



Strasbourg, 4 septembre 2020

CDPC(2020)3rév

COMITÉ EUROPÉEN POUR LES PROBLÈMES CRIMINELS (CDPC)

**ETUDE DE FAISABILITE QUANT A UN FUTUR INSTRUMENT DU CONSEIL DE L'EUROPE
SUR L'INTELLIGENCE ARTIFICIELLE ET LE DROIT PENAL**

Document préparé par le groupe de travail
sur l'intelligence artificielle et le droit pénal
avec la Dr Sabine Gless comme rapporteur général
et en coopération avec le Secrétariat du CDPC

SOMMAIRE

Introduction	3
1. Le phénomène de l'IA et ses impacts en droit pénal.....	5
1.1. Analyse des données, automatisation et capacité d'auto-apprentissage	5
1.2. Intelligibilité et marge d'erreur des systèmes pilotés par l'IA inconnus des humains	5
1.3. IA, droit pénal et droits de l'homme.....	6
1.4. Responsabilité pénale et IA : l'exemple des voitures autonomes	6
2. Réglementation existante en matière d'IA et de droit pénal	7
2.1. Législations en vigueur dans les Etats membres du Conseil de l'Europe	7
2.2. Initiatives internationales en matière d'IA et de droit pénal.....	8
2.3. Le CAHAI (Comité Ad hoc sur l'Intelligence Artificielle).....	9
3. Un instrument juridique international sur l'IA et la justice pénale	10
3.1. Analyse de la nécessité d'adopter un instrument juridique international	10
3.2. Le potentiel du Conseil de l'Europe pour permettre l'adoption d'un instrument juridique international sur l'IA et le droit pénal.....	11
4. Principaux éléments d'un instrument international du Conseil de l'Europe sur l'IA et le droit pénal	12
4.1. But, champ d'application et terminologie	12
4.1.1. But de l'instrument	12
4.1.2. Champ d'application de l'instrument	13
4.1.3. Terminologie de l'instrument	13
4.2. Droit pénal matériel : la responsabilité pénale des opérateurs et des prestataires de l'IA .	14
4.3. Droit procédural et coopération internationale : collecte des preuves dans les systèmes d'IA	15
4.4. Mesures de prévention.....	16
4.5. Mesures de protection.....	16
4.6. Mécanismes de suivi	16
Conclusion.....	17

Introduction

Le fait que les robots font désormais partie intégrante de notre quotidien soulève de nombreuses et différentes questions pratiques et juridiques, y compris dans le domaine de la responsabilité, en particulier concernant le droit pénal. Ceci est particulièrement visible avec l'automatisation de la conduite, qu'elle soit utilisée dans des voitures, des avions ou des drones, mais également pour des robots intervenant, par exemple, en matière médicale, en matière de transactions financières ou au service de personnes âgées.

Face à la perspective de la période digitale, les autorités étatiques ainsi que les organisations internationales travaillent en étroite collaboration car il n'existe pour le moment aucun cadre juridique commun ni aucune coopération internationale applicable à cette situation. Afin de pouvoir apporter les avantages promis aux citoyens, certains pays ont toutefois d'ores et déjà adopté des réglementations spécifiques s'agissant de la conduite en mode autonome, régissant les questions de l'utilisation de l'intelligence artificielle (ci-après « IA ») pour un emploi spécifique et traitant donc des questions de responsabilité pour l'utilisation de l'IA.

Cependant, des normes et des réglementations juridiques nationales différentes pourraient compromettre une avancée collective en Europe ainsi qu'une protection adéquate des intérêts individuels et le maintien de la sécurité juridique.

Le Conseil de l'Europe se trouve être dans une position privilégiée pour fournir une assistance dans l'élaboration de normes communes en raison de son expérience remarquable dans la définition de critères de référence et d'approches harmonisées, de protection des droits de l'homme, de questions de cybercriminalité et en matière d'entraide judiciaire tant au niveau collectif qu'au niveau individuel. Tout cela sur la base d'un réseau solide de coopération en matière de justice pénale de ses États membres, déterminés à prendre des mesures d'exécution efficaces tout en respectant les droits fondamentaux, notamment les droits de la défense des personnes poursuivies pour des crimes présumés et les droits des victimes.

Le Comité européen pour les problèmes criminels (ci-après « CDPC ») s'est vu confier par le Comité des Ministres la responsabilité de superviser et de coordonner les activités du Conseil de l'Europe en matière de prévention et de contrôle du crime¹. Il a fait des défis en droit pénal matériel que posent les innovations en matière de robotique, d'IA et d'appareils intelligents autonomes, susceptibles de causer des dommages physiques et matériels ou même la mort sans intervention humaine, une priorité². Le CDPC a ainsi organisé une session thématique sur l'IA et la responsabilité en matière pénale le 28 novembre 2018 à Strasbourg, avec notamment pour objectifs (i) d'examiner le champ d'application et la substance des lois pénales nationales et du droit international applicable à l'utilisation de véhicules autonomes (ou d'autres déploiements d'IA), (ii) de déterminer dans quelles circonstances certains comportements sont concernés par le droit pénal en ce qui concerne la délégation, la division ou l'attribution de tâches, de fonctions et de comportements à des technologies automatisées, ainsi que les conséquences transfrontalières éventuelles et (iii) d'examiner la portée et le contenu d'un instrument juridique international établissant des normes communes pour les aspects de droit pénal des technologies automatisées, en particulier concernant l'automatisation de la conduite.

A l'occasion de cette session thématique, le CDPC a reconnu l'importance de ce sujet et a décidé de mettre en place un groupe de travail restreint de 15 représentants des États membres,

¹ Le CDPC a pour mission d'identifier les éléments prioritaires de coopération juridique intergouvernementale, de proposer au Comité des Ministres les domaines d'action en matière de droit pénal et de procédure, criminologique et pénologique, et de conduire les activités dans ces domaines. Le CDPC élabore des conventions, des recommandations et des rapports.

² CDPC – Liste des décisions – 73^{ème} session plénière, CDPC (2017) 27, pt 3, p. 2.

soutenu par un certain nombre d'experts scientifiques (ci-après dénommé le « groupe de travail »)³. Le groupe de travail a été chargé (i) de faire le point sur les réglementations existantes, (ii) d'identifier les défis futurs liés au développement de l'IA à relever dans le domaine du droit pénal, en mettant l'accent sur la responsabilité pénale et les conditions de licence pour la commercialisation et l'utilisation de systèmes équipés de l'IA et (iii) de proposer des activités et normes possibles dans ce domaine.

Une première réunion du groupe de travail s'est déroulée à Paris le 27 mars 2019 avec pour premier objectif de préparer un questionnaire permettant de collecter auprès des États membres les informations clés au niveau national, en identifiant les éventuelles lacunes du droit pénal existant et les solutions pénales déjà en place, et avec pour second objectif d'évaluer l'opportunité d'un instrument international sur les aspects de droit pénal de l'IA. Le questionnaire a été envoyé aux États membres en mai 2019, en se basant sur l'exemple de l'automatisation de la conduite. Une évaluation des réponses au questionnaire a été publiée le 21 octobre 2019, après les réponses obtenues de 36 États membres (Andorre, Arménie, Autriche, Azerbaïdjan, Belgique, Bosnie-Herzégovine, Bulgarie, Croatie, Danemark, Estonie, Finlande, France, Allemagne, Grèce, Italie, Moldova, Monaco, Monténégro, Macédoine du Nord, Norvège, Lettonie, Lituanie, Luxembourg, Pologne, Portugal, République tchèque, Roumanie, Russie, Serbie, Slovaquie, Slovénie, Espagne, Suède, Suisse, Turquie et Ukraine).

Faisant suite à ces projets, la 77^{ème} réunion plénière du CDPC s'étant tenue à Strasbourg du 3 au 6 décembre 2019, il a été demandé au groupe de travail de « *réaliser une étude de faisabilité visant à déterminer la portée et les principaux éléments d'un futur instrument du Conseil de l'Europe, de préférence une convention, sur l'intelligence artificielle et le droit pénal* »⁴.

Le présent document vise donc à répondre aux questions suivantes : est-ce qu'un comité ad hoc d'experts du Conseil de l'Europe doit être établi afin de travailler à la rédaction d'un projet d'instrument établissant des normes pénales communes concernant les différentes questions soulevées par des véhicules se trouvant en mode autonome (ou par d'autres applications de l'IA) ? Le cas échéant, sur quels points juridiques spécifiques, identifiés comme problématiques au niveau international, ce groupe devrait-il se pencher ?

Après avoir brièvement rappelé les enjeux principaux du développement de l'IA en matière de responsabilité pénale (voir Chapitre 1) et les réglementations sur le sujet (voir Chapitre 2), seront étudiés l'opportunité d'adopter un instrument international juridiquement contraignant sur l'IA et la responsabilité pénale (voir Chapitre 3) ainsi que les points juridiques à traiter dans un tel instrument (voir Chapitre 4).

³ CDPC – Liste des décisions – 75^{ème} session plénière, CDPC (2018) 21, pt 3, p. 2.

⁴ CDPC – Liste des décisions - 77^{ème} réunion plénière, CDPC (2019) 23, pt 7, p.4.

1. Le phénomène de l'IA et ses impacts en droit pénal

1.1. Analyse des données, automatisation et capacité d'auto-apprentissage

Dans les situations où l'homme devait auparavant prendre des décisions et agir en conséquence, la technologie est aujourd'hui utilisée pour remplacer ces performances humaines par des systèmes informatiques automatisés, fonctionnant à l'aide d'algorithmes. Ces systèmes peuvent - basés sur l'apprentissage automatique - collecter et traiter de grandes quantités de données pour agir de manière autonome, et peuvent ainsi répliquer et assumer des fonctions de prise de décision précédemment exercées par l'homme. Cette nouvelle capacité à traiter une pléthore de données, basée sur des algorithmes et sur l'apprentissage automatique, permettant la détection et l'analyse de modèles pour générer automatiquement une activité autonome et éventuellement de nouvelles règles de prise de décision, non proposées par les humains, est souvent appelée Intelligence Artificielle (IA), voir également infra 4.1.3.

L'IA utilise principalement les techniques issues de la statistique, de l'informatique et de la psychologie cognitive pour reproduire des tâches habituellement attribuées aux humains. Les systèmes d'apprentissage autonome évoluent en s'adaptant aux informations transmises par leurs capteurs ou des mises à jour, les concepteurs ne déterminant et ne réglant que les paramètres initiaux et l'objectif global à atteindre de manière optimale.

Cette technologie a rapidement été appliquée dans de nombreux domaines, de plus en plus complexes, tels que la prévention des fraudes, l'analyse des risques notamment en matière judiciaire, financière et d'assurances, la collecte de renseignements sur les consommateurs, la publicité ciblée, le matériel médical ou encore la conduite automatisée sur terre ou dans les airs. La prolifération de services numériques et des systèmes d'IA est source d'avantages considérables, offrant aux hommes plus de commodité et d'efficacité dans un très large éventail de domaines et d'activités. Pourtant, dans l'état actuel des choses, l'IA n'a pas une compréhension holistique des situations et ne peut accomplir que des tâches limitées. Encore une fois, l'exemple le plus marquant de cette « IA limitée » est l'automatisation.

1.2. Intelligibilité et marge d'erreur des systèmes pilotés par l'IA inconnus des humains

Les solutions avantageuses promises par les applications de l'IA ne seront pas à l'abri d'erreurs : une voiture roulant en mode autonome peut manquer un panneau de signalisation et écraser un piéton. Un outil de surveillance des transactions financières pour la prévention de la fraude pourrait signaler un modèle de paiement qui est légal et, ce faisant, placer un être humain sous le soupçon d'actes criminels alors que tout ce que la personne a fait était de payer ses factures. Un appareil médical entraîné par l'IA pourrait confondre un cancer malin avec un début de tumeur et manquer de le signaler au médecin. Les solutions de l'IA ont leurs forces spécifiques, mais aussi leurs faiblesses qui peuvent évoluer en risques, voire causer des dommages et décès. De nos jours, on ne peut pas s'attendre à ce que les humains puissent prévoir toutes les tenants et les aboutissements possibles d'un emploi de l'IA.

La responsabilité, et en particulier la culpabilité pénale, exige cependant la prévisibilité du risque de préjudice éventuel. Le problème de la (im-)prévisibilité des résultats obtenus en utilisant des systèmes basés sur l'IA, s'est rapidement posé du point de vue d'une éventuelle absence de responsabilité ainsi, la mise en œuvre de technologies d'apprentissage profond a conduit au développement de systèmes opaques que l'homme ne parvient ni à comprendre ni à expliquer à partir des modèles initiaux et menant ainsi à la prise de décisions automatisées ou semi-automatisées qui sont difficiles à déchiffrer.

Cette problématique est d'autant plus prégnante que certains aspects importants des processus décisionnels humains ne peuvent pas être automatisés, ni assumés par l'IA. Les prises de décision humaine et des algorithmes sont fondamentalement différentes, avec des conséquences et des erreurs

distinctes. Les points faibles, les limites et les frontières des décisions prises par les algorithmes engendrent des risques inhérents à cette nouvelle technologie, tenant à la fiabilité des algorithmes en tant qu'outils, à la perception et à l'interprétation humaines de leur mise en œuvre et de leurs résultats, et l'acceptation du résultat d'une décision.

La question de la répartition de la responsabilité des risques créés par le développement des systèmes d'IA s'avère dès lors centrale, puisque la décision de savoir quels risques nous devrions prendre (par exemple sur la voie publique ou dans la prévention des fraudes) et comment attribuer les responsabilités est éventuellement un choix politique engageant toute la société.

1.3. IA, droit pénal et droits de l'homme

Avec la mise en œuvre de systèmes d'IA et la survenance des premiers accidents, il est apparu en droit pénal que, tant que l'attribution des responsabilités, que le respect de certains droits fondamentaux, notamment judiciaires, sont particulièrement concernés (ou « posent question ») car le développement de la technologie peut compromettre à court terme ou même, dans certains pays maintenant, la jouissance effective de ces droits.

En particulier, le respect d'un droit au procès équitable, protégé par les articles 5, 6 et 7 de la Convention européenne des droits de l'homme (ci-après « CEDH »), se conjugue avec celui du respect de la vie privée, protégé par l'article 8 de la CEDH, s'agissant des données collectées et utilisées par les systèmes d'IA et pouvant être utilisées dans le cadre de la mise en œuvre des poursuites pénales.

Dans les États membres du Conseil de l'Europe, le droit pénal est généralement considéré comme portant uniquement sur le comportement et les intentions de personnes humaines, soit physiques, soit agissant au nom d'entités (responsabilité des personnes morales). Du fait de leur grande complexité, les systèmes de haute technologie peuvent être facilement mal compris ou mal appréhendés par les concepteurs, les fabricants, les régulateurs et les usagers, ce qui oblige toutes les parties concernées à connaître leurs droits et devoirs respectifs. Dans ce contexte, la responsabilité pénale dérivant de situations où l'IA a causé des dommages sérieux à des êtres humains doit être clarifiée en définissant clairement les principes de l'état de droit en ce domaine et en établissant des garanties procédurales non ambiguës.

1.4. Responsabilité pénale et IA : l'exemple des voitures autonomes

L'exemple des voitures autonomes est particulièrement parlant s'agissant de la responsabilité pénale en matière d'IA ; utilisé par le groupe de travail lors de ses travaux, il sera brièvement repris ici afin d'illustrer les enjeux principaux liés à ces questions.

Parmi les nombreux accidents routiers enregistrés, certains de ces cas impliquent désormais des voitures se trouvant en mode autonome⁵, ayant même parfois mené jusqu'au décès de personnes physiques. Avec la présence d'un opérateur dans l'habitacle de la voiture roulant en mode autonome

⁵ L'industrie distingue différents niveaux d'automatisation de la conduite (Norme SAE J3016_201401). Au niveau 2, une voiture peut exécuter des tâches de conduite dynamiques, mais le conducteur doit surveiller et outrepasser le système, si nécessaire. Au niveau 3, un pilote n'a plus besoin de surveiller le système lorsqu'il est activé, mais doit répondre à une demande de prise de contrôle. L'automatisation de niveau 4 est utilisée pour divers « mélanges » de conduite hautement automatisée et entièrement automatisée en fonction de l'objectif ; elle couvre en particulier les situations dans lesquelles le conducteur ne répond pas à une demande de prise en charge et où la voiture est supposée « minimiser » le risque résultant de cette situation. Le niveau 5 envisage la conduite autonome sans conducteur humain. Dans leurs formes multiples, les assistants de conduite sont fondés sur différents concepts d'apprentissage automatique ciblant des fonctions variées (positionnement sur la voie, maintien de la distance de sécurité, stationnement, info-divertissement, détection de la somnolence, etc.). Pour fonctionner sur la voie publique, les assistants de conduite doivent être adaptatifs, autrement dit capables d'obtenir et d'évaluer des données de façon autonome en vue d'assurer une fonction spécifique. Ils intègrent donc tous les risques liés à l'IA à différentes échelles (l'autonomie, l'interconnectivité et l'interface homme-robot).

au moment de l'incident, la question se pose de savoir qui est pénalement responsable. Si traditionnellement la vitesse excessive et la consommation d'alcool ou de stupéfiants du conducteur étaient souvent des facteurs déterminants, ils n'apparaissent plus comme les seuls pertinents en la matière. Même si la technologie utilisée pour la conduite en mode autonome apparaît comme étant défectueuse, il n'y a pas de réponse simple à la question de la responsabilité pénale lorsque des systèmes d'IA sont impliqués. Pourtant, cette question est cruciale puisqu'une répartition claire de la responsabilité et sa mise en exécution par les juridictions assure une coopération fiable et pacifique au sein de la société.

Toutefois, la responsabilité à proprement parler n'est pas le seul problème. Ces voitures fortement automatisées engendrent de précieuses données lorsqu'elles roulent qui peuvent s'avérer intéressantes à bien des égards y compris pour la répression et les enquêtes pénales. Si une conduite en mode autonome s'achève par un accident, les questions se posent, en l'occurrence, de savoir si ces données peuvent servir de preuves lors d'un procès pénal, en particulier à l'encontre du conducteur humain, et de déterminer comment évaluer la crédibilité des systèmes qui produisent les données ou la fiabilité de ces données en tant que preuves. Or, les règles traditionnelles en la matière ne sont pas forcément adaptées pour évaluer adéquatement la fiabilité et la crédibilité de ces nouvelles preuves numériques.

2. Règlementation existante en matière d'IA et de droit pénal

2.1. Législations en vigueur dans les Etats membres du Conseil de l'Europe

D'après les réponses au questionnaire envoyé par le groupe de travail et complété par les États membres, seulement certains d'entre eux ont préparé voire adopté une législation générale applicable à la responsabilité pénale lorsque l'être humain cède le volant à un assistant de conduite, particulièrement sur le thème de la négligence. Un plus grand nombre d'États a adopté des règlements pour des projets pilotes de conduite automatisée, en concentrant les efforts législatifs sur l'application de normes techniques générales pour les permis spéciaux autorisant la conduite autonome, ainsi que sur une réglementation des fonctions dont les véhicules fortement automatisés devraient disposer. Les États membres qui ont choisi de réglementer uniquement l'essai de conduite prennent souvent leurs décisions au cas par cas. Dans ces conditions, il est plutôt difficile de tirer une conclusion générale qui définisse une direction commune pour cette réglementation⁶.

A titre d'illustration, l'exemple de la France sera pris pour les pays ayant adopté des dispositions sur l'utilisation de la conduite automatisée. La mise en circulation de véhicules à délégation de conduite est autorisée en France depuis l'entrée en vigueur de l'ordonnance n°2016-1057 du 3 août 2016 relative à l'expérimentation de véhicules à délégation de conduite sur les voies publiques, subordonnant toute expérimentation à la délivrance d'une autorisation des pouvoirs publics. La loi n° 2019-486 du 22 mai 2019 relative à la croissance et la transformation des entreprises (PACTE) a modifié l'ordonnance précitée, afin d'élargir le périmètre de l'expérimentation, tout en précisant son cadre juridique en prévoyant notamment le régime de responsabilité pénale applicable. Ainsi, le conducteur d'un véhicule n'est plus pénalement responsable des infractions commises pendant la conduite du véhicule lorsque le système de délégation de conduite, qu'il a activé conformément à ses conditions d'utilisation, est en fonctionnement et l'informe en temps réel être en état d'observer les conditions de circulation et d'exécuter sans délai toute manœuvre en ses lieux et place. Le conducteur redevient toutefois pénalement responsable après sollicitation du système de conduite et à l'issue d'un délai de reprise de contrôle du véhicule. De même, il redevient responsable lorsque qu'il ignore la circonstance évidente que les conditions d'utilisation du système de délégation de conduite, définies pour l'expérimentation, ne sont pas ou plus remplies. Par ailleurs, lorsque le système de délégation de conduite a été activé et fonctionne dans les conditions prévues, le titulaire de l'autorisation pour expérimenter le véhicule en

⁶ CDPC (2019) 17, p. 5.

marche autonome est pécuniairement responsable des amendes résultant des contraventions commises à l'occasion de la conduite et pénalement responsable des délits d'atteinte involontaire à la vie ou à l'intégrité de la personne si la conduite a provoqué un accident entraînant un dommage corporel, lorsqu'il est établi une faute dans la mise en œuvre du système de délégation de conduite.

Un autre exemple intéressant est l'Allemagne qui, en 2017, a adopté un amendement à sa loi sur la circulation routière, selon lequel les conducteurs sont libérés de l'obligation de surveiller et de prêter attention à la circulation lorsqu'un système de conduite automatisé sous licence est activé conformément aux conditions d'utilisation⁷. Cependant, les conséquences de cette régulation n'ont pas encore été clairement définies en ce qui concerne la responsabilité pénale d'un conducteur qui utilise une voiture avec une autonomie de niveau 3 lorsqu'un accident se produit, n'ont pas encore été clairement définies. Il est intéressant de noter que les véhicules autonomes doivent être équipés d'un dispositif d'enregistrement des données⁸.

Le cas de l'Allemagne souligne à cet égard plusieurs difficultés, parmi lesquelles notamment : Comment répartir la responsabilité entre le conducteur humain et un système d'aide à la conduite et éventuellement prouver qu'un accident a été causé par la défaillance du système et comment obtenir cette preuve des constructeurs des véhicules ? ; Ou la question de l'acceptation par la société du fait que, dans certains cas, aucune responsabilité pénale ne pourrait être attribuée, malgré le décès d'une personne causé par un véhicule roulant en mode autonome ?

2.2. Initiatives internationales en matière d'IA et de droit pénal

Pour le moment, il n'existe pas de réglementation régionale ou internationale en matière d'IA et de responsabilité pénale. Toutefois, de nombreuses organisations et institutions se penchent de manière générale sur le sujet de l'IA et abordent la question des responsabilités civiles et pénales en la matière, tandis que plusieurs universités et comités étudient en profondeur les problématiques évoquées ci-dessus.

En particulier, la Charte éthique européenne sur l'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement de 2018⁹.

S'agissant de l'Union Européenne, la Commission européenne a mis sur pied un groupe d'experts de haut niveau, qui a publié en avril 2019 des lignes directrices pour une IA digne de confiance, énumérant sept exigences essentielles, à savoir : (i) le facteur humain et le contrôle humain, (ii) la robustesse technique et la sécurité, (iii) le respect de la vie privée et la gouvernance des données, (iv) la transparence, (v) la diversité, la non-discrimination et l'équité, (vi) le bien-être sociétal et environnemental, et (vii) la responsabilisation.

Le 19 février 2020, la Commission Européenne a adopté un livre blanc intitulé : « *Intelligence artificielle- Une approche européenne axée sur l'excellence et la confiance* », prônant une approche axée sur la régulation et l'investissement, qui poursuit le double objectif de promouvoir le recours à l'IA et de tenir compte des risques associés à certaines utilisations de cette nouvelle technologie. Ce livre blanc vise

⁷ Voir §. Cf 1b de la loi allemande sur la circulation routière : (1) Le conducteur du véhicule peut détourner son attention de la circulation et du contrôle du véhicule pendant la conduite au moyen de fonctions de conduite hautement ou entièrement automatisées conformément au § 1a ; ce faisant, il doit rester prêt et alerte afin de pouvoir remplir à tout moment ses obligations au titre du paragraphe 2. (2) Le conducteur du véhicule est tenu de reprendre immédiatement le contrôle du véhicule, 1. si le système de conduite hautement ou entièrement automatisé l'y invite, ou 2. s'il reconnaît ou, en raison de circonstances évidentes, doit reconnaître que les exigences relatives à l'utilisation prévue des fonctions de conduite hautement ou entièrement automatisées ne sont plus remplies.

⁸ Voir § 63a du Code de la Route allemand

⁹ Accessible sur <<https://rm.coe.int/charte-ethique-fr-pour-publication-4-decembre-2018/16808f699b>>

à définir des options stratégiques concernant la manière d'atteindre ces objectifs, la Commission invitant notamment ses États membres et les autres institutions européennes à réagir aux options présentées et à contribuer à la prise de décision future de la Commission dans ce domaine.

A cette occasion, la Commission souligne qu'il est capital que l'IA européenne soit fondée sur les valeurs et les droits fondamentaux européens, tels que la dignité humaine et la protection de la vie privée, et retient qu'il faut adopter une approche européenne commune en matière d'IA pour parvenir à une échelle suffisante et éviter la fragmentation du marché unique, reconnaissant que des travaux importants sur l'IA sont en cours notamment au sein du Conseil de l'Europe.

S'agissant de la problématique de la responsabilité du fait des produits défectueux, qui constitue le seul élément de responsabilité étudié par la Commission Européenne, celle-ci recommande d'adapter ou de clarifier les dispositions législatives existantes dans ce domaine voir de prévoir de nouvelles mesures législatives spécifiquement consacrées à l'IA, avec des exigences obligatoires dans ses applications à haut risque, afin d'assurer un recours effectif aux personnes ayant subi un préjudice causé par les systèmes d'IA et d'assurer la sécurité juridique et la compétitivité des entreprises qui commercialisent leurs produits reposant sur l'IA dans l'Union Européenne¹⁰. Le livre blanc identifie toutefois en la matière les mêmes problématiques que celles précédemment indiquées en matière de responsabilité pénale, notamment la nécessité de définir l'IA¹¹, l'opacité des systèmes d'IA et la question de la répartition des obligations entre les opérateurs économiques impliqués.

2.3. Le CAHAI (Comité Ad hoc sur l'Intelligence Artificielle)

Conformément à son mandat, le CAHAI devra « examiner, sur la base de larges consultations multipartites, la faisabilité et 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 dans le domaine des droits de l'homme, de la démocratie et de l'État de droit [...] en coordination et consultation avec d'autres comités intergouvernementaux qui mènent des travaux sur ce sujet ».

Suite à sa première réunion (novembre 2019), le CAHAI réalise un vaste exercice de cartographie de tous les travaux et instruments applicables à l'intelligence artificielle, y compris la Convention européenne des droits de l'homme et l'impact de l'intelligence artificielle (risques et opportunités) sur les droits de l'homme, l'État de droit et la démocratie. Cette cartographie permettra d'identifier les éventuelles lacunes de ces instruments, mais aussi les principes communs et transversaux pour la conception, le développement et l'application de l'intelligence artificielle.

Le CAHAI examinera les options possibles en termes d'instruments et d'outils juridiques, mais une position claire sur la question de savoir s'il doit orienter ses travaux vers la préparation d'un instrument juridique contraignant ne pourra être prise qu'une fois que les implications auront été pleinement analysées (novembre 2020 au plus tôt).

Les États membres du Conseil de l'Europe ont des attentes claires sur une coordination des travaux du CAHAI avec d'autres organisations internationales, en particulier l'Union européenne (CE, FRA, CEPD), l'UNESCO ou l'OCDE, afin de promouvoir les synergies et d'éviter tout doublon. A cet égard,

¹⁰ Voir le rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité qui accompagne le livre blanc.

¹¹ La définition retenue par le groupe d'experts de haut niveau est la suivante : « *Les systèmes d'intelligence artificielle (IA) sont des systèmes logiciels (et éventuellement matériels) conçus par des êtres humains et qui, ayant reçu un objectif complexe, agissent dans le monde réel ou numérique en percevant leur environnement par l'acquisition de données, en interprétant les données structurées ou non structurées collectées, en appliquant un raisonnement aux connaissances, ou en traitant les informations, dérivées de ces données et en décidant de la/des meilleure(s) action(s) à prendre pour atteindre l'objectif donné. Les systèmes d'IA peuvent soit utiliser des règles symboliques, soit apprendre un modèle numérique. Ils peuvent également adapter leur comportement en analysant la manière dont l'environnement est affecté par leurs actions antérieures* ».

le Conseil de l'Europe et ces organisations participent aux comités respectifs sur l'intelligence artificielle de chacun d'entre eux.

3. Un instrument juridique international sur l'IA et la justice pénale

3.1. Analyse de la nécessité d'adopter un instrument juridique international

Il ressort de l'analyse des réponses au questionnaire que les États membres, qu'ils aient adopté ou non une réglementation spécifique en matière d'IA et de droit pénal, restent généralement ancrés dans les notions traditionnelles relatives aux régimes de responsabilité. Le groupe de travail a ainsi identifié deux tendances¹² :

- Soit le conducteur est responsable de toutes ses actions, ce qui peut mener à une responsabilité que les États membres définissent eux-mêmes comme pouvant être injuste notamment lorsque l'accident survient alors que la conduite automatique est en fonctionnement et que le conducteur a respecté l'ensemble des directives¹³ ;
- Soit le conducteur peut détourner son attention du trafic et remettre le véhicule aux assistants de conduite pourvu qu'il utilise correctement les fonctions de conduite automatique et qu'il soit prêt à répondre à tout moment à une demande de prise de contrôle. Dans ce cas, le groupe de travail a identifié un enjeu juridique d'une part lorsque, malgré le respect de toutes les consignes, il y a un accident mortel mais également, d'autre part, s'agissant du degré de négligence requis du conducteur pour engendrer sa responsabilité pénale, une responsabilité automatique en cas de non-respect des instructions données par le système d'IA par manque de temps ou défaut de connaissance technique apparaissant également comme pouvant être injuste.

Le groupe de travail a ainsi relevé que les États membres conviennent de la nécessité impérieuse d'établir une nouvelle réglementation, au plus tard lorsque « *les voitures se conduiront elles-mêmes* »¹⁴. En raison des prises de décisions autonomes, de la grande complexité des processus d'auto-apprentissage et du fait que le conducteur humain ne peut plus être considéré comme responsable de toutes les activités de conduite, le cadre législatif existant n'est plus en mesure de couvrir suffisamment ces questions.

Bien que ces questions relèvent d'éléments cruciaux de la justice pénale des États membres, matière régalienne relevant de la souveraineté de chaque État, il apparaît essentiel que l'élaboration de toute réglementation nationale s'inscrive dans un cadre international et collaboratif, et ce pour plusieurs raisons.

Premièrement, puisque les véhicules et les êtres humains traversent les frontières, l'élaboration d'un instrument juridique international sur ces questions permettra de donner des éléments législatifs de base communs aux États membres pour permettre l'élaboration de législations cohérentes sur le territoire européen. Cette nécessité de cohérence est d'autant plus nécessaire que certaines applications de l'IA, et notamment les véhicules autonomes, touchent à la mobilité concrète des citoyens sur le territoire européen : les frontières étatiques se franchissent sans obstacle et la technologie ne prend pas nécessairement en compte le changement de cadre légal et de souveraineté. Il ne s'agira pas d'élaborer tout un nouveau système de responsabilité bouleversant le droit pénal propre à chaque État membre mais bien de définir un cadre général pour le droit pénal et les applications de l'IA dans lequel inscrire le développement des réglementations étatiques.

¹² CDPC (2019) 17, p. 5.

¹³ Notamment lorsque le système d'IA présente une défaillance qui a mené à l'accident, dans la collecte de données ou dans les règles appliquées.

¹⁴ CDPC (2019) 17, p. 5.

Deuxièmement, un cadre international dans les législations nationales en matière d'IA et de droit pénal apportera une certaine sécurité juridique aux citoyens et justiciables européens ainsi qu'à l'industrie fournissant la technologie, accompagnant ainsi son développement dans le respect des droits fondamentaux au lieu de la freiner par des législations incompatibles. Un instrument commun favorise la libre circulation au-delà des frontières nationales et prévient la situation de décisions prises au cas par cas des juridictions, ce qui laisserait aux juges la tâche de trouver une cohérence parmi les cas d'espèce. L'idée selon laquelle il serait trop tôt pour régler ces questions car la technologie en est encore à ses commencements, doit être comparée avec les avantages de travailler maintenant sur un cadre international qui donnera des orientations dans le processus de développement de ces technologies. Il est important de rappeler que les questions de responsabilité en matière pénale se posent d'ores et déjà s'agissant de l'utilisation de l'IA dans les véhicules à conduite automatisée et que l'élaboration d'un instrument juridique international sur ces questions prendra au minimum plusieurs mois.

Troisièmement, un tel instrument juridique international facilitera une coopération accrue des Etats concernant les différentes problématiques liées à l'utilisation de l'IA en matière de justice pénale, une telle coopération internationale apparaissant comme étant essentielle notamment pour recueillir des preuves de l'étranger (pour les cas concernant à la fois les personnes physiques et morales) et pour le transfert de procédures pénales. Il ressort ainsi des réponses au questionnaire que tous les États membres reconnaissent la nécessité d'assurer le suivi et le stockage des données¹⁵. La plupart des autorités étatiques utilisent déjà divers outils d'analyse numérique pour assurer la sécurité sur la voie publique (radars de contrôle de la vitesse, éthylomètres numériques, lecteurs automatiques de plaques minéralogiques, tachygraphes intelligents ou éthylotests antidémarrage, positionnement GPS d'un véhicule). Sans pour autant répondre aux préoccupations qui subsistent en matière d'échange de ces données lorsque l'automatisation de la conduite devient une norme, davantage de questions non résolues se poseront concernant, par exemple, la protection de la vie privée, le respect des droits fondamentaux et de respect des droits fondamentaux dans l'utilisation de ces données comme preuves ou de la fiabilité de ses preuves.

3.2. Le potentiel du Conseil de l'Europe pour permettre l'adoption d'un instrument juridique international sur l'IA et le droit pénal

Le Conseil de l'Europe a été créé afin de garantir la démocratie fondée sur la liberté des personnes et de prévenir de nouvelles violations massives des droits de l'homme comparables à celles commises pendant la seconde guerre mondiale. Il s'intéresse à tous les grands problèmes auxquels les pays européens font face, hormis ceux liés à la défense militaire. Sa mission consiste à promouvoir la démocratie, les droits de l'homme et l'état de droit, ainsi qu'à élaborer des réponses communes aux défis politiques, sociaux, culturels et juridiques de ses Etats membres. Il est impératif et urgent de veiller à la primauté du droit, notamment dans le cadre du développement des systèmes d'IA. De par le travail qu'il a déjà accompli dans ce domaine, le Conseil de l'Europe fait figure de pionnier intergouvernemental sur les questions d'IA et de droit pénal.

La préparation et l'adoption d'un instrument juridique spécifiquement dédié aux enjeux liés à l'IA dans la justice pénale permettra d'appréhender efficacement et de manière cohérente les infractions présumées en la matière et les problèmes qui y sont liés (qui apparaissent au compte-gouttes notamment en Europe), en permettant l'élaboration de législations nationales comblant les vides juridiques constatés par les experts selon des standards internationaux communs. Cela permettra également d'assurer et de promouvoir une coopération internationale accrue sur ces nouveaux sujets, en envoyant un signal fort fournisseurs de systèmes pilotés par IA quant à la nécessité de développer cette technologie dans le respect des droits fondamentaux protégés au niveau international et en

¹⁵ CDPC (2019) 17, p. 6.

collaboration avec eux. Une convention internationale pourrait être l'un des outils sur lequel débiter les négociations, pour souligner l'urgence de la situation.

En raison du grand nombre de ses Etats membres (47 pays), le Conseil de l'Europe dispose d'une influence telle que les instruments qu'il élabore ont un poids considérable. En adoptant un instrument sur l'IA et le droit pénal s'inspirant de son approche des droits de l'homme, le Conseil de l'Europe la ferait reposer sur la CEDH, qui confère à chaque partie l'obligation positive de protéger ses citoyens contre les violations des droits de l'homme. Par ailleurs, l'adoption d'une convention du Conseil de l'Europe sur l'IA et le droit pénal, en tenant compte des conventions internationales existantes (telles que la Convention de Genève sur la circulation routière de 1949 ou la Convention de Vienne sur la circulation routière de 1968¹⁶) et en mettant à profit les travaux en cours au sein des autres institutions internationales, pourrait avoir un effet d'entraînement au niveau de l'Union européenne et pourrait inspirer d'autres instruments internationaux.

4. Principaux éléments d'un instrument international du Conseil de l'Europe sur l'IA et le droit pénal

Reprenant la structure traditionnelle des instruments du Conseil de l'Europe, les questions suivantes, identifiées comme importantes à traiter pour les Etats membres du Conseil de l'Europe mais également des Etats tiers, pourraient cristalliser les futures négociations en vue de l'adoption d'un instrument international en matière d'IA et de droit pénal.

4.1. But, champ d'application et terminologie

4.1.1. But de l'instrument

Tel qu'exposé ci-dessus et au regard de l'analyse des réponses des Etats membres au questionnaire portant sur l'IA et le droit pénal (s'appuyant sur l'exemple de la conduite automatisée), quatre objectifs ont été identifiés, à savoir :

1. Etablir un cadre international pour le développement des législations nationales s'agissant des problématiques du droit pénal en matière d'IA (plus particulièrement concernant la responsabilité pénale dans le cadre de la conduite automatisée) ;
2. Inciter les Etats membres à prendre en compte l'enjeu juridique constaté en matière de droit pénal et d'IA, en traitant la question au niveau législatif, à l'aide de principes normatifs communs ;
3. Anticiper les problématiques probatoires et juridiques d'ores et déjà identifiées en matière de responsabilité pénale et d'IA et assurer les principes du procès équitable ainsi qu'une coopération internationale efficace en la matière ;
4. Assurer le développement des systèmes d'IA dans le respect des droits fondamentaux protégés par les instruments du Conseil de l'Europe.

Ces objectifs généraux pourront être discutés, modifiés et complétés par les Etats membres.

¹⁶ Les États membres de la Convention de Vienne ont voté un amendement qui ajoute deux définitions (d'un système de conduite automatisé et du contrôle dynamique du véhicule) ainsi qu'un article qui, en substance, précise qu'une obligation pour tout véhicule en mouvement d'avoir un conducteur est considérée comme satisfaite lorsque le véhicule utilise un système de conduite automatisé conforme aux réglementations techniques nationales et internationales et que les réglementations nationales régissent son fonctionnement / sa circulation sur la route.

4.1.2. Champ d'application de l'instrument

Il s'agit ici de définir la portée exacte d'un futur instrument du Conseil de l'Europe sur l'IA et le droit pénal et de ne pas déborder sur les champs de compétence d'autres organes qui travaillent de leur côté sur d'autres problématiques connexes.

Ainsi et par souci de clarification, les questions de responsabilité civile et administrative résultant de l'usage de l'IA ne seront pas abordées dans le cadre de l'élaboration d'un instrument juridique du Conseil de l'Europe sur l'IA et le droit pénal, ces questions impliquant des enjeux distincts et relevant d'un cadre légal différent. De même, l'utilisation de l'IA par les forces armées des Etats membres ne saurait relever de ce futur instrument, seule l'utilisation à des fins civiles posant la question de la mise en jeu de la responsabilité pénale des citoyens.

Par ailleurs, il ressort des réponses au questionnaire qu'aucun Etat membre n'envisage pour le moment la création d'une personnalité juridique des robots dotés de l'IA en matière pénale¹⁷, la responsabilité pénale étant de façon unanime basée sur une intention ou une négligence coupable ne pouvant être rattachée *in fine* qu'à une personne physique, et cette question sera donc exclue des discussions, les recherches et les débats étant en cours à cet égard.

En outre, le Bureau du CDPC a décidé lors de sa réunion du 10/11 octobre 2019 « de charger le PC-CP de produire une étude pour la réunion plénière du CDPC en 2020, portant sur l'utilité de rédiger un texte normatif visant à établir le cadre nécessaire à l'utilisation de plus en plus intensive de l'intelligence artificielle par les services pénitentiaires et de probation en Europe »¹⁸. L'utilisation de l'IA par le système judiciaire lui-même (évaluation des risques, profilage, police prédictive, identification faciale, prisons et services de probation etc.) sera donc exclue des discussions, d'autres organes étant chargés de conduire une étude indépendante sur ce point.

Enfin et comme lors de l'élaboration et de la diffusion du questionnaire destiné aux Etats membres en matière d'IA et de droit pénal¹⁹, les questions relatives à la cybersécurité seront également exclues de l'analyse, pour éviter tout chevauchement avec les autres activités du Conseil de l'Europe et notamment du T-CY.

4.1.3. Terminologie de l'instrument

Il n'existe pas, à ce jour, de définition universelle du terme « Intelligence Artificielle ». Lors des différents travaux précédents du CDPC, les participants ont utilisé la définition suivante de l'IA : une combinaison de disciplines – dont la logique mathématique, la statistique, le calcul de probabilités, la neurobiologie informatique et l'informatique – visant à permettre à la machine d'imiter, voire de surpasser, les facultés cognitives de l'être humain.

L'établissement d'une définition de travail commune de l'IA s'avère être une prémisse indispensable à toute discussion et élaboration de normes communes à cet égard en matière pénale. Il s'agira donc d'un point nécessaire de discussion dans le cadre de l'établissement d'un instrument juridique international en matière d'IA et de droit pénal, de même que les définitions corrélatives des termes « Robot »²⁰ et « Preuve électronique »²¹. Toutefois, il suffira de s'entendre sur des définitions de travail

¹⁷ CDPC (2019)17, p. 5.

¹⁸ CDPC-BU (2019) 4, pt c, p. 3.

¹⁹ CDPC (2019) 11, p. 2.

²⁰ Le groupe de travail a utilisé pour ses travaux la définition suivante : dispositif doté d'une intelligence artificielle intégrée dans une structure matérielle, capable d'accomplir des actions ayant des incidences sur le monde réel. Le « bot » était quant à lui défini comme suit : logiciel autonome capable d'interagir avec d'autres programmes ou avec un utilisateur humain.

²¹ Le groupe de travail a utilisé pour ses travaux la définition suivante : donnée générée automatiquement lors d'une coopération homme-machine fondée sur l'IA, présentée comme preuve dans une enquête pénale.

dans le cadre de l'instrument, c'est-à-dire que pour un instrument sur l'automatisation de la conduite, il est possible de s'appuyer sur des normes techniques (par exemple les normes ISO).

4.2. Droit pénal matériel : la responsabilité pénale des opérateurs et des prestataires de l'IA

Un dénominateur commun européen pour le droit pénal matériel est l'enjeu principal de l'instrument juridique du Conseil de l'Europe sur l'IA et le droit pénal proposé, à savoir l'établissement d'un cadre international commun pour les règles matérielles nationales en matière de droit pénal et d'IA. Deux problématiques essentielles, parmi plusieurs autres, exposées ci-dessous pourraient faire l'objet de discussions entre les Etats membres, les questions centrales et sous-jacentes étant celles de l'approche de la responsabilité concernant le risque possible émanant de l'utilisation de l'IA, et celles autour du fait que le concept actuel de la négligence soit suffisamment adapté pour couvrir tous les comportements fautifs ou bien qu'une nouvelle législation spécifique soit nécessaire.

Tout d'abord, en ce qui concerne les avantages escomptés de l'emploi de l'intelligence artificielle et les caractéristiques de l'interaction homme-robot, les États membres pourraient examiner s'ils veulent se mettre d'accord sur un point de référence, une forme spécifique de négligence et / ou l'étendue du préjudice causé ou un niveau d'automatisation mis en place comme déclencheur d'une enquête pénale. Si un instrument opte pour une procédure pénale spécifique, des cas typiques doivent être pris en compte, où les producteurs, les utilisateurs ou d'autres personnes ne prennent pas les mesures nécessaires pour contrôler les risques émanant des robots, comme les défauts de conception ou de formation de l'IA ou lorsque les utilisateurs ne tiennent pas compte des instructions d'un système, par exemple en ne reprenant pas le contrôle de la machine (délai raisonnable de reprise en main de la machine) ou, dans le cadre des véhicules pilotés par l'IA, en refusant de s'arrêter malgré la détection par le système de signes de fatigue et la suggestion d'une pause. Il faut peut-être définir de nouvelles formes d'infractions, comme la vitesse excessive par la formation défectueuse d'un assistant de vitesse (qui, sans raison plausible, accélère dans une zone résidentielle) ou une ingérence dangereuse dans le trafic routier par le piratage.

Ensuite, l'existence et le développement de l'IA pose la question si une nouvelle approche de la responsabilité pénale est nécessaire, lorsque l'infraction est commise alors que le robot agit de façon complètement autonome et/ou que l'utilisateur a respecté l'ensemble des directives, mais que les applications spécifiques de l'IA ont causé un dommage. La prise en compte de l'IA comme contrepartie agissante lors de l'attribution de la responsabilité pénale, à première vue, est contraire aux principes de base du droit pénal, qui a été conçu par rapport à l'action humaine. Cependant, si le conducteur humain a complètement disparu du siège du conducteur, le passager actuel ne peut plus être tenu entièrement responsable des accidents causés par la voiture. Certains veulent établir un parallèle avec la responsabilité des entreprises. Cependant, le parallèle n'est pas évident, car les sociétés sont des personnes morales dans tous les États membres, et donc responsables en vertu de la loi, mais elles ne peuvent pas comparaître en tant que telles devant une juridiction pénale pendant un procès. En outre, dans la plupart des pays, la responsabilité pénale des personnes morales est liée aux actes répréhensibles d'un être humain les représentant. Il semble toutefois évident que la question doit être abordée et les points à discuter sont les suivants : un producteur de « produits intelligents » (smart products) ou un prestataire de services d'IA pourrait-il être tenu responsable, par principe, et si oui quel niveau de négligence de sa part et/ou quel degré de préjudice seraient exigés pour entraîner sa responsabilité pénale. A titre d'exemple, il s'agirait des cas de responsabilité pénale résultant d'une alimentation incorrecte des données cartographiques ou des informations sur les capteurs ou de la création d'un système d'IA dangereux. Dans ce contexte, les Etats membres pourraient également s'accorder sur un cadre normatif relatif à la multiplicité des responsabilités pénales que pourraient engendrer une conjonction de négligences fautives ayant causé l'infraction. Ce cadre permettrait ensuite la réforme au niveau national des infractions routières qui se concentrent actuellement sur l'action humaine.

Enfin, les Etats membres pourraient appréhender la nature et la mesure des sanctions pénales les plus appropriées en cas de responsabilité pénale impliquant un système d'IA, dans un souci de cohérence législative au niveau européen.

4.3. Droit procédural et coopération internationale : collecte des preuves dans les systèmes d'IA

Résoudre le problème de la responsabilité pénale n'a de sens que si les problématiques de procédure intrinsèquement liées sont également résolues. Cela inclut - entre autres - la question de l'utilisation des données générées par l'interaction homme-robot comme preuve.

Au regard des récents développements et applications de l'IA, les éléments de preuve pertinents en matière pénale proviendront très certainement de la machine, sous forme de données générées par le robot prenant souvent part à l'interaction humain-robot. Cette situation peut poser plusieurs questions :

Tout d'abord, la disponibilité des données pourrait être un problème. Sans réglementation supplémentaire, ce serait des produits de consommation, tels que les voitures, qui génèrent des données qui seraient nécessaires comme preuve dans un procès pénal. Ces données pourraient être stockées chez les fabricants de robots compatibles avec l'IA ou auprès de prestataires ou de sociétés de services cloud. Il s'agira alors de savoir comment obtenir l'accès à l'intérieur avec un besoin de dispositions procédurales nationales (harmonisées) et ne pas avoir accès à l'étranger avec des instruments spécifiques d'entraide judiciaire. Dans ce contexte, la cohérence avec les règles contenues dans d'autres instruments internationaux traitant de cette question devrait être assurée, en particulier avec le deuxième protocole additionnel à la Convention de Budapest, qui est actuellement en cours de négociation au sein du Conseil de l'Europe (T-CY). En ce qui concerne la coopération internationale, les problèmes classiques liés à la territorialité peuvent être particulièrement inquiétants, tout comme les lois nationales protégeant le secret commercial.

Une solution pourrait être la nécessité d'installer une boîte d'enregistrement des données. L'obligation d'incorporer un appareil d'enregistrement (parfois appelé "boîte noire", à ne pas confondre avec le problème d'opacité dans les robots également appelé problème de boîte noire) lors de l'utilisation de l'IA doit être coordonnée entre les États. Il sera nécessaire de prévoir une coopération entre les États membres pour obtenir des données de la "boîte noire" lorsque celle-ci ne se trouve plus sur le territoire où l'infraction a été commise, d'établir une obligation pour le fabricant de divulguer les codes et les informations nécessaires, le paramétrage d'un système d'apprentissage machine, les données d'entraînement, etc. et la nécessité d'un contrôle en la matière par des autorités judiciaires.

Deuxièmement, la question de la fiabilité et de sa vérification au sein des procédures pénales et autres se pose, si les produits de consommation, comme les voitures, génèrent des données qui doivent être utilisées comme preuves dans un procès pénal. Il pourrait être nécessaire non seulement de réglementer la collecte, le stockage, le chiffrement et l'accès aux données, mais également d'envisager la portée du privilège du secret commercial et d'autres protections commerciales. Les règles traditionnelles peuvent ne pas être conçues pour offrir des moyens adéquats pour tester de manière significative la fiabilité et la crédibilité de ces nouvelles preuves numériques ni une protection suffisante des intérêts commerciaux ni les instruments nécessaires pour s'attaquer à de nouveaux problèmes, par exemple le problème de la boîte noire de l'IA.

Troisièmement, la question liée au respect des droits fondamentaux en la matière (droit à la vie privée et respect des droits de la défense, notamment du droit au silence et du droit d'interroger les témoins soi-même conféré par l'article 6 de la CEDH) doit être traité dans le développement d'un instrument juridique sur l'IA.

4.4. Mesures de prévention

Afin de minimiser les risques liés à une utilisation de plus en plus massive des systèmes d'IA, aussi bien par rapport à la responsabilité pénale qu'aux questions d'utilisation des preuves, le besoin d'obligations générales de transparence quant aux systèmes déployés et d'information sur leur fonctionnement à la charge des entreprises privées impliquées dans le développement et la mise en circulation des systèmes d'IA pourraient être discutées entre les Etats membres. De même, la mise en place d'une obligation générale de formation des utilisateurs des systèmes d'IA pourrait faire l'objet de négociations entre les Etats membres, par exemple à la charge des écoles de conduite en matière de véhicules autonomes dans le cadre de l'obtention du permis de conduire.

Par ailleurs, les Etats membres peuvent envisager de (i) recenser les cas de responsabilité pénale devant ses juridictions nationales impliquant les systèmes d'IA, la façon dont la preuve a été obtenue dans ces cas et les standards légaux selon lesquels la décision a finalement été prise, (ii) recenser toute nouvelle norme nationale en matière d'IA et de responsabilité pénale et (iii) s'assurer de la conformité de ces normes nationales avec les dispositions du futur instrument du Conseil de l'Europe. Cet organe indépendant national serait ainsi en mesure d'alerter l'Etat membre concerné de toute problématique en matière de responsabilité pénale et d'IA nécessitant l'adaptation voire la modification du cadre normatif national ou international.

4.5. Mesures de protection

La mise en place d'un mécanisme de permis national ou régional pour les développeurs de systèmes d'IA pourrait être encouragée par les Etats membres, permettant la suspension d'un tel permis en cas d'identification d'un risque criminel majeur, notamment par l'organe national indépendant susmentionné, au lieu et place d'une interdiction pure et simple de l'utilisation de systèmes d'IA sur le territoire national, annihilant le développement de la technologie.

4.6. Mécanismes de suivi

Il convient enfin d'aborder la question du suivi de la mise en œuvre du futur instrument du Conseil de l'Europe. Celle-ci tirerait clairement profit d'un mécanisme de suivi, s'assurant de sa mise en œuvre.

En dépit du fait que la plupart des comités de suivi ne peuvent produire que des avis non contraignants, contrairement à la Cour européenne des Droits de l'Homme, ils jouent néanmoins un rôle crucial dans la production d'un recueil des meilleures pratiques qui pourraient ultérieurement servir de modèle à d'autres. Ces avis de haut niveau pourraient également être utilisés par d'autres organes du Conseil de l'Europe, tels que la Cour européenne des Droits de l'Homme, dans leur jurisprudence à venir.

Conclusion

D'après le plan et la marche à suivre établis par le CDPC²², il ressort du premier produit réalisé, à savoir le questionnaire réalisé dans le cadre du projet de recherche sur le droit pénal national et le cadre juridique international concernant les véhicules autonomes (ou les autres applications de l'IA), qu'il est hautement souhaitable et souhaité par les Etats membres qu'un instrument juridique international soit négocié en matière d'IA et de droit pénal, conformément aux enjeux et problématiques développés ci-dessus, et ce afin d'établir un cadre international permettant le développement de législations nationales spécifiques. En effet, se mettre d'accord sur des normes communes pour répartir clairement et correctement la responsabilité pénale éventuelle et clarifier les questions de procédure connexes ainsi que les implications possibles pour les droits de l'homme doit relever d'un effort combiné entre la sphère publique et les acteurs privés, encourageant le développement de la technologie dans de bonnes conditions et dans le respect des principes fondateurs de la société civile. Il appartient aux Etats membres d'élaborer des mécanismes efficaces garantissant la responsabilité algorithmique, en travaillant en collaboration avec ceux qui développent et exercent le pouvoir numérique.

Le CDPC pourrait donc s'engager dans la mise en œuvre des prochaines étapes du plan (« [Document de réflexion](#) »), à savoir le produit n°2 consistant dans l'organisation d'une conférence internationale sur des normes pénales communes concernant les dommages causés par des véhicules autonomes (ou par d'autres applications de l'IA), afin d'offrir aux États membres et aux acteurs publics et privés un forum où discuter des évolutions dans le domaine de l'IA, des lacunes du droit pénal existant en la matière et des solutions juridiques d'ores et déjà mises en place ou à établir via un instrument international ainsi que le produit n°3, consistant en la création d'un groupe *ad-hoc* d'experts, qui travailleraient sur un instrument, dont la forme et les contours demeurent à déterminer, établissant des normes communes internationales, en premier lieu dans le domaine pénal concernant les dommages causés par des véhicules autonomes (ou par d'autres applications de l'IA).

²² CDPC (2018) 14rév, p. 9.