systèmes d'IA, ainsi que la recherche sur ces systèmes devraient être réalisés librement, en tenant dûment compte de la sécurité et de la sûreté*, et dans le plein respect des normes du Conseil de l'Europe en matière de droits de l'homme.

24. En outre, le CAHAI recommande de prévoir une disposition encourageant les Parties à créer des « *bacs à sable réglementaires* » afin de stimuler l'innovation responsable en matière de systèmes d'IA en permettant l'essai de ces systèmes, sous la supervision de l'autorité de réglementation nationale compétente, tout en assurant le respect des normes énoncées dans la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE n° 108) et son Protocole d'amendement (STCE n° 223), ainsi qu'avec les normes énoncées dans cet instrument juridiquement contraignant transversal sur la conception, le développement et l'application de l'IA, et toutes autres normes applicables.

25. Pour promouvoir une approche multipartite et afin de sensibiliser la société à l'impact du développement, de la conception et de l'application des systèmes d'IA, le CAHAI considère qu'il est utile de prévoir une disposition appelant les Parties à promouvoir les *délibérations publiques* fondées sur des preuves et l'engagement sur ce sujet. Le libellé de cette disposition pourra s'inspirer de l'article 28 de la Convention pour la protection des droits de l'homme et de la dignité de l'être humain à l'égard des applications de la biologie et de la médecine (STE n° 164).

26. Le CAHAI propose de prévoir une disposition sur la *prévention des préjudices illégaux* pouvant être causés par le développement, la conception et l'application des systèmes d'IA, y compris clarifiant le concept de "préjudice illégal" dans le contexte de l'instrument transversal sur l'IA, les droits de l'homme, la démocratie et l'État de droit.

27. Le CAHAI propose en outre de prévoir une disposition sur le respect de *l'égalité de traitement et de la non-discrimination* des personnes en lien avec le développement, la conception et l'application des systèmes d'IA, afin d'éviter que des biais injustifiés ne soient intégrés dans ces systèmes et l'utilisation de systèmes d'IA entraînant des effets discriminatoires.

28. Pour les mêmes raisons, un instrument juridiquement contraignant transversal devrait contenir des dispositions visant à garantir que *l'égalité de genre* et les droits relatifs aux *groupes vulnérables et aux personnes en situation de vulnérabilité*, y compris *les enfants*, sont respectés tout au long du cycle de vie des systèmes d'intelligence artificielle.

29. Le CAHAI considère également qu'il est prudent de prévoir une disposition sur la *gouvernance des données* applicable aux systèmes d'IA, en conformité avec et en s'appuyant sur la Convention du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE n° 108) et son Protocole d'amendement (STCE n° 223). Cela peut inclure l'obligation de mettre en place des mécanismes de gouvernance des données pour évaluer et garantir l'exactitude, l'intégrité, la sécurité et la représentativité des données, d'une manière qui soit adaptée à la finalité prévue du système et qui soit proportionnée.

30. Enfin, le CAHAI recommande de prévoir des dispositions en matière de *robustesse, sécurité et cybersécurité, transparence, explicabilité, audibilité et responsabilité* tout au long de leur cycle de vie. Il est à noter que, pour le CAHAI, les concepts de « transparence », d'« explicabilité » et de « responsabilité » revêtent une importance capitale pour la protection des droits des individus dans le contexte des systèmes d'IA. En outre, le CAHAI recommande que la question de la durabilité des systèmes d'IA tout au long de leur cycle de vie soit examinée de manière appropriée.

31. Enfin et surtout, un instrument juridiquement contraignant transversal devrait contenir une disposition visant à garantir le niveau nécessaire de *contrôle humain* sur les systèmes d'IA et leurs effets, et ce tout au long de leur cycle de vie.

VII Éléments concernant le développement, la conception et l'application des systèmes d'intelligence artificielle dans le secteur public

32. Le développement, la conception et l'application de systèmes d'IA dans le secteur public suscitent certaines inquiétudes quant à la manière de garantir le respect des droits de l'homme, de la démocratie et de l'État de droit lorsque des systèmes d'IA sont utilisés pour prendre ou informer des décisions qui ont un impact sur les droits et obligations des personnes physiques et morales. Cela dit, le CAHAI souligne que toutes les applications de l'IA dans le secteur public ne présentent pas de risques pour la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit. Il est donc important d'examiner attentivement et au cas par cas le risque potentiel posé par l'application d'un système d'IA donné. En conséquence, la distinction devrait être faite entre, d'un côté, les systèmes d'IA qui peuvent avoir une incidence sur les droits de l'homme, la démocratie ou l'État de droit, et, de l'autre, les systèmes d'IA qui, bien qu'utilisés par les mêmes autorités publiques, ne présentent pas de tels risques.

33. Partant de l'hypothèse qu'un instrument juridiquement contraignant transversal devrait être par nature général, le CAHAI recommande que cet instrument se penche sur les risques potentiels liés au développement, à la conception et à l'application de systèmes d'IA aux fins du *maintien de l'ordre*, de *l'administration de la justice* et de *l'administration publique*. En ce qui concerne l'« administration publique », le CAHAI note qu'un instrument juridiquement contraignant transversal ne devrait pas aborder la pléthore d'activités administratives spécifiques entreprises par les autorités publiques, tels les soins de santé, l'éducation, les prestations sociales, etc., mais se limiter à des prescriptions générales relatives à l'utilisation responsable des systèmes d'IA dans l'administration. Les questions liées aux différents secteurs de l'administration peuvent, le cas échéant, être traitées dans des instruments sectoriels appropriés.

34. Le CAHAI estime qu'un instrument juridiquement contraignant transversal, s'agissant du développement, de la conception et de l'application de systèmes d'IA dans le secteur public, devrait, au minimum, comprendre des dispositions sur *l'accès à un recours effectif*, le nécessaire *droit à un contrôle humain* des décisions prises ou guidées par un système d'IA, sauf si des motifs concurrents légitimes et impérieux l'excluent, ainsi que *l'obligation pour les autorités publiques de mettre en œuvre un examen humain adéquat pour les processus qui sont informés ou soutenus par des systèmes d'IA et de fournir aux personnes physiques ou morales concernées des informations pleinement pertinentes* sur le rôle des systèmes d'IA dans la prise ou l'aide à la prise de décisions les concernant, sauf si des raisons concurrentes légitimes et impérieuses s'opposent à ou limitent l'examen ou la communication de ces informations. En outre, les Parties devraient être tenues de veiller à ce que des *garanties suffisantes et efficaces contre les pratiques arbitraires et abusives* dues à l'application d'un système d'IA dans le secteur public soient prévues dans leur droit interne.

35. The CAHAI note également que les Parties devraient veiller au respect des normes relatives aux systèmes d'IA en ce qui concerne les droits de l'homme, la démocratie et l'État de droit dans la mesure où les acteurs privés agissant en leur nom sont concernés.

VIII Éléments concernant la démocratie et la gouvernance démocratique

36. Tout en reconnaissant que l'IA peut jouer un rôle positif en matière de fonctionnement de la démocratie et de gouvernance démocratique pour favoriser les processus inclusifs et participatifs, le CAHAI s'inquiète également de l'utilisation potentielle de l'IA dans le but d'influer illégalement ou indûment sur les processus démocratiques. Le façonnement de l'opinion publique aux moyens de l'IA, ainsi que les effets dissuasifs potentiels résultant de l'utilisation de l'IA, devrait en conséquent être considérés dans le contexte d'un potentiel instrument juridiquement contraignant, alors que les questions spécifiques ayant égard à la manipulation des élections comme le *microciblage*, le *profilage* et la *manipulation de contenus* (y compris les « hypertrucages »/« *deep fakes* ») pourraient être abordées dans des instruments plus sectoriels.

37. Le rôle des entités privées, par exemple des plateformes en ligne qui contribuent à façonner la sphère publique, devrait également être considéré à cet égard, dans la mesure où la concentration croissante du pouvoir économique et des données pourrait saper les processus démocratiques.

38. À cet égard, le CAHAI souligne la nécessité de respecter *le droit à la liberté d'expression, notamment la liberté de se forger et d'avoir des opinions et de recevoir et de communiquer des idées et des informations politiques, ainsi que le droit à la liberté de réunion et d'association*, dans le but d'assurer que tous les partis et les groupes d'intérêt ont accès aux processus démocratiques dans les mêmes conditions et qu'un espace libre de débat public puisse être assuré.

IX Éléments concernant les garanties

39. Le CAHAI recommande qu'un instrument juridiquement contraignant transversal comprenne une série de dispositions sur les garanties juridiques qui doivent s'appliquer à toutes les applications des systèmes d'IA utilisés pour prendre ou guider des décisions ayant une incidence sur les droits et autres intérêts significatifs des personnes physiques et morales.

40. Ces garanties devraient, au minimum, comprendre : le droit à *un recours effectif devant une autorité nationale* (y compris les autorités judiciaires) contre de telles décisions, le droit d'*être informé de l'application d'un système d'IA dans le processus décisionnel*, et le droit de *choisir d'interagir avec un humain en plus de ou plutôt qu'avec un système d'IA*, ainsi que le droit de *savoir lorsque l'on interagit avec un système d'IA plutôt qu'avec un humain*. D'autres garanties peuvent être opportunes en fonction des spécificités des systèmes d'IA utilisés. Les modalités d'exercice de ces droits devraient être prévues par le droit national. Des exceptions légitimes à ces droits peuvent être prévues par la loi, lorsqu'elles sont nécessaires et proportionnées dans une société démocratique.

41. Enfin, le CAHAI est d'avis qu'un instrument juridiquement contraignant transversal devrait également comprendre une disposition sur la *protection des lanceurs d'alerte* en lien avec le développement, la conception et l'application de systèmes d'IA susceptibles d'avoir des effets négatifs sur la jouissance des droits de l'homme, le bon fonctionnement de la démocratie et le respect de l'État de droit. Une telle disposition devrait respecter les limites juridiques légitimes en matière de divulgation.

X Éléments concernant la responsabilité civile

42. Tout en reconnaissant que les questions relatives à la responsabilité civile et au développement, à la conception et à l'application des systèmes d'IA seraient en général couvertes par le droit interne en vigueur dans les Parties à un éventuel instrument juridiquement contraignant, le CAHAI considère néanmoins qu'il est utile d'étudier la question plus en détail afin d'examiner le besoin d'assurer le partage d'une approche commune de base entre toutes les Parties en matière de responsabilité civile en relation avec l'IA.

XI Éléments concernant les autorités de contrôle, la conformité et la coopération

43. Le CAHAI considère qu'un instrument juridiquement contraignant transversal devrait comprendre des dispositions faisant obligation aux Parties de prendre toutes les mesures nécessaires et appropriées pour *garantir la conformité effective* avec l'instrument, en particulier par la *mise en place de mécanismes et de normes de conformité*. En outre, des dispositions relatives à la mise en place ou la désignation d'*autorités de contrôle nationales*, définissant leurs pouvoirs, leurs tâches et leur fonctionnement et garantissant leur expertise, leur indépendance et leur impartialité dans l'exercice de leurs fonctions, ainsi que l'allocation de ressources financières et humaines, devraient être considérées pour inclusion. De plus, l'instrument juridiquement contraignant transversal devrait comprendre des dispositions régissant la *coopération entre les Parties* ainsi que l'*entraide judiciaire et autre, notamment l'échange de données et d'autres formes d'informations*, tout en assurant la cohérence avec les autres instruments déjà applicables du Conseil de l'Europe dans le domaine de l'entraide judiciaire internationale.

44. Un instrument juridiquement contraignant transversal devrait également comprendre des dispositions sur la création d'un « *comité des Parties* » chargé d'accompagner sa mise en œuvre. À cet égard, le CAHAI renvoie aux dispositions standard utilisées dans d'autres instruments juridiquement contraignants du Conseil de l'Europe, qui pourront, si cela est nécessaire, être modifiées pour mieux répondre aux objectifs du présent instrument juridiquement contraignant.

PARTIE III : Éléments pour de potentiels instruments juridiques supplémentaires

XII Analyse d'impact sur les droits de l'homme, la démocratie et l'État de droit

45. Le CAHAI considère qu'il est utile et nécessaire de compléter un instrument juridiquement contraignant transversal par un *modèle juridiquement non contraignant pour analyser l'impact des systèmes d'IA sur la jouissance des droits de l'homme, le fonctionnement de la démocratie et le respect de l'État de droit*.

46. Une analyse d'impact sur les droits de l'homme, la démocratie et l'État de droit bien menée peut faire avancer l'analyse de la manière dont le déploiement des systèmes d'IA peut affecter la jouissance des droits de l'homme, le fonctionnement de la démocratie et le respect de l'État de droit. Il convient cependant de noter que ce type d'analyse d'impact n'est pas conçu pour équilibrer les impacts négatifs et positifs, ce qui peut dépendre des spécificités du système juridique de la juridiction dans laquelle le système d'IA est destiné à être appliqué. Dans une étape ultérieure, il peut alors être examiné si et comment les risques identifiés par le biais de l'HUDERIA peuvent être atténués, et si et comment un intérêt légitime peut légitimer l'utilisation du système malgré des interférences avec les normes relatives aux droits de l'homme, à la démocratie et à l'État de droit, lorsque de telles limitations sont prescrites par la loi, proportionnées, et nécessaires dans une société démocratique.

47. En effet, une HUDERIA ne devrait *pas être autonome*, mais être complétée, au niveau du droit interne ou international, par d'autres mécanismes de conformité, tels que la certification et la labellisation qualité, les audits, les bacs à sable réglementaires et la surveillance régulière, comme souligné dans l'étude de faisabilité. Il est important que l'analyse d'impact soit alignée avec ces autres mécanismes de conformité, car il ne serait pas justifiable en termes de coûts et de charges d'exiger des analyses d'impact sur les droits de l'homme, la démocratie et l'État de droit qui s'écartent des approches de supervision ou de réglementation publiques prévues par le droit national. En plus des mécanismes de conformité, il faut également veiller à ce que des recours efficaces restent disponibles pour ceux qui pourraient être affectés négativement par le déploiement de systèmes d'IA.

48. Compte tenu du temps et des ressources nécessaires pour entreprendre une telle analyse, et afin de préserver la proportionnalité d'une approche fondée sur les risques, le CAHAI estime qu'en règle générale, une analyse approfondie et formalisée d'impact sur les droits de l'homme, la démocratie et l'État de droit ne devrait être mandatée que s'il existe des indications claires et objectives de risques pertinents émanant de l'application d'un système d'IA. Cela nécessite que tous les systèmes d'IA fassent l'objet d'un examen initial afin de déterminer s'ils doivent ou non être soumis à une telle analyse formalisée. Il est recommandé que les indications quant à la nécessité d'une analyse plus approfondie soient développées davantage. Il convient également de considérer que l'utilisation d'un système d'IA dans un contexte nouveau ou différent ou pour un objectif nouveau ou différent, ou d'autres changements pertinents nécessiteraient une réévaluation.

49. Le CAHAI souligne que l'adoption d'une approche fondée sur les risques implique que tout impact pertinent de l'application d'un système d'IA sur la jouissance des droits de l'homme, le fonctionnement de la démocratie et le respect de l'État de droit devrait être dûment évaluée et examinée sur une base systématique et régulière en vue d'identifier des mesures d'atténuation adaptées aux risques encourus, et si de telles mesures d'atténuation ne sont pas jugées suffisantes, d'appliquer des mesures d'interdiction, si nécessaire. De plus, étant donné la nécessité d'un processus d'analyse itératif, une telle analyse devrait dans tous les cas être effectuée à nouveau chaque fois qu'un système d'IA donné subit des changements substantiels.

50. Le CAHAI recommande que, a minima, les *principales étapes* suivantes soient incluses dans une analyse d'impact sur les droits de l'homme, la démocratie et l'État de droit, sous réserve qu'un examen initial ait été effectué, et avec la participation des parties prenantes, le cas échéant :

- (1) *Identification des risques* : identification des risques pertinents pour les droits de l'homme, la démocratie et l'État de droit ;
- (2) *Analyse d'impact* : analyse de l'impact, en tenant compte de la probabilité et de la gravité des effets sur ces droits et principes ;
- (3) *Analyse de la gouvernance* : analyse des rôles et responsabilités des détenteurs d'obligations, des titulaires de droits et des parties prenantes dans la mise en œuvre et la gouvernance des mécanismes d'atténuation de l'impact ;
- (4) *Atténuation et évaluation* : identification des mesures d'atténuation appropriées et garantie d'une évaluation continue.

51. En ce qui concerne l'étape de l'*analyse d'impact*, le CAHAI recommande en outre que l'analyse d'un système d'IA pourrait comprendre, a minima, les éléments suivants : analyse du *contexte* et de la *finalité* du système d'IA, *niveau d'autonomie* du système d'IA, *technologie sous-jacente* du système d'IA, *utilisation* du système d'IA (à la fois intentionnelle et potentiellement non intentionnelle), *complexité* du système d'IA (partie de multiples réseaux de neurones profonds/s'appuyant sur d'autres systèmes d'IA), *transparence* et *explicabilité* du système et de la manière dont il est utilisé, *mécanismes de surveillance et de contrôle humains* pour le fournisseur d'IA et l'utilisateur d'IA, *qualité des données*, *robustesse/sécurité du système*, implication de *personnes ou de groupes vulnérables*, l'*échelle* à laquelle le système est utilisé, son champ d'application *géographique et temporel*, l'évaluation de la *probabilité* et l'*étendue* du préjudice potentiel, la *réversibilité* potentielle d'un tel préjudice et s'il s'agit d'une application « ligne rouge » telle qu'établie par le droit national ou international.

52. En outre, le CAHAI note que si l'analyse d'impact des systèmes d'IA est relativement simple en ce qui concerne les droits de l'homme en raison de l'existence d'obligations clairement définies et universelles dans ce domaine, l'analyse d'impact des systèmes d'IA sur la démocratie et l'État de droit pourrait s'avérer plus difficile dans certains cas. Néanmoins, étant donné le lien étroit entre les droits de l'homme d'une part et la démocratie et l'État de droit d'autre part, dans certaines situations, un impact négatif sur les premiers peut également fournir une indication d'un impact négatif sur les seconds. Par exemple, lorsque le droit à la liberté de réunion et d'association, ou le droit à des élections libres est entravé, le fonctionnement de la démocratie est entravé. De la même manière, une interférence avec le droit à un procès équitable impacte négativement l'État de droit. De plus, d'autres éléments peuvent également être pris en compte, tels que la finalité et la fonction du système au sein d'une société démocratique, son domaine d'application (avec une attention particulière à l'utilisation des systèmes d'IA dans le secteur public ou la sphère publique), et la manière dont il peut entraver certains principes démocratiques et d'État de droit (tels que le principe de légalité, la prévention de l'abus de pouvoir, ou l'impartialité et l'indépendance judiciaires).

53. Enfin, le CAHAI est d'avis que *l'implication des parties prenantes* dans l'analyse d'impact devrait être assurée. Plus l'impact est considéré comme grave, ou plus son ampleur est grande, plus l'engagement des parties prenantes doit être important. À cet égard, une attention particulière devrait être accordée à l'implication des parties prenantes externes et des membres de la société (c'est-à-dire ceux qui ne sont pas couvertes par les catégories de « fournisseurs d'IA » et « utilisateurs d'IA », cités au Chapitre III) qui pourraient potentiellement être négativement affectés par le déploiement du système d'IA.

XIII Éléments complémentaires concernant l'intelligence artificielle dans le secteur public

54. Comme indiqué au chapitre VII, le développement, la conception et l'application de systèmes d'IA dans le secteur public devraient être abordé dans un instrument juridiquement contraignant transversal, couvrant les droits et obligations transversaux les plus importants qui devraient être respectés dans ce domaine. De plus, le CAHAI est d'avis qu'en raison de la spécificité contextuelle des risques posés par l'IA dans le secteur public à la lumière du rôle spécifique de celui-ci dans la société, un tel cadre transversal pourrait être complété au niveau sectoriel par des instruments juridiquement contraignants ou non contraignants supplémentaires.

55. Ces instruments pourraient par exemple élaborer d'autres principes et exigences, spécifiques aux services publics, concernant la *transparence*, l'*équité*, la *responsabilité*, l'*obligation de rendre compte*, l'*explicitabilité* et les *recours* afin de garantir une utilisation responsable de l'IA. Le CAHAI recommande que l'utilisation, la conception, la passation de marché, le développement et le déploiement des systèmes d'IA dans le secteur public soient soumis à des mécanismes de *contrôle* adéquats afin de garantir le *respect* des droits de l'homme, des principes démocratiques et de l'État de droit, et de renforcer la *confiance du public* en rendant l'utilisation des systèmes d'IA digne de confiance, c'est-à-dire *intelligible*, *traçable* et *auditable*.

56. En outre, étant donné que la distinction entre la participation des secteurs public et privé est souvent ambiguë, et compte tenu des problèmes de responsabilité liés à la sous-traitance de services publics à des acteurs privés, toute disposition s'appliquant à la conception, au développement, et à l'application de l'IA dans le secteur public devrait également s'appliquer aux acteurs privés qui agissent pour le compte du secteur public.

57. Le CAHAI considère que les éléments suivants, relatifs à la conception, à la passation de marchés, au développement et au déploiement d'un système d'IA par une entité publique pourraient, en plus des éléments déjà décrits au chapitre VII, être traités dans le cadre d'un instrument juridiquement contraignant ou non juridiquement contraignant sur l'IA dans le secteur public.

58. Dans la *phase de conception* du système, le CAHAI est d'avis qu'un instrument juridiquement contraignant ou non juridiquement contraignant pourrait traiter de la façon dont une attention particulière pourrait être accordée à l'analyse du problème que l'entité publique entend résoudre, afin d'évaluer si un système d'IA est l'ajustement approprié au problème et, dans l'affirmative, de quelles caractéristiques il devrait être doté. Un instrument juridiquement contraignant ou non juridiquement contraignant pourrait en outre aborder les questions suivantes : les ensembles de données à utiliser pour le système d'IA devraient être clairement identifiés, et la protection de ces données ainsi que leur origine respectées. Les choix de conception du système devraient alors être rendus explicites et documentés. Les utilisateurs prévus du système, à la fois les fonctionnaires et le public, ainsi que ceux qui pourraient potentiellement être affectés par le système, devraient être impliqués dès le début, et leurs capacités à utiliser le système d'IA en question devraient être prises en compte. Une approche ouverte et transparente de la co-conception devrait être privilégiée. Enfin, une analyse d'impact sur les droits de l'homme, la démocratie et l'État de droit devrait être réalisée pour anticiper, prévenir et atténuer les risques potentiels. Cela nécessite également la mise en place de cadres de gestion et d'atténuation des risques, qui sont pertinents dans toutes les phases.

59. Au cours de la *phase de passation de marché*, un examen approfondi de la législation applicable et des mesures politiques en place devrait être effectué. Si nécessaire, des processus de passation des marchés publics devraient être adaptés et des lignes directrices sur les marchés publics pour l'IA devraient être adoptées, afin de garantir que les systèmes d'IA acquis soient conformes aux normes relatives aux droits de l'homme, de la démocratie et de l'État de droit. Une approche multidisciplinaire et multipartite devrait être assurée afin d'impliquer des perspectives et des angles différents, y compris ceux des groupes vulnérables. Étant donné que les entités publiques sont responsables des systèmes qu'elles adoptent et appliquent, une attention particulière doit être accordée à l'impact potentiel sur la responsabilité publique.

60. Pendant la phase particulièrement sensible du *développement* du système, les processus de documentation et de journalisation doivent être méticuleusement conservés afin d'assurer la transparence et la traçabilité du système. Des processus de test et de validation adéquats, ainsi que des mécanismes de gouvernance des données devraient être mis en place. Parmi d'autres risques, le risque potentiel d'accès ou de traitement inégal, les différentes formes de préjugés et de discrimination, ainsi que l'impact sur l'égalité de genre devraient être évalués.

61. Les cadres de gestion et d'atténuation des risques mis en place lors des phases précédentes devraient être évalués, adaptés et maintenus pendant la phase de *déploiement*. Compte tenu de la nature du risque, il peut être nécessaire de garantir l'implication humaine afin d'assurer une surveillance appropriée du système. Le cas échéant, le système d'IA devrait être initialement et régulièrement audité par un acteur indépendant, et les résultats rendus publics pour encourager la confiance du public. À cette fin, le CAHAI considère que l'établissement de registres publics répertoriant les systèmes d'IA utilisés dans le secteur public, contenant des informations essentielles concernant le système telles que son but, les acteurs impliqués dans son développement et son déploiement, des informations de base sur le modèle, et des indicateurs de performance le cas échéant, ainsi que le résultat de l'HUDERIA, devrait être traité dans le contexte d'un instrument juridiquement contraignant ou non juridiquement contraignant concernant l'IA dans le secteur public. En outre, l'instrument susmentionné pourrait traiter de l'établissement d'un mécanisme de retour d'information afin de recueillir des commentaires sur la manière d'améliorer le système directement auprès de ses utilisateurs et de ceux qui pourraient en être affectés. De plus, l'instrument pourrait traiter de la nécessité pour le système d'IA d'être soumis à une évaluation et à une mise à jour régulières, notamment en tenant compte des commentaires. Le processus d'évaluation pourrait être périodique. La transparence et la communication envers les utilisateurs et les citoyens devraient de même être traitées, tout comme devrait l'être la possibilité d'avoir accès à des mécanismes de responsabilité et de recours personnels et collectifs. Enfin et surtout, l'instrument devrait traiter du droit du public d'être informé du fait qu'il interagit avec un système d'IA plutôt qu'avec un être humain, ainsi que du droit d'interagir avec un être humain plutôt qu'avec un système d'IA *uniquement*, en particulier lorsque les droits et intérêts des individus et des personnes morales peuvent être négativement impactés. Des exceptions légitimes à ces droits pourraient être prévues par la loi lorsqu'elles sont nécessaires et proportionnées dans une société démocratique.

62. Enfin, un instrument juridiquement contraignant ou non juridiquement contraignant concernant l'IA dans le secteur public pourrait traiter des mesures pour accroître la culture et les compétences numériques tant des fonctionnaires que du grand public, notamment à travers des investissements dans le renforcement des capacités (formation et éducation initiale et continue) des agents publics et la sensibilisation sur les avantages, les risques, les capacités et les limites des systèmes d'IA, et en permettant la recherche d'intérêt public. Ces compétences devraient englober des connaissances théoriques et pratiques sur l'interaction entre la conception, le développement et l'application de systèmes d'IA d'une part, et les droits de l'homme, la démocratie et l'État de droit d'autre part. En outre, l'instrument susmentionné pourrait également traiter de la manière dont ces systèmes devraient être supervisés et les risques qui en découlent devraient être gérés.