Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Centre for Information Policy Leadership

Dix Recommandations pour une réglementation mondiale de l'IA

Octobre 2023.

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Dix recommandations pour une réglementation mondiale de l'IA

Sommaire

INTRODUCTION 3

I. DES RÈGLES FONDÉES SUR DES PRINCIPES ET DE RÉSULTATS À ATTEINDRE 5

1. Créer un cadre flexible et adaptable qui définit les résultats à atteindre, plutôt que de prescrire en détail la manière de les atteindre 5

2. Adopter une approche fondée sur les risques considèrant risques et bénéfices de manière holistique 6

3. S'appuyer sur les fondamentauxde la législation contraignante et non contraignante 7

4. Responsabiliser les individus grâce à la transparence, ' l' explicabilité

et aux mécanismes de recours 9

II RESPONSABILITÉ ORGANISATIONNELLE A DÉMONTRER 10

5. Faire de la responsabilité organisationnelle (accountability) et de sa démontration un élément central de la réglementation en matière d'IA 10

6. Faire progresser l'adoption de pratiques de gouvernance de l'IA responsables 11

7. Répartir soigneusement les responsabilités, en mettant l'accent sur la partie la plus étroitement associée à l'origine du préjudice 11

III. UNE SURVEILLANCE RÉGLEMENTAIRE PARTENARIALE ET EMPIRIQUE 12

8. Créer des mécanismes de coordination et de coopération entre les organismes de réglementation 12

9. Instaurer une surveillance réglementaire fondée sur la coopération et permettre une innovation réglementaire permanente 13

10. S'efforcer de parvenir à une interopérabilité mondiale 15

ANNEXE I – CADRE DE RESPONSABILITÉ DU CIPL 16

ANNEXE II - MISE EN CORRESPONDANCE DES MEILLEURES PRATIQUES EN MATIERE DE GOUVERNANCE DE L'IA AVEC LE CADRE DE RESPONSABILISATION DU CIPL 17

INTRODUCTION

L'intelligence artificielle (IA)[i] génère des avantages sociétaux importants et croissants, notamment en alimentant la recherche médicale, en luttant contre le changement climatique, en transformant les industries et en modernisant les gouvernements. Dans le même temps, le déploiement et l'adoption rapides de nouvelles applications, telles que les chatbots d'IA générative et les générateurs d'images, ont intensifié des préoccupations de longue date et soulevé de nouvelles questions liées à la vie privée et à la protection des données, à la transparence et à l'explicabilité, aux droits de l'homme, à la propriété intellectuelle, à la sécurité, à la partialité, à l'impact sur la main-d'œuvre, à la production et à la diffusion de fausses informations et de désinformations, ainsi qu'à d'autres effets sociétaux. En réponse, les organisations développent des contrôles opérationnels et des cadres de gouvernance pour garantir un développement et un déploiement responsables de l'IA ; les experts du secteur travaillent à l'élaboration de normes ; les décideurs politiques rédigent de nouvelles lois ; et les régulateurs testent les limites des pouvoirs existants et en proposent de nouveaux[ii]. Cependant, il n'y a pas de consensus entre les pays sur la meilleure approche à adopter pour réglementer l'IA : faut-il privilégier une réglementation stricte et spécifique, des modèles de corégulation, des certifications et des assurances, des normes sectorielles, ou tout à la fois[iii] ?

Depuis plus de 20 ans, le CIPL est un leader d'opinion en matière de responsabilité organisationnelle et d'approche fondée sur le risque de la politique et des pratiques en matière de données. Il a contribué très tôt à l'évaluation des défis et à la définition de solutions pour la gouvernance de l'IA et les pratiques de l'industrie. Parmi les principales contributions du CIPL dans ce domaine, citons *Artificial Intelligence and Data Protection in Tension* (*Des tensions entre l'Intelligence Artificielle et la protection des données)* (octobre 2018), *Hard Issues and Practical Solutions* (*Problèmes essentiels et solutions pratiques)* (février 2020) et *Artificial Intelligence and Data Protection: How the GDPR Regulates AI* (*L'Intelligence Artificielle et la protection des données* : *comment le GDPR réglemente l'IA)* (mars 2020)[iv]. Le CIPL a également préparé des réponses détaillées aux consultations publiques sur la politique en matière d'IA au Brésil, dans l'Union européenne, au Royaume-Uni et aux États-Unis[v].

S'inspirant de cette expérience et de notre engagement auprès des leaders du secteur privé qui développent et déploient des technologies d'IA, des décideurs politiques et des régulateurs, le CIPL propose dans ce document, dix recommandations pour guider l'élaboration de politiques et de réglementations en matière d'IA afin de permettre une IA responsable et digne de confiance. Ces dix recommandations résument le point de vue du CIPL sur une approche progressive ou en trois étapes de la réglementation de l'IA :

a) des règles fondées sur des principes et les résultats à attenindre,
b) une responsabilité organisationnelle à démontrer, et
c) une surveillance réglementaire appuyée et intelligente.

Une telle approche fournit des règles à l'épreuve du temps, fondées sur des principes fondamentaux, qui peuvent guider le développement et le déploiement éthiques de l'IA, même si la technologie et les cas d'utilisation évoluent. Nous décrivons cette approche ci-dessous.

Recommandations pour la réglementation de l'IA

Le CIPL recommande une approche de la réglementation de l'IA fondée sur le risque et par étapes, qui s'appuie sur les lois et les normes existantes et sur les pratiques responsables des organisations. Cette approche devrait être soutenue par une surveillance réglementaire innovante et des instruments de corégulation.

C'est pourquoi il nous semble que toute approche législative ou réglementaire de l'IA devrait suivre ces recommandations générales :

A. Règles fondées sur les principes et les résultats
   1. Créer un cadre flexible et adaptable qui définit les résultats à atteindre, plutôt que de prescrire des détails sur la manière de les atteindre
   2. Adopter une approche fondée sur les risques qui considère les risques mais aussi les bénéfices de manière holistique
   3. S'appuyer sur les fondamentaux existants de la législation contraignante et non contraignante
   4. Responsabiliser les individus grâce à la transparence, au droit de demander des explications ciblées et aux mécanismes de réparation

B. Responsabilité organisationnelle (accountability) démontrable
   5. Faire de la responsabilité organisationnelle et de sa démonstration un élément central de la réglementation en matière d'IA
   6. Faire progresser l'adoption de pratiques de gouvernance de l'IA responsables
   7. Répartir soigneusement les responsabilités, e et concentrer sur la partie a plus étroitement associée à l'origine du préjudice

C. Une surveillance réglementaire partenariale et empirique
   8. Créer des mécanismes de coordination et de coopération entre les organismes de réglementation
   9. Instaurer une surveillance réglementaire fondée sur la coopération et permettre une innovation réglementaire permanente
   10. S'efforcer de parvenir à une interopérabilité mondiale

I. DES RÈGLES FONDÉES SUR DES PRINCIPES ET DES RÉSULTATS À ATTEINDRE

1. Créer un cadre flexible et adaptable qui définit les résultats à atteindre, plutôt que de prescrire en détail la manière de les atteindre

Pour être efficaces, les réglementations sur l'IA doivent pouvoir rester pertinentes à mesure que la technologie et les cas d'utilisation continuent de progresser. Toute règle doit être *neutre sur le plan technologique* : un cadre trop normatif et spécifique aux technologies individuelles ou aux modèles et pratiques d'entreprise actuels risque d'être rapidement dépassé et d'entraver les innovations et les bénéfices que celles-ci sont susceptible d'engendrer. En effet, une approche basée sur des listes de technologies spécifiques nécessitera des modifications fréquentes pour suivre l'évolution technologique. Si les règles incluent des listes de technologies et d'applications présumées à haut risque, elles devraient permettre à ces présomptions d'être réfutées et d'évoluer dans le temps.

Les règles doivent également être *fondées sur des principes et des résultats* à atteindre . Elles devraient permettre aux organisations de garantir les résultats requis (par exemple, l'équité, l'impartialité, la transparence, l'exactitude, la sécurité, la surveillance humaine) grâce à des politiques, des procédures et des contrôles internes vérifiables et fondés sur les risques, adaptés à leur contexte spécifique, sans prescrire la manière d'atteindre ces résultats. Une telle approche offre aux développeurs la souplesse nécessaire pour innover, y compris en ce qui concerne les contrôles, les outils techniques et les garanties, tout en maintenant une cohérence avec les principes fondamentaux applicables et les résultats -[vi].

De même, un cadre réglementaire ne devrait pas être trop prescriptif quant aux méthodologies d'évaluation de l'impact et des risques de l'IA, mais décrire les critères à prendre en compte lors de l'évaluation des risques et des avantages d'une application de l'IA et laisser aux régulateurs compétents le soin de fournir des orientations plus adaptées, réalistes et pratiques en collaboration avec ceux qui développent et déploient les technologies de l'IA.

Dans le même temps, les règles devraient offrir autant de certitude que possible quant à leur champ d'application. Par exemple, un cadre réglementaire pour l'IA doit définir l'IA afin que les parties prenantes puissent clairement comprendre quels systèmes entrent dans le champ de ces règles. En l'absence d'une telle clarté et d'une focalisation sur les résultats, l'ambiguïté réglementaire et la surprescription réglementaire risquent de freiner l'investissement et l'innovation, en particulier pour les petites et moyennes entreprises (PME) et les start-ups, qui sont de puissants moteurs de l'innovation et de l'investissement en matière d'IA.

2. Adopter une approche fondée sur les risques qui considère risques et bénéfices de manière holistique

Toute approche réglementaire de l'IA devrait viser à protéger les droits fondamentaux de l'homme et à minimiser les risques pour les individus et la société, tout en permettant le développement et l'utilisation de l'IA dans l'intérêt collectif. En effet, les risques pour les individus et la société peuvent également se matérialiser par la *non-utilisation* de technologies d'IA efficaces, telles que celles qui ont la capacité de prédire et de prévenir les maladies, ou de réduire les préjudices en ligne, les menaces de cybersécurité et la fraude. Une approche holistique fondée sur les risques favorise cet objectif en facilitant l'adoption de mesures de protection pratiques proportionnelles aux risques et aux bénéfices d'un système d'IA particulier. Elle se concentre sur les impacts potentiels de la technologie de l'IA dans le contexte de cas d'utilisation spécifiques.

Un cadre réglementaire de l'IA reposant sur l'évaluation des risques fournirait des critères non exhaustifs pour aider les organisations à déterminer la probabilité et la gravité de tout dommage qui en résulterait, et les mesures nécessaires pour l'atténuer. L'évaluation et la compréhension de l'impact potentiel de leurs systèmes d'IA permettent aux organisations d'adapter leurs mesures d'atténuation des risques réels et d'éviter la mise en œuvre de mesures inutiles. Par exemple, **l'algorithme "k plus proches voisins » (KNN** ou **k-NN** pour « K-nearest neighbors »**)** est un algorithme d'apprentissage automatique utilisé dans diverses applications[vii] : dans le commerce de détail pour recommander des produits, dans les soins de santé pour prédire le risque de crise cardiaque et de cancer de la prostate, dans la finance pour détecter les activités frauduleuses, dans l'agriculture pour prédire le rendement des récoltes et dans les transports pour prédire et optimiser la circulation. Ces différentes utilisations du même algorithme KNN présentent des niveaux de risque différents - la probabilité et la gravité d'un préjudice dans la recommandation de chansons ou de vêtements diffèrent de celles de la médecine d'urgence.

Un cadre fondé sur les risques doit également évaluer les avantages potentiels d'un système d'IA pour les individus, les organisations et la société. Ceux-ci peuvent ensuite être mis en balance avec les risques identifiés liés au déploiement (ou au non-déploiement) de l'IA. Par exemple, les risques liés aux véhicules autonomes (VA) dépendent des différents environnements dans lesquels ils sont déployés. On peut dire que le risque d'atteinte aux personnes est plus faible lorsque les véhicules autonomes sont déployés dans les mines et les exploitations agricoles que dans les zones urbaines ou résidentielles. Dans le même temps, les véhicules autonomes dans les premiers contextes offrent différents avantages, tels que l'allègement du besoin de main-d'œuvre, le soutien à l'agriculture durable et l'amélioration de la productivité[viii]. Le même exemple est utile pour souligner l'importance de bien peser les paramètres permettant de mesurer le risque : on pourrait évaluer le risque des VA par rapport au statu quo (un monde où la plupart des voitures sont conduites par des personnes) et/ou par rapport à une solution optimale désignée pour les VA.

En résumé, les résultats des évaluations holistiques des risques liés à l'utilisation d'une technologie d'IA spécifique peuvent varier considérablement d'un cas d'utilisation à l'autre. Du point de vue de l'élaboration des politiques, cela signifie qu'il peut être difficile d'identifier définitivement à l'avance les utilisations à haut ou à faible risque, car le contexte du déploiement est essentiel. Une approche fondée sur les risques est préférable à une approche catégorique consistant à définir les systèmes d'IA qui sont automatiquement considérés comme présentant un risque élevé. Par exemple, le fait de considérer comme à haut risque tous les systèmes d'IA faisant des déductions sur la base de données biométriques pourrait englober des utilisations auxiliaires à risque relativement faible, comme lorsque l'IA est utilisée pour appliquer des filtres ou améliorer la qualité de la vidéo lors d'appels vidéo.

Une approche appropriée pourrait être la suivante :

A) décrire les facteurs, les critères et les dommages potentiels que l'évaluation des risques doit prendre en compte ;
B) fournir, tout au plus, une liste indicative d'utilisations potentiellement plus ou moins risquées qui peuvent être réfutées dans chaque cas particulier ;
C) fournir des conseils continus sur la manière d'évaluer les risques et les avantages sur la base des enseignements tirés au fil du temps.

3. S'appuyer sur les fondamentaux de la législation contraignante et non contraignante

Un régime d'IA souple et adaptable devrait s'appuyer sur les cadres juridiques existants, y compris les réglementations et la législation (« lois contraignantes ») et les « lois non contraignantes » (par exemple, les principes de l'OCDE en matière d'IA). De nombreux secteurs dans lesquels l'IA trouve des applications sont déjà très réglementés (par exemple, les soins de santé, la finance) et les lois et règlements existants prévoient déjà des exigences, des structures de conformité et des recours qui s'appliquent à l'utilisation de l'IA. Toutefois, les lois et règlements existants peuvent également devoir être interprétés d'une nouvelle manière et adaptés aux réalités de l'IA. Les lacunes réglementaires concernant les risques liés à l'IA devraient être comblées par des interventions réglementaires et coréglementaires ciblées, en accordant la priorité aux secteurs où les réglementations existantes ne trouvent pas à s'appliquer.

Le fait de s'appuyer, dans la mesure du possible, sur les cadres juridiques existants réduit le risque de créer des règles qui se chevauchent ou qui sont contradictoires, ce qui pourrait entraîner une incertitude juridique et des protections incohérentes. Les règles existantes en matière de lutte contre la discrimination, de protection des consommateurs, de propriété intellectuelle et, surtout, de protection des données et de la vie privée sont pertinentes pour faire face à bon nombre des risques les plus importants liés à l'IA. Par exemple, en mars 2020, le CIPL a produit une analyse complète

décrivant la manière dont le Règlement général sur la protection des données de l'UE réglemente déjà l'IA en ce qui concerne l'utilisation des données à caractère personnel[ix]. En cas de lacunes dans les cadres législatifs – comme aux États-Unis, qui ne disposent pas actuellement d'une loi fédérale complète sur la protection de la vie privée, les combler est une base importante pour une réglementation solide de l'IA[x]. Lorsque les cadres existants sont pertinents, les organismes de réglementation peuvent favoriser la conformité en publiant des orientations sur la manière dont ces règles s'appliquent à l'IA. En consultant une série de parties prenantes, les régulateurs peuvent identifier les circonstances dans lesquelles de telles orientations seront les plus utiles.

En outre, il est important de reconnaître que les règles existantes peuvent nécessiter une certaine adaptation et une interprétation réglementaire évoluée afin de les aligner sur les développements de la technologie de l'IA. Par exemple, certains principes énoncés dans de nombreuses lois sur la protection des données, tels que la base légale du traitement, la spécification de la finalité et la limitation de l'utilisation des données, peuvent être en contradiction avec les besoins des systèmes d'IA et leur mode de fonctionnement.

En ce qui concerne le concept de base légale, il se peut que les lois actuelles sur la protection des données ne prévoient pas suffisamment de bases légales pour permettre aux développeurs d'IA d'utiliser des catégories sensibles de données à caractère personnel, telles que la santé, le sexe et l'origine ethnique, afin de garantir que le modèle d'IA soit formé et fonctionne d'une manière qui n'aboutisse pas à des résultats biaisés ou discriminatoires. En outre, si les données à caractère personnel sont traitées en fonction d'une base juridique spécifique pour des finalités déterminées, et utilisées uniquement à ces fins ou à des fins « compatibles », en retenant une interprétation étroits de ce qui constitue une « compatibilité », cela peut s'avérer en contradiction avec la nature même du fonctionnement et de l'apprentissage des algorithmes d'IA. Compte tenu du potentiel de l'IA à découvrir des utilisations nouvelles et imprévues des données, ces principes peuvent inutilement contrecarrer des applications bénéfiques de l'IA, à moins qu'ils ne soient interprétés de manière plus large dans le contexte de l'IA. Une telle interprétation élargie consisterait, par exemple, à appliquer une définition plus large de la notion de « compatibilité », qui inclurait toutes les finalités qui n'annulent pas ou n'entrent pas en conflit avec la finalité initiale et qui n'augmentent pas le risque de préjudice pour les personnes. Une autre solution consisterait à considérer la formation algorithmique comme un objectif à part entière, distinct de l'objectif de déploiement de l'algorithme dans un cas d'utilisation particulier. Cela permettrait d'élargir la collecte et l'utilisation des données au cours de la phase de formation afin de garantir une formation et un fonctionnement corrects de l'algorithme. Enfin, des contradictions similaires apparaissent en ce qui concerne les principes de minimisation des données et de limitation de leur conservation, qui peuvent limiter les possibilités d' « apprentissage » des algorithmes par la reconnaissance de corrélations. En résumé, si certains principes traditionnels de protection des données sont interprétés de manière trop rigide, ils peuvent bloquer le développement et le déploiement d'applications d'IA bénéfiques – ou avoir des conséquences involontaires telles que l'introduction de biais indésirables, en limitant l'accès à des données de formation diverses[xi]. Les régulateurs doivent être en mesure de faire évoluer l'interprétation des principes fondamentaux de protection des données par le biais d'orientations réglementaires élaborées en consultation avec les développeurs et les déployeurs d'IA.

Enfin, les règles existantes doivent être complétées par des cadres juridiques non contraignants, des normes industrielles et des outils de régulation fondamentaux élaborés en partenariat avec les parties prenantes, tels que des codes de conduite, des certifications et des modèles d'assurance. Les normes internationales peuvent contribuer à établir des exigences de base pour le développement et le déploiement de l'IA, qui reflètent des conceptions et des valeurs communes issues de processus de développement multipartites. Lors de leur sommet d'Hiroshima en avril 2023[xii], les ministres du numérique et de la technologie du G7 ont réaffirmé le rôle clé des normes, et, en septembre, ils ont

convenu d'élaborer un Code de conduite pour les organisations développant des systèmes d'IA avancés[xiii], tandis que le groupe de travail multipartite sur la certification mène des travaux prometteurs sur la certification de l'IA[xiv]. L'exploitation de cadres juridiques non contraignants tels que les principes de l'OCDE en matière d'IA peut favoriser l'alignement international sur les réglementations relatives à l'IA : par exemple, la version du Parlement européen sur le réglement européen sur l'IA tire sa définition de l' « intelligence artificielle » de ces principes.

4. Responsabiliser les individus grâce à la transparence, l'explicabilité et aux mécanismes de recours

Le CIPL a plaidé en faveur de l'autonomisation de l'individu en tant que principe fondamental d'une réglementation saine en matière de protection de la vie privée, et il en va de même pour l'IA. Pour que l'IA soit digne de confiance et bénéfique pour tous, les réglementations, les cadres de corégulation et les pratiques de l'industrie doivent donner aux individus les moyens d'agir :

- **Transparence.** Les développeurs et les déployeurs de l'IA devraient fournir une transparence appropriée et significative sur les entrées et les opérations des systèmes d'IA, tout en préservant la vie privée et la protection des données, la sécurité, la sûreté et les secrets commerciaux. Cette transparence contextualisée devrait s'étendre aux utilisateurs professionnels des systèmes d'IA, aux auditeurs, aux régulateurs et au grand public.

    - Les **systèmes d'IA à haut risque** doivent documenter la manière dont le système est destiné à être utilisé, les utilisations inappropriées et les risques connus, ainsi que les recommandations à l'intention des déployeurs sur la manière de gérer ces risques.
    - L'**IA générative** nécessite des mesures pour s'assurer que les utilisateurs de ces systèmes comprennent les modalités et les limites des modèles en matière de données. Dans l'idéal, les développeurs et les déployeurs devraient assurer la transparence par le biais de plusieurs mécanismes, notamment des politiques, des conditions de service, des notifications sur les produits et des centres de ressources centralisés.

- **Explicabilité.** L'explicabilité est un aspect de la transparence et un moyen de renforcer la responsabilité et la confiance. Elle exige que les développeurs et les déployeurs expliquent clairement comment les systèmes d'IA affectent les décisions et les résultats ayant un impact sur les individus, tout en gardant à l'esprit l'équilibre à trouver entre notamment l'explicabilité et la sécurité/sûreté du système, ou entre l'explicabilité et l'exactitude des modélisations. Plus l'algorithme est complexe et précis, plus il peut être difficile d'expliquer son fonctionnement. Dans certaines circonstances, des contraintes techniques peuvent également s'opposer à l'explicabilité. Par exemple, il n'est pas toujours possible d'expliquer comment les grands modèles de langage (large language models, LLM) génèrent des résultats spécifiques sur la base d'entrées individuelles ou de paramètres de modèle. Les organisations devront documenter les compromis pertinents afin de démontrer comment et pourquoi elles ont privilégié l'exactitude par rapport à l'explicabilité. De même, les algorithmes d'IA utilisés dans le domaine de la santé et de la médecine en sont un bon exemple : l'IA peut permettre d'obtenir certains bénéfices pour la santé qui ne pourraient pas être obtenus avec des outils autres que l'IA, mais ces algorithmes d'IA ne sont pas toujours explicables en tant que tels. Dans ces circonstances, l'exactitude, et le contrôle continu de cette exactitude, peut primer sur l'explicabilité. En résumé, en fonction du contexte, des risques et des avantages potentiels d'un cas d'utilisation spécifique, l'exigence

d'une explicabilité complète comme condition d'utilisation peut ne pas être appropriée dans tous les cas.

- **Retour de l'information et réparation.** Lorsque des personnes ne comprennent pas une décision prise par l'IA ou estiment avoir subi un préjudice du fait de l'IA, des options claires doivent être prévues pour recueillir les commentaires des utilisateurs, leurs demandes de renseignements, leurs plaintes, permettre une plus grande transparence, le droit de contester la décision, l'exigence d'un contrôle par un humain et, enfin, la réparation, ainsi que l'action des autorités chargées de l'application de la loi, le cas échéant. Dans le cadre de la conception de solutions de bout en bout qui tirent parti de l'IA, les développeurs et les utilisateurs professionnels devraient réfléchir à la manière de renforcer la transparence, le contrôle par un humain, en cas d'utilisation contestée de l'IA, ainsi qu'aux possibilités de saisie de plaintes et de réparation.

II RESPONSABILITÉ ORGANISATIONNELLE (ACCOUNTABILITY) DÉMONTRABLE

5. Faire de la responsabilité organisationnelle et de sa démontrabilité un élément central de la réglementation en matière d'IA

Pour garantir la responsabilité au sein de l'écosystème au sens large, les réglementations devraient faciliter l'utilisation par les organisations de cadres de responsabilité et de programmes de gouvernance qui fournissent les outils et les processus permettant aux organisations de mettre en œuvre toutes les exigences légales et autres normes pertinentes. Comme dans d'autres domaines de la conformité traditionnelle des entreprises et de l'éthique des affaires, et plus récemment dans les espaces de données , de la sécurité et du numérique, la responsabilité doit être intégrée et mise en œuvre à tous les stades du cycle de vie et de « l'infrastructure technologique » de l'IA, y compris l'infrastructure du centre de données, les modèles et les applications de l'IA[xv].

Il existe divers cadres de responsabilité qui fournissent des modèles utiles pour concevoir des programmes de responsabilité organisationnelle et de gouvernance de l'IA, notamment le cadre américain conçu par le NIST « AI Risk Management Framework », le « Model AI Governance Framework » de Singapour et l'« Accountability Framework » du CIPL décrits dans les annexes 1 à 3 du présent rapport[xvi].

Les organisations doivent également être en mesure de *faire preuve de* responsabilité organisationnelle proactive (accountability) en interne, vis-à-vis de leur direction et de leur conseil d'administration, et en externe, vis-à-vis des actionnaires, des investisseurs, des régulateurs et du grand public. Les certifications, les audits, les codes de conduite et les évaluations sont autant d'outils utiles pour démontrer cette responsabilité proactive. En effet, ces mécanismes de responsabilisation sont essentiels dans la politique et la réglementation numériques, y compris pour les développeurs et les déployeurs d'intelligence artificielle, pour les raisons suivantes :

- Ils démontrent à tous les acteurs de l'organisation un engagement et une capacité à garantir que les produits et les services répondent à des critères spécifiques.
- Ils permettent aux organisations de traduire les exigences légales fondées sur des principes et des résultats à atteindre en contrôles démontrables, fondés sur les risques -garantissant ainsi une réglementation plus efficace et une mise en conformité plus effective.
- Ils jouent un rôle important dans la sécurité juridique et le renforcement de la confiance, y compris dans les contextes interentreprises.

Toute réglementation de l'IA devrait explicitement inclure la responsabilité organisationnelle (accountability) et sa démonstration comme des éléments centraux, et également permettre le

développement et l'utilisation de cadres de corégulation, tels que les systèmes de certification et les codes de conduite, qui facilitent et démontrent cette responsabilité.

6. Faire progresser l'adoption de pratiques de gouvernance de l'IA responsables

S'il est vrai que les organisations qui développent et déploient l'IA devraient être tenues de respecter un ensemble de pratiques responsables, les décideurs politiques et les régulateurs devraient également encourager et inciter à l'adoption de pratiques, de cadres, d'outils et de technologies plus larges en matière de responsabilité. Ils devraient collaborer avec les parties prenantes pour élaborer des outils et des cadres permettant d'établir et de démontrer la responsabilité de l'IA. L'objectif devrait être de créer un environnement dans lequel les organisations considèrent l'adoption de cadres de responsabilité rigoureusement développés comme un facteur de différenciation pour créer de la valeur et renforcer la confiance dans leurs pratiques en matière de données - au-delà du respect des obligations légales et réglementaires de base.

Les responsables politiques et les régulateurs devraient également comprendre la force motivante et les défis concernant les pratiques et les solutions technologiques responsables, telles que les technologies renforçant la protection de la vie privée (Privacy Enhancing, PET), et prendre des mesures pour encourager leur développement et leur adoption à plus grande échelle[xvii].

Il convient d'envisager un large éventail d'incitations potentielles à l'obligation de rendre des comptes[xviii], notamment :

- Reconnaître formellement la responsabilité démontrée ou certifiée comme un facteur atténuant dans les actions coercitives et dans l'évaluation des sanctions et/ou des niveaux d'amendes ;
- Utiliser la responsabilité démontrée comme une forme de « licence d'exploitation », en donnant aux organisations responsables une plus grande liberté pour développer et déployer des modèles d'IA de manière responsable ;
- Permettre une utilisation plus large des données dans les projets d'IA pour une recherche socialement bénéfique qui a été validée par des évaluations de risques, des mesures d'atténuation, une surveillance et des contrôles pertinents dans le cadre de programmes de responsabilisation ;
- Permettre aux parties qui acquièrent des systèmes d'IA de satisfaire aux exigences de diligence raisonnable en achetant des systèmes certifiés conformes à des normes reconnues en matière d'IA responsable.
- Utiliser la démonstration de la responsabilité de l'IA comme critère d'éligibilité pour les projets de marchés publics, afin d'inciter les entrepreneurs à obtenir une certification en matière d'IA responsable.

7. Répartir soigneusement les responsabilités, privilégiant des échanges avec et concentrer sur la partie la plus étroitement associée à l'origine du préjudice

L'adoption de mécanismes de responsabilité organisationnelle (accountability) par tous les acteurs de l'écosystème de l'IA permettra d'améliorer le respect des règles et les résultats sur le terrain, et il sera probablement moins nécessaire de recourir à des questions de responsabilité. Cependant, lorsque l'identification de la responsabilité d'une ou plusieurs entités pose problème, des débats animés sont en cours concernant la répartition appropriée des responsabilités entre les parties de l'écosystème de l'IA .

En principe, la responsabilité devrait être attribuée principalement à la partie la plus étroitement associée à la réalisation du dommage en question. Cependant, en pratique,l'attribution

desresponsabilités peut s'avérer plus complexe. L'analyse à effectuer sera influencée par les normes juridiques existantes et des précédents, ainsi que de la diligence des parties pour divulguer des informations pertinentes par la transparence et les obligations de rapports.

Selon les circonstances, la responsabilité peut être attribuée au développeur, au déployeur, aux utilisateurs finaux ou à l'ensemble de ceux-ci. La responsabilité des développeurs pourrait être engagée pour des systèmes qui n'ont pas été suffisamment testés en termes de dommages potentiels, ou qui ont été fournis aux utilisateurs avec des indications trompeuses sur leurs capacités. D'autre part, les utilisateurs partagent la responsabilité de la manière dont ils utilisent les systèmes d'IA, puisqu'ils déterminent s'ils utilisent un système pour un usage à haut risque ou d'une manière expressément contre-indiquée par les orientations fournies par les développeurs.

Comme dans d'autres domaines du commerce, les contrats – y compris les pratiques contractuelles spécifiques et émergentes en matière d'IA – vont jouer un rôle important dans la répartition des responsabilités des parties dans le cycle de vie du développement et du déploiement de l'IA. Par exemple, si un développeur interdit par contrat un cas d'utilisation à haut risque de son produit, le risque de mauvaise utilisation devrait être transféré à l'utilisateur qui a délibérément enfreint les termes de ce contrat. Dans les scénarios où des tiers fournissent des modèles d'IA ou des solutions basées sur l'IA, la ventilation des responsabilités entre développeurs de modèles et déployeurs devrait être précisée dans les contrats.

III. UNE SURVEILLANCE RÉGLEMENTAIRE PARTENARIALE ET EMPIRIQUE

8. Créer des mécanismes de coordination et de coopération entre les organismes de réglementation

L'IA est utilisée dans des secteurs régis par des réglementations et des régulateurs différents. Par exemple, les autorités de protection des données (APD) auront une compétence générale sur le traitement des données à caractère personnel utilisant l'IA. D'autres régulateurs auront des attributions plus sectorielles sur les applications de l'IA – comme dans les soins de santé, le logement, les services financiers, les télécommunications, les produits pharmaceutiques et des disciplines transversales, telles que la propriété intellectuelle. Dans la plupart des cas, il ne devrait pas être nécessaire de créer une nouvelle autorité de régulation de l'IA, car cela entraînerait probablement des chevauchements, des incohérences et un manque de sécurité juridique. Il convient plutôt de :

a) renforcer les compétences et les capacités des régulateurs existants afin qu'ils soient prêts à assurer la surveillance et le contrôle de l'IA ; et
b) permettre une coordination et une collaboration de haut niveau en matière de politique d'IA entre les autorités existantes.

Si chaque autorité de régulation doit rester compétente dans son propre domaine (par exemple, à des fins de sécurité juridique, les APD doivent conserver une compétence générale sur les applications d'IA impliquant le traitement de données à caractère personnel et/ou ayant une incidence sur la vie privée des personnes), **un organisme gouvernemental central permanent de coordination doit être créé pour définir des politiques et des objectifs de haut niveau en matière d'IA** applicables à tous les secteurs et à toutes les industries, et faciliter l'alignement, la coordination réglementaire et l'action conjointe entre les différents organismes de régulation, le cas échéant. L'organe de coordination offre aux régulateurs un espace dans lequel ils peuvent discuter des compromis entre les différents objectifs politiques, notamment l'efficacité, la productivité, l'équité, la protection de la vie privée, la sécurité et la résilience. Il peut également apporter des

éclaircissements précisant vers qui les parties doivent se tourner pour obtenir des orientations dans des circonstances spécifiques de développement et de déploiement de l'IA.

Cette approche serait bénéfique à la fois pour les organisations et les régulateurs en favorisant la cohérence des approches réglementaires ainsi qu'une politique et des orientations holistiques et interdisciplinaires plus faciles à mettre en œuvre et à contrôler par les régulateurs spécialisés et l'industrie au fil du temps. Cette approche peut également être utile pour harmoniser les nouvelles lois et réglementations avec celles qui existent déjà.

Le forum de coopération sur la réglementation numérique (Digital Regulation Cooperation Forum, DRCF) du Royaume-Uni est un exemple de coordination interréglementaire à analyser. Il comprend un directeur général et un personnel permanent, des activités communes, des orientations communes et d'autres mesures réglementaires, ainsi que des projets de collaboration formels et des détachements de personnel. Comme en témoigne son programme de travail pluriannuel sur la transparence algorithmique[xix], l'IA est un point focal dans les travaux du DRCF, D'autres pays, tels que l'Australie, la France, l'Irlande et les Pays-Bas, ont également mis en place des mécanismes de coopération pour les régulateurs[xx].

9. Instaurer une surveillance réglementaire fondée sur la coopération et permettre une innovation réglementaire permanente

Comme la technologie continue d'évoluer, les régulateurs, les techniques de régulation et les outils doivent également évoluer.

a) Dans un monde où des intérêts multiples et concurrents sont en jeu, les régulateurs doivent renforcer leurs capacités et leur mode de fonctionnement. Par exemple, la tâche des autorités chargées de la protection des données ne devrait pas se limiter à la protection des droits fondamentaux des individus, mais également permettre une utilisation responsable et avisée des données et le développement de la technologie de l'IA au profit de la société et de l'économie, d'une manière qui protège les droits fondamentaux. Il faut pour cela changer l'état d'esprit, les priorités et l'action en matière de réglementation. Cette évolution est essentielle si les régulateurs actuels veulent rester pertinents et efficaces dans un nouveau monde numérique.

b) Pour être stratégiques et efficaces, les régulateurs doivent adopter une approche fondée sur le risque. Il faut pour cela comprendre les risques et les avantages des systèmes d'IA. Il convient également de se concentrer sur les domaines qui présentent les risques les plus élevés pour les individus et la société, tout en préservant les avantages de la technologie de l'IA et de ses progrès. Il est enfin nécessaire que les régulateurs établissent des priorités dans l'ensemble de leur travail – stratégie réglementaire, orientation, supervision et application – et qu'ils se concentrent sur les domaines qui créent les risques les plus élevés pour les individus et la société.

c) Dans une société fondée sur le numérique et l'IA, les mécanismes de contrôle traditionnels fondés exclusivement ou principalement sur l'exécution d'un contrôle *a posteriori* pourraient ne plus suffire. En effet, le fait de s'en remettre à la seule application de la loi pour remédier aux défaillances du marché n'aboutira pas aux résultats souhaités. Compte tenu du rythme des progrès de la technologie de l'IA et de la nécessité de comprendre ses risques et ses avantages, il est urgent d'adopter une approche plus coopérative fondée sur un engagement constructif permanent entre les régulateurs et les entités réglementées, sur le partage

d'expériences et d'informations sur les évolutions technologiques et sur la collaboration en vue d'élaborer des objectifs de conformité réalistes et des interprétations des règles applicables. Pour ce faire, les régulateurs et les entités réglementées doivent être transparents et prêts à s'engager dans un échange d'informations constructif en temps réel, à mesure que les technologies et les pratiques commerciales évoluent[xxi]. Investir dans des mesures *ex ante* telles que l'incitation à une responsabilisation proactive et démontrable permettra probablement d'obtenir de meilleurs résultats qu'une mise en œuvre *ex post,* onéreuse. Bien entendu, l'application de la législation doit rester une option réglementaire et un levier important pour les infractions répétées, graves et négligentes qui causent un préjudice réel aux personnes et à la société.

d) Les outils réglementaires innovants, tels que les bacs à sable et le prototypage de politiques, peuvent s'avérer efficaces pour la surveillance réglementaire des nouvelles technologies telles que l'IA. Ils permettent aux régulateurs d'acquérir une connaissance plus approfondie et une expérience directe des applications et des développements de l'IA, destinées au marché général. Ils prévoient également une sphère de sécurité permettant à l'industrie de tester les risques et les avantages de l'innovation responsable avec un lien direct avec le régulateur compétent. Bien sûr, dans ce cadre, les gouvernements doivent fournir un financement et des ressources appropriés aux régulateurs pour qu'ils développent et s'engagent dans des bacs à sable réglementaires et qu'ils soient en mesure d'étendre ces activités à un plus grand nombre de participants, y compris sur une base sectorielle.

Les bacs à sable réglementaires : Les bacs à sable réglementaires sont des mécanismes importants pour l'exploration et l'expérimentation réglementaires, car ils constituent un banc d'essai pour l'application des lois à des produits et services innovants dans des contextes réels sous la supervision d'un régulateur[xxii]. Ils peuvent être utilisés pour aider à aborder et à résoudre certains des aspects les plus difficiles du déploiement de l'IA dans le contexte des exigences juridiques en vigueur, en particulier celles qui semblent incohérentes ou en contradiction avec les nouvelles technologies. En voici quelques exemples :

- Depuis 2020, le bureau de l'ICO au Royaume-Uni (information Commissioner Office) a mis en place un programme de bac à sable, en mettant l'accent sur les technologies émergentes et la biométrie[xxiii] ;
- A Singapour, le bac à sable pour la réglementation des données (Data Regulatory Sandbox) de l'Infocomm Media Development Authority (IMDA) permet aux entreprises d'obtenir des orientations en matière de réglementation pour les projets innovants et à grand volumes de données. L'IMDA gère également un bac à sable spécifique destiné à favoriser le développement et l'adoption de technologies renforçant la protection de la vie privée (Privacy-enhancing Technologies,PET)[xxiv] ;
- L'autorité norvégienne de protection des données (Datatilsynet) a lancé un bac à sable réglementaire spécial pour les applications d'IA[xxv] ;
- Le gouvernement colombien a mis en place un bac à sable réglementaire pour promouvoir le respect de la vie privée dès la conception et par défaut dans les projets d'IA[xxvi] ;
- La CNIL française gère un bac à sable qui a mené à bien des projets dans les domaines de la santé numérique et des technologies éducatives. En 2023, elle a annoncé une nouvelle initiative axée sur l'IA[xxvii] ;
- le projet de loi sur l'IA de l'UE permettra, après adoption, aux États membres d'établir des bacs à sable réglementaires pour l'IA, et pourrait même les y obliger. A ce stade, l'Espagne a été le premier État membre à piloter un bac à sable pour l'IA[xxviii].

Les bacs à sable réglementaires devraient être conçus de manière à encourager l'innovation, le partage d'informations et d'autres modes de coopération. Tout cadre réglementaire en matière d'IA devrait fournir une base statutaire explicite permettant aux régulateurs de mettre en place des bacs à sable, y compris des bacs à sable interréglementaires avec les régulateurs appropriés et pertinents, notamment les régulateurs de la protection des données, de la concurrence, des médias, des consommateurs, de la santé/pharmacie, des télécommunications et des finances. Les règlements doivent tenir compte de la manière dont les autorités juridiques ou les priorités en matière d'application de la loi peuvent affecter la participation des entreprises à ces initiatives. En même temps, pour que le public ait confiance dans les résultats, les bacs à sable devraient inclure des garanties -par exemple, que les individus continueront à être protégés contre les préjudices, même si l'expérimentation politique a lieu.

Prototypage de politiques : Il s'agit de projets pilotes qui mobilisent des acteurs publics et privés pour explorer, évaluer et développer conjointement différents modèles législatifs de gouvernance avant leur mise en œuvre effective. Le processus consiste généralement à sélectionner un groupe de participants, tels que des entreprises technologiques en phase de démarrage, pour développer et appliquer des prototypes de politiques en partenariat avec des experts gouvernementaux, industriels et universitaires. Le programme OpenLoop de Meta a été l'un des principaux praticiens du prototypage de politiques, notamment pour la proposition de loi sur l'IA de l'UE[xxix]. Par ailleurs, l'IMDA de Singapour dispose d'un programme de prototypage de politiques dans le cadre de son bac à sable pour la réglementation des données (Data Regulatory Sandbox) qui s'est concentré sur la notification, le consentement et la divulgation,la transparence et l'explicabilité de l'IA et la transparence et le consentement dans le métavers -y compris les contextes d'application de l'intérêt légitime en tant que base de traitement[xxx].

10. S'efforcer de parvenir à une interopérabilité mondiale

Compte tenu de la nature mondiale de la technologie de l'IA – des données qu'elle utilise pour la formation à la recherche et au développement, en passant par l'infrastructure informatique et les applications transfrontalières – il est clair qu'aucun gouvernement ne peut aborder de manière satisfaisante la politique et la réglementation en matière d'IA de manière isolée. La coopération au niveau international est essentielle pour garantir que les individus et les sociétés du monde entier puissent compter sur les avantages d'une IA digne de confiance et responsable, et que les nouveaux risques soient évalués et atténués de manière continue. Ces travaux bénéficieraient d'un forum international spécifique permettant aux gouvernements et aux autres parties prenantes de coopérer sur la politique en matière d'IA.

En outre, la coopération internationale doit favoriser l'interopérabilité des politiques et des réglementations en matière d'IA. Comme l'a noté le CIPL dans le contexte de la protection des données, l'interopérabilité mondiale permet de fournir des services transfrontaliers de manière responsable, d'élargir l'accès, de réduire les coûts de mise en conformité, d'accroître la sécurité juridique et de garantir une protection cohérente des droits et des intérêts des personnes[xxxi]. Les différentes juridictions auront leurs propres priorités, traditions juridiques et corpus de réglementations existantes, mais elles pourront peut-être se rallier autour de principes et de fondamentaux communs dans l'examen de la politique et de la réglementation en matière d'IA – similaires à ceux que le CIPL a mis en avant dans le présent document. Elles peuvent également prendre des mesures pour codifier l'interopérabilité par le biais de mécanismes de reconnaissance et de certification, notamment en participant au système des règles mondiales de protection de la vie privée dans le contexte de la protection des données et des flux de données transfrontaliers (Global Cross-Border Privacy Rules, CBPR)[xxxii]. Des efforts encourageants ont été déployés en faveur de l'interopérabilité de l'IA dans le cadre de l'initiative du G7 susmentionnée, des principes de l'OCDE en matière d'IA[xxiii], des accords commerciaux et économiques tels que l'accord de partenariat

économique numérique (Digital Economic Partnership Agreement, DEPA)[xxxiv] et du partenariat mondial sur l'IA[xxxv].

ANNEXE I – CADRE DE RESPONSABILITÉ DU CIPL

ANNEXE II - MISE EN CORRESPONDANCE DES MEILLEURES PRATIQUES EN MATIERE DE GOUVERNANCE DE L'IA AVEC LE CADRE DE RESPONSABILISATION DU CIPL

Le tableau suivant présente des exemples d'activités d'IA responsables lancées par des organisations sélectionnées de différents secteurs, géographies et tailles, basées sur le cadre de responsabilité du CIPL et mises en correspondance avec chaque élément de responsabilité. Ces pratiques ne constituent pas des normes obligatoires du secteur, mais servent d'exemples spécifiques calibrés en fonction des risques, du contexte sectoriel, du modèle d'entreprise, de la taille et du niveau de maturité des organisations.

ÉLÉMENT DE RESPONSABILITÉ	**PRATIQUES CONNEXES**
Encadrement et supervision	• Engagement public et ton donné par la direction pour respecter l'éthique, les valeurs et les principes spécifiques dans le développement, le déploiement et l'utilisation de l'IA • Processus et prise de décision institutionnalisés en matière d'IA avec des critères d'escalade • Conseils et comités (internes ou externes) de surveillance de l'IA/de l'éthique - pour examiner les cas d'utilisation risqués de l'IA et pour améliorer en permanence les pratiques en matière d'IA • Nomination d'un membre du conseil d'administration pour la supervision de l'IA • Nomination d'un dirigeant de l'IA, d'un responsable de l'IA ou d'un champion de l'IA • Mise en place d'un conseil ou d'un comité interdisciplinaire interne sur l'IA • Garantie de l'inclusion et la diversité dans les équipes de développement de modèles et de produits d'IA
Évaluation des risques	• Outils d'évaluation de l'impact algorithmique ou de l'équité pour contrôler et tester en permanence les algorithmes afin d'éviter les préjugés humains, les discriminations injustes et les dérives conceptuelles tout au long du cycle de vie de l'IA • Évaluation de l'impact sur l'éthique / évaluation de l'impact sur les droits de l'homme / évaluation de l'impact sur la protection des données • Élaborer des méthodes normalisées d'évaluation des risques, qui tiennent compte des avantages, de la probabilité et de la gravité des facteurs de risque pour les individus et/ou la société, du niveau de contrôle humain impliqué dans les décisions individuelles automatisées ayant des effets juridiques, ainsi que de la possibilité d'expliquer ces décisions en fonction du contexte et de leur auditabilité • Documentation des compromis (par exemple, précision - minimisation des données, sécurité - transparence, impact sur un petit nombre - avantages pour la société) pour les traitements à haut risque dans le cadre de l'évaluation des risques • Évaluation de la qualité des données par le biais d'indicateurs de performance clés • Évaluation des données par rapport à l'objectif - qualité, provenance, personnelles ou non, synthétiques, sources internes ou externes • Cadre pour la préparation des données et l'évaluation des modèles - y compris l'ingénierie des caractéristiques, la validation croisée, le backtesting, les KPI validés par l'entreprise • Travail en étroite collaboration avec les experts en affaires et en données (analystes de données, ingénieurs de données, ingénieurs informatiques et logiciels) afin d'évaluer régulièrement les besoins et la précision des résultats pour s'assurer que le modèle peut être utilisé correctement

Politiques et Procédures	• Adoption de politiques et des procédures spécifiques en matière d'IA sur la manière de concevoir, d'utiliser ou de vendre l'IA • Politiques sur l'application de la protection de la vie privée et de la sécurité dès la conception dans le cycle de vie de l'IA • Règle fixant le niveau de vérification de l'entrée et de la sortie des données • Essais pilotes des modèles d'IA avant leur diffusion • Utilisation de données protégées (par exemple, des données cryptées, pseudonymisées, symbolisées ou synthétiques) dans certains modèles • Utilisation d'ensembles de données de haute qualité mais plus restreints • Utilisation de modèles fédérés d'apprentissage de l'IA, en tenant compte des compromis avec la sécurité des données et les responsabilités des utilisateurs • Considérations particulières pour les organisations qui créent et vendent des modèles, des logiciels et des applications d'IA • Listes de contrôle ou outils de diligence raisonnable/d'auto-évaluation pour les partenaires commerciaux utilisant l'IA • Définition des étapes de l'escalade en ce qui concerne les rapports, la gouvernance et l'analyse des risques • Phase d'idéation entre toutes les parties prenantes (scientifiques des données, entreprises, utilisateurs finaux, fonctions de contrôle) au cours de laquelle les besoins, les résultats, les règles de validation, la maintenance, le besoin d'explication, le budget sont discutés
Transparence	• Différents besoins de transparence envers les individus, les régulateurs, les partenaires commerciaux et en interne aux différents stades du cycle de vie de l'IA, en fonction du contexte • Des informations adéquates communiquées de manière simple et facile à comprendre • Tenir compte du fait que l'IA doit être inclusive et accessible aux personnes ayant des besoins particuliers ou des handicaps • Mettre en place une piste de transparence pour expliquer les décisions et le fonctionnement général de l'algorithme afin de rendre le système d'IA contrôlable • Expliquer qu'il s'agit d'une décision de l'IA/ML, s'il y a risque de confusion (test de Turing) • Fournir des informations contrefactuelles • Comprendre les attentes des clients et déployer en fonction de leur degré de préparation à l'IA • Mettre en place une transparence progressive • De la boîte noire à la boîte de verre - examiner les données ainsi que l'algorithme/le modèle • L'aspiration à l'explicabilité permet de comprendre la boîte noire et d'instaurer la confiance • Définir les critères de déploiement des technologies de l'IA au sein de l'organisation sur la base de scénarios d'utilisation et les communiquer à l'utilisateur • Produire des fiches de modèle (courts documents accompagnant les modèles d'IA pour décrire le contexte dans lequel le modèle doit être utilisé, quelle est la procédure d'évaluation) • Centre de données pour la transparence en matière de gouvernance des données, d'accessibilité des données, de lignage des données, de modification des données, de qualité des données, de définition, etc. • Adapter la transparence au risque identifié : par exemple, filigrane pour les résultats de l'IA générative
Formation et Sensibilisation	• Formation des scientifiques des données, y compris la façon de limiter et de traiter les préjugés • Formation transversale – professionnels de la protection de la vie privée et ingénieurs • Formation à l'éthique et à l'équité pour les équipes technologiques • Cas d'utilisation où le déploiement problématique de l'IA a été interrompu • Rôle des « traducteurs » dans les organisations, expliquant l'impact et le fonctionnement de l'IA

Suivi et Vérification	• Capacité de l'être humain à participer à la conception, à la surveillance et à la réparation • Capacité de compréhension humaine de l'activité et des processus grâce à l'IA • Possibilité d'audit humain des données d'entrée et de sortie • Possibilité d'un contrôle humain des décisions individuelles ayant des effets juridiques • Suivi de l'écosystème de l'entrée des données, du traitement des données et de la sortie des données • Recours à différentes techniques d'audit • Recours à des techniques de test contrefactuel • Prédéfinition des contrôles d'audit de l'IA • Équipe d'audit interne spécialisée sur l'IA et les autres technologies émergentes • Les processus doivent permettre un contrôle ou une intervention humaine dans le système d'IA lorsque cela est techniquement possible et raisonnablement nécessaire • Suivi du modèle (backtesting et boucle de retour de l'information) et processus de maintenance
Réponse et application de la loi	• Processus et procédures pour recevoir et traiter le retour d'information et les plaintes • Mécanismes de réparation pour remédier à une décision de l'IA • Réparation par un être humain, non pas un robot • Canal de retour de l'information

[i] Dans le présent rapport, le CIPL utilise le terme « intelligence artificielle » conformément à la définition des « systèmes d'intelligence artificielle » élaborée par le National Institute of Standards and Technology (NIST) des États-Unis dans son Risk Management Framework 1.0, adaptée d'une définition comparable élaborée par l'Organisation de coopération et de développement économiques (OCDE) : « *Un système d'IA est un système technique ou mécanique qui peut, pour un ensemble donné d'objectifs, générer des résultats tels que des prédictions, des recommandations ou des décisions influençant des environnements réels ou virtuels. Les systèmes d'IA sont conçus pour fonctionner avec différents niveaux d'autonomie*. » Voir https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[ii] Par exemple : L'autorité italienne de protection des données « Garante » interdit ChatGPT le 30 mars 2023, disponible à l'adresse https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870832 ; Le Conseil européen de protection des données (European Data Protection Board, EDPB) crée un groupe de travail sur Chat GPT, disponible à l'adresse https://edpb.europa.eu/news/news/2023/edpb-resolves-dispute-transfers-meta-and-creates-task-force-chat-gpt_en ; L'autorité britannique de la concurrence et des marchés (Competition and Markets Authority) lance un examen initial des modèles d'intelligence artificielle, disponible à l'adresse https://www.gov.uk/cma-cases/ai-foundation-models-initial-review ; Le Bureau du commissaire à la protection de la vie privée du Canada (The Office of the Privacy Commissioner) a lancé une enquête sur ChatGPT, disponible à l'adresse https://www.priv.gc.ca/en/opc-news/news-and-announcements/2023/an_230404/ ; Déclaration commune sur les efforts d'application contre la discrimination et les préjugés dans les systèmes automatisés du Bureau de protection financière des États-Unis (Consumer Financial Protection Bureau), Division des droits civils du ministère de la Justice (Department of Justice's Civil Rights Division), Commission de l'égalité des chances en matière d'emploi (Equal Employment Opportunity Commission) et Commission fédérale du commerce (Commission and Federal Trade Commission), disponible à l'adresse https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf.

[iii] Selon l'OCDE, plus de 800 initiatives et stratégies politiques en matière d'IA ont été élaborées dans 69 pays, régions et dans l'Union européenne https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf.

[iv] CIPL, "First Report: Artificial Intelligence and Data Protection in Tension" (« Premier rapport : Contradictions entre l'Intelligence Artificielle et la protection des données »), octobre 2018, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_first_ai_report_-_ai_and_data_protection_in_tension_2.pdf ; "Second Report: Hard Issues and Practical Solutions" (« Second rapport : Questions difficiles et solutions pratiques »), février 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_report_-artificial_intelligence_and_data_protection_-_hard_issues_and_practical_solutions 27_february_2020.pdf ; "Artificial Intelligence and Data Protection: How the GDPR Regulates AI" (« Intelligence artificielle et protection des données : Comment le GDPR réglemente l'IA »), mars 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl- hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020.pdf.

[v] CIPL, "Response to NTIA Request for Comment on AI Accountability Policy" (« Réponse à la demande de commentaire au NTIA concernant la politique de responsabilité en matière d'IA »), juin 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_ntia_ai_accountability_policy_june2023.pdf ; "CIPL's Top Ten Recommendations for Regulating AI in Brazil" (« Les dix principales recommandations du CIPL pour réglementer l'IA au Brésil »), octobre 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_ntia_ai_accountability_policy_june2023.pdf ; "Response to UK DCMS Proposed Approach to Regulating AI" (« Réponse à l'approche proposée par le DCMS britannique pour réglementer l'IA »), septembre 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_uk_dcms_proposed_approach_to_regulating_ai_23_09_22.pdf; "CIPL Response to the EU Commission's Consultation on the Draft AI Act" (« Réponse du CIPL à la Consultation de la Commission européenne sur le projet de loi sur l'IA »), juillet 2021, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_consultation_on_the_draft_ai_act_29_july_2021.pdf.

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

087493.0000002 DMS 303819731v3

[vi] « Il n'existe pas de définitions universellement acceptées pour les développeurs, les déployeurs et les utilisateurs d'IA ; en effet, une taxonomie élaborée conjointement par l'UE et les États-Unis en 2023 mentionne que les définitions de déploiement, de développeur et d'utilisateur sont « en attente » (voir "EU-U.S. Terminology and Taxonomy for Artificial Intelligence: First Edition" (« Terminologie et taxonomie UE-États-Unis pour l'intelligence artificielle : Première édition »), https://www.nist.gov/system/files/documents/noindex/2023/05/31/WG1%20AI%20Taxonomy%20and%20Terminology%20Subgroup%20List%20of%20Terms.pdf). Dans le présent document, nous utilisons le terme « développeurs » pour désigner les parties qui conçoivent et construisent des systèmes d'IA, le terme « déployeurs » pour désigner les parties qui mettent ces systèmes à disposition et le terme « utilisateurs » pour désigner les utilisateurs finaux qui exploitent ces systèmes de manière régulière. Une seule entité peut jouer chacun de ces rôles à différents moments ou simultanément.

[vii] *K-Nearest Neighbors Algorithm (L'algorithme des plus proches voisins)*, IBM, disponible à l' adresse https://www.ibm.com/uk-en/topics/knn#:~:text=Next%20steps-,K%2DNearest%20Neighbors%20Algorithm,of%20an%20individual%20data%20point

[viii] *3 ways autonomous farming is driving a new era of agriculture (3 façons dont l'agriculture autonome ouvre une nouvelle ère pour l'agriculture, Forum économique mondial)*, 2022 Disponible à l'adresse : https://www.weforum.org/agenda/2022/01/autonomous-farming-tractors-agriculture/

[ix] CIPL, "Artificial Intelligence and Data Protection: How the GDPR Regulates AI" « Intelligence artificielle et protection des données : comment le GDPR réglemente l'IA », mars 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl- hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020.pdf.

[x] Pour en savoir plus sur l'intersection entre l'IA et la réglementation en matière de protection des données, voir le Premier rapport du CIPL sur l'IA – *Artificial Intelligence and Data Protection in Tension (Contradiction entre l'Intelligence Artificielle et la protection des données)* ; le Deuxième rapport du CIPL sur l'IA - *Hard Issues and Practical Solutions (Questions difficiles et solutions pratiques)* ; et Livre blanc du CIPL/Hunton Andrews Kurth - *How the GDPR Regulates AI (Comment le GDPR réglemente l'IA)* – tous disponibles ici https://www.informationpolicycentre.com/ai-project.html.

[xi] Pour plus d'informations sur ce sujet, voir CIPL, "First Report: Artificial Intelligence and Data Protection in Tension" (« Premier rapport : Contradiction entre l'Intelligence Artificielle et la protection des données »), octobre 2018, cipl_first_ai_report_-_ai_and_data_protection_in_tension 2.pdf (informationpolicycentre.com), et CIPL, "Second Report: Hard Issues and Practical Solutions" (« Second rapport : Questions difficiles et solutions pratiques »), février 2020, AI Project - Centre for Information Policy Leadership (informationpolicycentre.com).

[xii] Ministerial Declaration The G7 Digital and Tech Ministers' Meeting 30 April 2023 (g7digital-tech-2023.go.jp)

[xiii] "G7 Hiroshima AI Process: G7 Digital & Tech Ministers Statement" (« Processus d'Hiroshima du G7 en matière d'IA : Déclaration des ministres du G7 chargés du numérique et de la technologie »), septembre 2023, disponible à 3e39b82d-464d-403a-b6cb-dc0e1bdec642-230906_Ministerial-clean-Draft-Hiroshima-Ministers-Statement68.pdf (politico.eu).

[xiv] Groupe de travail sur la certification, "Unlocking the Power of AI – Steps for Effective Certification to Help Drive Innovation and Trust" (« Libérer le pouvoir de l'IA - Étapes d'une certification efficace pour favoriser l'innovation et la confiance »), juin 2023, https://www.responsible.ai/post/white-paper-draft-from-the-certification-working-group.

[xv] Pour une discussion plus approfondie sur la gouvernance à travers ces couches de la pile technologique de l'IA, voir Microsoft, "Governing AI: A Blueprint for the Future" (« Gouverner l'IA : Un plan pour l'avenir »), mai 2023, httpshttps://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw.

[xvi] National Institute of Standards and Technology (NIST), "AI Risk Management Framework" (« Cadre de gestion des risques liés à l'IA »), https://www.nist.gov/itl/ai-risk-management-framework ; Commission de protection des données personnelles (Personal Data Protection Commission, PDPC), "Singapore's Approach to AI Governance" (« L'approche de Singapour en matière de gouvernance de l'IA »), https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI- Governance-Framework, CIPL, "Organizational Accountability" (« Responsabilité organisationnelle »), https://www.informationpolicycentre.com/organizational-accountability.html.

[xvii] Voir le rapport du CIPL sur les technologies renforçant la protection de la vie privée (à paraître, automne 2023).

[xviii] Incitation à la responsabilité : Comment les autorités de protection des données et les législateurs peuvent encourager la responsabilité, CIPL, 23 juillet 2018, disponible à https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_paper_2_-incentivising_accountability_-_how_data_protection_authorities_and_law_makers_can_encourage_accountability.pdf.

[xix] L'Autorité de la concurrence et des marchés (La Competition and Markets Authority, CMA), le bureau du commissaire à l'information (Information Commissioner's Office, ICO), le Bureau des communications (Office of Communications, Ofcom) et l'autorité de conduite financière (Financial Conduct Authority, FCA) participent au DRCF. Voir https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum. Pour en savoir plus sur le chantier de la transparence algorithmique du DRCF, veuillez consulter le site https://www.gov.uk/government/publications/transparency-in-the-procurement-of-algorithmic-systems- findings-from-our-workshops.

[xx] Aux Pays-Bas, l'Autorité pour les consommateurs et les marchés (Authority for Consumers and Markets, ACM), l'Autorité néerlandaise de protection des données (Dutch Data Protection Authority, AP), l'Autorité néerlandaise pour les marchés financiers (Dutch Authority for the Financial Markets, AFM) et l'Autorité néerlandaise des médias (Dutch Media Authority, CvdM) ont lancé le Forum de coopération sur la réglementation numérique (Digital Regulation Cooperation Forum, SDT). Voir https://www.acm.nl/en/about-acm/cooperation/national-cooperation/digital-regulation-cooperation-platform-sdt. Le Centre d'expertise français pour la régulation des plateformes numériques (PEReN) a été créé sous l'autorité des ministères de l'économie, de la culture et du numérique. S'appuyant sur une expertise en science des données, il est une source d'expertise technique et de soutien aux régulateurs numériques français. Voir https://www.peren.gouv.fr/en/. Le forum des régulateurs australiens des plateformes numériques (Digital Platform Regulators Forum) réunit la Commission australienne de la concurrence et de la consommation (Australian Competition and Consumer Commission, ACCC), Autorité australienne des communications et des médias (Australian Communications and Media Authority, ACMA), le Commissaire à l'eSécurité (eSafety Commissioner - eSafety) et le Bureau de la Commission australienne de l'information (Office of the Australian Information Commissioner, OAIC). Voir https://www.accc.gov.au/about-us/media/media-updates/communique-digital-platforms-regulators-forum. L'Irlande a créé le réseau des régulateurs économiques (Economic Regulators Network), qui regroupe sept régulateurs. Voir https://www.econreg.ie/about/our-members/.

[xxi] Voir Christopher Hodges et CIPL, "Organizational Accountability in Data Protection Enforcement" (« Responsabilité organisationnelle dans l'application de la protection des données »), octobre 2021, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_organizational_accountability_in_data_protection_enforcement_-. how_regulators_consider_accountability_in_their_enforcement_decisions_6_oct_2021.pdf

[xxii] Voir l'article du CIPL "Regulatory Sandboxes in Data Protection – Constructive Engagement and Innovative Regulation in Practice" (« Bacs à sable réglementaires en matière de protection des données - Engagement constructif et réglementation innovante dans la pratique) - 8 mars 2019 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protectionconstructive_engagement_and_innovative_regulation_in_practice 8_march_201 9.pdf.

[xxiii] Bac à sable réglementaire de ICO https://ico.org.uk/for-organisations/regulatory-sandbox/.

[xxiv] Bac à sable pour la réglementation des données de IMDA, https://www.imda.gov.sg/how-we-can-help/data-innovation/data-regulatory-sandbox.

[xxv] Bac à sable réglementaire sur l'IA de Datatilsynet, disponible à l' adresse https://www.datatilsynet.no/en/regulations-andtools/sandbox-for-artificial-intelligence/.

[xxvi] Bac à sable sur la protection de la vie privée dès la conception et par défaut dans les projets d'intelligence artificielle, Surintendance colombienne de l'industrie et du commerce (Columbian Superintendence of Industry and Commerce), disponible à l'adresse https://globalprivacyassembly.org/wp-content/uploads/2021/07/B6.SIC-Colombia-Sandbox-on-privacy-by-design-and-by-default-in-AI-projects.pdf

[xxvii] CNIL, "Digital health and EdTech: the CNIL publishes the results of its first sandboxes" (« Santé numérique et EdTech : la CNIL publie les résultats de ses premiers « bacs à sable »), juillet 2023, https://www.cnil.fr/en/digital-health-and-edtech-cnil-publishes-results-its-first-sandboxes.

[xxviii] Au moment de la rédaction du présent rapport, les décideurs politiques de l'UE n'avaient pas encore décidé s'il fallait rendre la création de bacs à sable pour l'IA par les États membres facultative ou obligatoire. Voir Luca Bertuzzi, "EU Council sets path for innovation measures in AI Act's Negotiations" (« Le Conseil de l'UE ouvre la voie à des mesures d'innovation dans le cadre des négociations relatives à la loi sur l'IA »), Euractiv, 10 juillet 2023. https://www.euractiv.com/section/artificial-intelligence/news/eu-council-sets-path-for-innovation-measures-

in-ai-acts-negotiations/. Voir également "First Regulatory Sandbox on Artificial Intelligence Presented" (« Présentation du premier « bac à sable » réglementaire sur l'intelligence artificielle »), Digibyte, 27 juin 2022, https://digital-strategy.ec.europa.eu/en/news/firstregulatory-sandbox-artificial-intelligence-presented.

[xxix] Voir "Introducing Open Loop, a global program bridging tech and policy innovation" (« Open Loop, un programme mondial qui fait le lien entre l'innovation technologique et l'innovation politique »), disponible à l'adresse https://ai.facebook.com/blog/introducing-open-loop-a-global-program-bridging-tech-and-policy-innovation / ; Évaluation de l'impact de l'IA : Une expérience de prototypage de politiques, https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3772500_code715910.pdf?abstractid=3772500&mirid=1 et https://openloop.org/programs/open-loop-eu-ai-act-program/.

[xxx] IMDA, "Policy Prototyping" (« Prototypage de politiques »), Policy Prototyping | IMDA - Infocomm Media Development Authority. TC Labs de Meta est un partenaire du programme de l'IMDA.

[xxxi] CIPL, "Ten Principles for a Revised US Privacy Framework" (« Dix principes pour un cadre révisé de protection de la vie privée aux États-Unis »), mars 2019, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_principles_for_a_revised_us_privacy_framework.pdf.

[xxxii] CIPL, "International Data Flows: Cross Border Privacy Rules, Privacy Recognition for Processors, and Global CBPR and PRP: Frequently Asked Questions" (« Flux de données internationaux : Règles transfrontalières de protection de la vie privée, reconnaissance de la protection de la vie privée pour les sous-traitants, et CBPR et PRP au niveau mondial : Foire aux questions »), juin 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_cpbr_and_prp_faq_jun23.pdf.

[xxxiii] OCDE, "Recommendation of the Council on Artificial Intelligence" (« Recommandations du Conseil sur l'intelligence artificielle »), mai 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[xxxiv] L'accord de partenariat économique numérique (Digital Economic Partnership Agreement, EPA) est un accord conclu entre la Nouvelle-Zélande, Singapour et le Chili. Voir https://www.mfat.govt.nz/en/trade/free-trade-agreements/free-trade-agreements-in-force/digital-economy-partnership-agreement-depa/.

[xxxv] Partenariat mondial sur l'intelligence artificielle, https://gpai.ai/.