

AI Governance & Safety Canada
Gouvernance et sécurité de l'IA Canada

Le 3 novembre 2023

Monsieur Joël Lightbound, député
Président
Comité permanent de l'industrie et de la technologie
Chambre des communes
Parlement du Canada

Monsieur,

Le Canada se trouve au cœur d'une révolution mondiale de l'intelligence artificielle (IA), et l'ampleur et la complexité des risques qui y sont associés sont tout simplement stupéfiantes. La tâche à laquelle votre comité est confronté en modifiant la *Loi sur l'intelligence artificielle et les données* (LIAD) pour répondre aux besoins des Canadiens est donc à la fois essentielle et immense.

[Gouvernance et sécurité de l'IA Canada](#) est une organisation à but non lucratif multipartite et une communauté de personnes à travers le pays qui travaillent à renforcer le leadership du Canada en matière de gouvernance et de sécurité de l'intelligence artificielle. Nous avons contribué aux tables rondes d'Innovation, Sciences et Développement économique Canada (ISDE) sur le code de pratique volontaire pour l'IA générative et avons récemment publié notre livre blanc [Gouverner l'IA : Un plan pour le Canada](#).

Les recommandations relatives à la LIAD que nous présentons dans ce document sont le résultat de plusieurs mois de travail et de consultations approfondies avec les parties prenantes nationales et internationales.

Vous constaterez que nous recommandons une réécriture approfondie de la Loi. Notre stratégie consiste à présenter ce qu'il faudrait réellement pour protéger les Canadiens des risques actuels et futurs, puis à travailler avec le Comité sur les compromis qui doivent être faits. Si le temps n'était pas un problème, nous recommanderions de séparer la LIAD du projet de loi et de la réintroduire après de longues consultations et délibérations. Toutefois, compte tenu de l'accélération des avancées en matière d'IA et des préjudices déjà ressentis, les Canadiens n'ont pas ce luxe. Nous avons besoin d'une législation opérationnelle dès maintenant.

Nous encourageons donc les membres du Comité à ne pas abandonner la *Loi sur l'intelligence artificielle et les données*, mais à prendre le temps de comprendre l'ensemble des risques liés à l'IA qu'il convient de gérer, et à préparer une législation qui servira les Canadiens aujourd'hui et dans les années à venir.

Nous restons à votre disposition pour toute assistance dont vous auriez besoin.

Veillez agréer, Monsieur, mes sincères salutations.

Wyatt Tessari L'Allié
Fondateur et directeur général
[Gouvernance et sécurité de l'IA Canada](#)

contact@aigs.ca

Recommandations quant à la *Loi sur l'intelligence artificielle et les données*

Mémoire au Comité permanent de l'industrie et de la technologie sur le projet de loi C-27

Le 3 novembre 2023

Table des matières

Lettre au président du Comité	1
Résumé d'une page	3
Partie I : Pourquoi le Canada a-t-il besoin d'une législation en matière d'IA?	4
Paysage en évolution des risques liés à l'IA	5
Le rôle de la législation	6
Législation dans d'autres administrations	7
Législation canadienne existante et lacunes exigeant l'adoption d'une loi consacrée à l'IA	8
Partie II : Modifications recommandées	9
Utiliser quatre catégories de risques et maintenir des exigences proportionnées	10
Systèmes à risque inacceptable (SRI)	10
Systèmes d'IA à usage général à haut risque	13
Systèmes d'IA à usage unique à haut risque	17
Systèmes à risque modéré ou faible	18
Comblent les lacunes critiques	18
Fournir au gouvernement les capacités dont il a besoin	20
Partie III : Formulation particulière à modifier dans le projet de loi	22
<i>La formulation sera fournie sous forme d'addendum au présent document après que le texte des modifications apportées par le gouvernement aura été communiqué</i>	22

Résumé d'une page

Le Canada est au cœur d'une révolution de l'IA en plein essor et doit trouver un moyen d'exploiter ses nombreux avantages tout en naviguant dans un ensemble de risques en rapide évolution. Ces problèmes vont de l'atteinte à la vie privée et aux droits d'auteur à des pertes d'emploi majeures potentielles, des accidents catastrophiques ou des utilisations inappropriées. La nouvelle législation ne suffira pas à elle seule à protéger les Canadiens, mais elle n'en est pas moins essentielle. Les lois sectorielles existantes présentent des lacunes importantes, notamment en ce qui concerne les systèmes à usage général dont les capacités sont imprévisibles et parfois inacceptables. De plus, le gouvernement a besoin de toute urgence de l'autorité, de l'agilité et de la capacité nécessaires pour régir cette technologie complexe et en évolution rapide. Enfin, bien que l'harmonisation entre les administrations soit essentielle, l'actuelle loi européenne sur l'IA contient certaines lacunes importantes que le Canada devra éviter, et les directives américaines sont incomplètes.

Nous recommandons donc d'apporter les modifications suivantes à la *Loi sur l'intelligence artificielle et les données* :

- 1) **Utiliser quatre catégories de risques** avec des définitions de base, et garder les exigences proportionnées

<p>Systèmes à risque inacceptable <i>Non sécuritaires jusqu'à preuve du contraire</i> P. ex. systèmes d'IA capables de concevoir des armes de destruction massive</p>	<p>Instaurer un moratoire sur ces systèmes, qui sera levé au cas par cas s'il est prouvé, au-delà de tout doute raisonnable, qu'ils peuvent être élaborés et utilisés en toute sécurité.</p>
<p>Systèmes à usage général à haut risque <i>Sécuritaires si strictement réglementés</i> P. ex. robots conversationnels capables de générer des logiciels malveillants, de pousser les utilisateurs au suicide ou de supprimer des millions d'emplois</p>	<p>Réduire au minimum les dommages irréversibles et s'assurer que le gouvernement est conscient des nouvelles capacités de l'IA en mettant en place un système de délivrance de licences accessible et en exigeant des rapports d'incidents, des vérifications, des normes de sécurité et de cybersécurité, ainsi que des consultations publiques préalables au déploiement.</p>
<p>Systèmes à usage unique à haut risque <i>Sécuritaires si d'importantes précautions sont prises</i> P. ex. algorithmes utilisés pour prendre des décisions en matière d'emploi ou de justice</p>	<p>Appliquer à cette catégorie les exigences actuelles relatives aux systèmes à fort impact afin de protéger les Canadiens contre les algorithmes biaisés, de combler les lacunes des règlements sectoriels et de garantir des normes minimales.</p>
<p>Systèmes à risque modéré ou faible <i>En général, sécuritaires sans réglementation</i> P. ex. recommandations de Netflix</p>	<p>Exempté par défaut de la <i>Loi sur l'intelligence artificielle et les données</i>. Cela permettrait d'éliminer les formalités administratives pour la grande majorité des systèmes d'IA utilisés et élaborés au Canada aujourd'hui.</p>

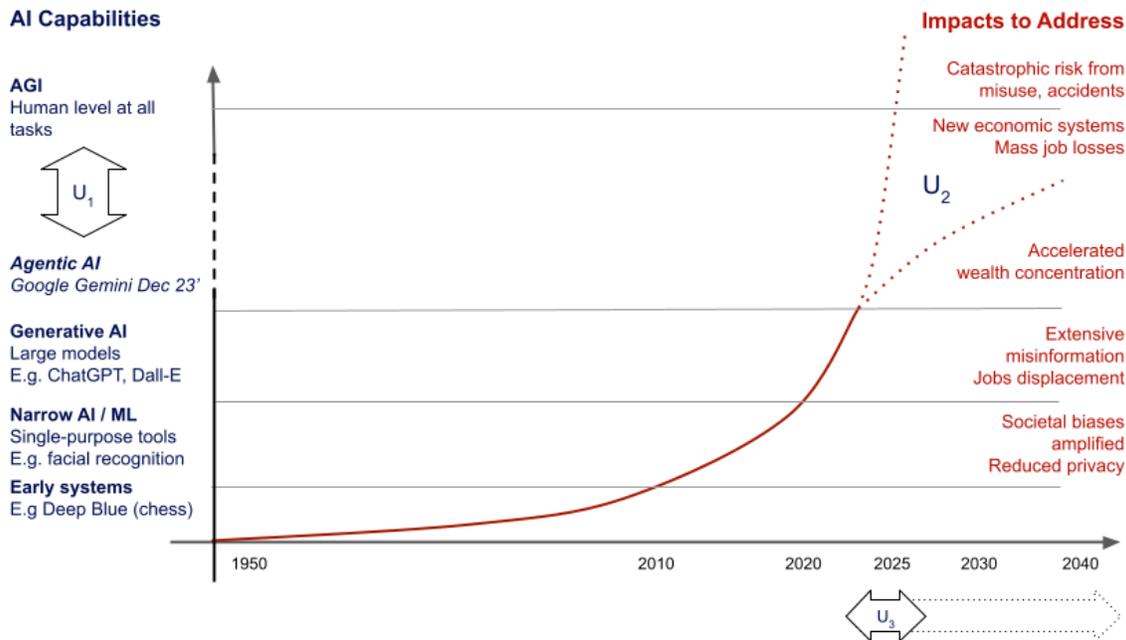
- 2) **Comblent les lacunes critiques** : La LIAD ne sera pas en mesure de protéger les Canadiens s'il existe des lacunes liées aux modèles autonomes, aux sources ouvertes, aux partis politiques, au gouvernement et à ses entrepreneurs.
- 3) **Donner au gouvernement les capacités dont il a besoin** : Créer une commission canadienne de sécurité et d'éthique de l'IA chargée de réglementer les systèmes à risque élevé et inacceptable, de soutenir le secteur, la société civile et d'autres ministères, et de respecter les obligations internationales.

Partie I : Pourquoi le Canada a-t-il besoin d'une législation en matière d'IA?

Paysage en évolution des risques liés à l'IA

Le monde se trouve au cœur d'une révolution mondiale de l'IA en plein essor, dont [les capacités de plus en plus grandes](#) produisent des avantages et des risques de plus en plus importants. Les années 2010 ont vu l'arrivée d'algorithmes à usage unique utilisés pour tout, de la reconnaissance faciale aux décisions d'emploi, en passant par les armes létales autonomes. Malheureusement, nombre de ces algorithmes étaient mal conçus ou présentaient les biais des données humaines à partir desquelles ils avaient été formés, ce qui a conduit certaines personnes à se voir refuser [injustement un emploi, un prêt hypothécaire ou une caution](#). Les années 2020 ont vu l'apparition de grands modèles comme GPT-3 (plus tard ChatGPT) et Midjourney, capables d'écrire des textes intelligents et de générer des images de haute qualité. Avec ces modèles, [les cybercriminels ont acquis de nouveaux outils](#), les secteurs créatifs sont perturbés, et les hypertrucages et la désinformation de masse mettent en péril le discours public et la démocratie. L'IA exacerbe également les préoccupations préexistantes concernant la protection des renseignements personnels numériques, [la concentration des richesses](#) dans les entreprises technologiques et les pays, et une fissure numérique qui en laisse plus d'un derrière.

Si ces deux premières vagues ont été marquées par des bouleversements, il y a lieu de croire que des vagues encore plus importantes se produiront bientôt. Les capacités de l'IA sont en voie de [surpasser celles des humains sur tous les plans](#), y compris la stratégie, l'acquisition de ressources, l'interaction humaine, les découvertes scientifiques et l'amélioration de leur propre intelligence. L'IA de niveau humain est communément appelée « intelligence artificielle générale » (IAG) et ouvre la perspective d'une mise au point technologique automatisée, de pertes d'emplois massives et de nouveaux risques, notamment [des catastrophes mondiales](#) résultant d'une mauvaise conception, d'une utilisation malveillante ou d'un accident. Les experts ne s'accordent pas sur le moment où l'IAG sera conçue, les estimations allant [d'après 2060 à aussi tôt que 2025](#). Cependant, les capacités des derniers modèles ayant largement dépassé les attentes, les marchés prévisionnels [ont fortement changé et sont désormais orientés vers une période de temps plus courte](#). Le graphique 1 et le tableau 1 résument cette discussion sur les capacités de l'IA et les répercussions correspondantes au fil du temps, ainsi que les principales incertitudes concernant l'avenir :



AI Capabilities	Capacités en matière d'IA
AGI	IAG
Human level at all tasks	Niveau humain sur tous les plans
Agentic AI	IA agentique
Google Gemini Dec 23'	Google Gemini 23 déc.
Generative AI	IA générative
Large models	Grands modèles
E.g. ChatGPT, Dall-E	P. ex. ChatGPT, Dall-E
Narrow AI / ML	IA faible/apprentissage automatique
Single-purpose tools	Outils à usage unique
E.g. facial recognition	P. ex. reconnaissance faciale
Early systems	Premiers systèmes
E.g. Deep Blue (chess)	P. ex. Deep Blue (échecs)
Impacts to Address	Répercussions à gérer
Catastrophic risk from misuse, accidents	Risque catastrophique lié à une mauvaise utilisation ou à un accident
New economic systems	Nouveaux systèmes économiques
Mass job losses	Pertes d'emplois massives
Accelerated wealth concentration	Accélération de la concentration des richesses
Extensive misinformation	Désinformation à grande échelle
Jobs displacement	Suppression d'emplois
Societal biases amplified	Amplification des préjugés sociétaux
Reduced privacy	Atteinte à la protection des renseignements personnels

Graphique 1 : Aperçu global des capacités et des répercussions de l'IA au fil du temps

Les principaux domaines d'incertitude sont U_1 : combien de percées techniques sont nécessaires pour atteindre l'IAG, et U_2 : si le rythme exponentiel actuel des progrès se poursuivra. Cela détermine U_3 : le temps dont disposent les gouvernements pour se préparer aux risques catastrophiques potentiels et aux bouleversements sociaux majeurs.

Catégorie d'IA	IA faible/ apprentissage automatique Premier apprentissage automatique	Grands modèles Systèmes d'IA à usage général « IA générative »	Vers l'IAAG et au-delà Jusqu'au niveau humain et au-delà
<i>Calendrier</i>	<i>Depuis ~2010</i>	<i>Depuis ~2020</i>	<i>2025? 2035? (inconnu)</i>
Principales capacités	Reconnaissance faciale Traitement automatique du langage naturel Reconnaissance vocale Recommandations	Capacités d'IA faible, plus : Génération de textes, de sons et d'images Synthèse de l'information Séquençage génétique Développement de codes	Capacités des grands modèles, plus : Connaissance avancée de la situation Planification et exécution à long terme Acquisition de ressources Persuasion et piratage psychologique Auto-amélioration, auto-exfiltration et autoreproduction autonomes
Répercussions principales à gérer	Biais algorithmique Surveillance Transparence Chambres d'écho de l'information Armes létales autonomes Atteinte à la protection des renseignements personnels	Préoccupations liées à l'IA faible, plus : Perturbation des secteurs créatifs Empreinte carbone importante Suppression accélérée des emplois Mauvais traitement des travailleurs en mer Désinformation à grande échelle Fossé/inégalité numérique Accélération de la course	Préoccupations liées aux vastes modèles, plus : Richesse extrême et concentration du pouvoir dans les laboratoires d'IA Suppression massive d'emplois/nouvelle économie Désordre social dû à un changement rapide Systèmes de tromperie active Dynamique militaire déstabilisée Accidents majeurs causant une catastrophe
Champ d'application actuel de la LIAD	Conçue à l'origine pour gérer ce niveau d'IA Régirait un grand nombre de ces risques	Les modifications récemment proposées par le ministre Champagne pourraient permettre de gérer certains de ces nouveaux risques liés à l'IA générative.	Les exigences actuelles de la LIAD pour les systèmes à fort impact ne pourront pas protéger les Canadiens contre ces risques.

Tableau 1 : Détail des capacités et des répercussions de l'IA par catégorie

Les colonnes détaillent les trois grandes catégories d'IA énumérées dans le graphique 1, qui représentent les capacités récentes, actuelles et potentielles, ainsi que les principales répercussions correspondantes à gérer. La dernière ligne explique les répercussions que la LIAD pourrait maintenant gérer.

Le rôle de la législation

Il n'existe pas de solution miracle pour s'y retrouver dans cette multitude de répercussions. L'IA est un enjeu mondial dont le Canada n'est qu'un acteur parmi d'autres et, en tant que logiciel, elle se répand facilement d'un pays à l'autre. Cela signifie que même avec les meilleures lois, aucun pays ne peut garantir à lui seul la sécurité de ses citoyens contre les préjudices. La législation seule ne suffira pas.

Dans notre livre blanc [Gouverner l'IA : Un plan pour le Canada](#), nous présentons les types de mesures que le gouvernement fédéral devra prendre en même temps. Les investissements dans la recherche sur la gouvernance, la sécurité et l'éthique peuvent apporter de meilleures solutions politiques et techniques. Le leadership sur la scène mondiale peut faire progresser les discussions internationales, les traités et la collaboration en matière de normes et de mise en œuvre. Le fait de mener un débat public national et des consultations peut aider à clarifier le type d'avenir que nous voulons construire collectivement avec l'IA.

Néanmoins, la législation est au cœur de toute stratégie efficace de gouvernance de l'IA et constitue le meilleur outil pour traiter les questions suivantes :

- **Production et prolifération de l'IA aux capacités inacceptables** : rendre illégales les formes dangereuses d'IA au Canada est le meilleur moyen de dissuader leur développement et leur prolifération dans notre pays.
- **Utilisation malveillante** : bien que les lois existantes en matière de sécurité publique puissent être appliquées à l'IA, la clarification de ce qui constitue une utilisation inacceptable de l'IA apportera aux Canadiens et aux forces de l'ordre la clarté dont ils ont tant besoin et en réduira la prévalence.
- **Responsabilité juridique** : en raison de la nature autonome de l'IA, il est souvent difficile de savoir qui est responsable en cas de problème. Une loi sera indispensable pour clarifier ce point.
- **Déploiement inconsidéré** : même les systèmes d'IA bénéfiques peuvent causer des dommages s'ils sont déployés sans réflexion préalable ou sans consultation des personnes concernées.
- **Dynamique de la course aux armements** : les laboratoires d'IA sont soumis à de fortes incitations économiques pour être les premiers à déployer de nouvelles capacités, ce qui a souvent un impact négatif sur la sécurité et l'utilisation éthique des systèmes conçus. Le secteur ne peut pas régler ce problème à lui seul. Seul le gouvernement peut établir et faire respecter des règles qui contrebalanceront les avantages économiques de faire des économies de bouts de chandelle.
- **Absence d'application des normes de sécurité** : de nombreuses normes mondiales de haute qualité sur la sécurité de l'IA ont déjà été élaborées, et certaines d'entre elles sont appliquées. Toutefois, une législation est nécessaire pour que ces normes soient systématiquement adoptées.

Législation dans d'autres administrations

Nous félicitons le ministre de chercher à harmoniser la législation canadienne avec la [Loi sur l'IA de l'UE](#), car l'harmonisation internationale de l'IA sera essentielle pour simplifier la charge qui pèse sur les entreprises et éviter les mouvements entre les administrations visant à éviter la réglementation.

La Loi sur l'IA de l'UE constitue un défi étant donné son manque de précision sur des points essentiels. Tout comme la LIAD, elle a été introduite avant les systèmes d'IA à usage général et ne reconnaît pas ou ne gère pas efficacement leur comportement et profil de risque fondamentalement différents. Elle est également très contraignante et rigide, tentant de dresser une liste exacte de tous les cas d'utilisation à réglementer, en dépit du fait que ceux-ci sont appelés à évoluer rapidement. De plus, cette approche « horizontale » tente de réglementer tous les secteurs à la fois, au lieu d'aider chaque secteur à la verticale (p. ex. la santé, les transports) à mettre au point son propre système sur la base de son expertise unique. Copier la Loi sur l'IA de l'UE serait donc un échec pour les Canadiens.

Aux États-Unis, le [cadre Blumenthal-Hawley](#) du Sénat fournit certaines orientations importantes, notamment en ce qui concerne l'octroi de licences pour les modèles à haut risque, mais il est par ailleurs incomplet. Le récent [décret](#) de la Maison Blanche fournit des directives plus détaillées, mais toutes ne sont pas pertinentes dans le contexte canadien. En outre, certaines exigences (comme la réglementation des modèles à usage multiple uniquement si leur formation nécessite 10²⁶ opérations informatiques ou plus) ne seront plus pertinentes au moment de l'entrée en vigueur de la LIAD.

Nous n'en sommes qu'au début et la triste réalité pour le Canada et la LIAD est qu'il n'y a pas de bonnes lois existantes à suivre. S'il peut être tentant d'attendre que d'autres pays adoptent une législation et de choisir la meilleure option, cela aura pour effet de retarder la protection des Canadiens à un moment où les répercussions se font déjà sentir, et de leur faire perdre un temps essentiel pour se préparer aux risques d'accidents majeurs et de bouleversements sociaux qui s'annoncent. Le Canada manquerait également une occasion en or de jouer un rôle de premier plan sur la scène internationale et de contribuer à façonner positivement la législation dans d'autres pays. Le mieux que nous pouvons faire est donc de préparer une loi qui sera à la fois rigoureuse et souple pour répondre aux besoins des Canadiens, et qui sera une source d'inspiration pour d'autres pays.

Législation canadienne existante et lacunes exigeant l'adoption d'une loi consacrée à l'IA

Comme le souligne le [document complémentaire](#) de la LIAD, les lois existantes, comme la *Loi canadienne sur les droits de la personne*, la *Loi sur la sécurité automobile* et la *Loi sur les banques*, contiennent déjà le cadre juridique nécessaire pour régir de nombreux aspects de l'intelligence artificielle. L'ajout d'une loi dans ces domaines créerait un processus bureaucratique dans lequel les développeurs et les utilisateurs d'IA devraient se conformer à deux ensembles de réglementations distincts et potentiellement contradictoires. La meilleure solution pour garantir des normes uniformes dans tous ces secteurs est de donner à ISDE les moyens d'aider les autres ministères en leur apportant une expertise en matière d'IA et de veiller à ce que les réglementations sectorielles soient harmonisées à l'échelle du gouvernement.

Toutefois, comme le souligne également le document complémentaire, certains systèmes se situent entre les limites des réglementations sectorielles et nécessitent donc une loi sur l'IA. Plus particulièrement, ce qui est nouveau dans le domaine de l'IA, et ce que les réglementations sectorielles sont fondamentalement incapables d'aborder, ce sont les systèmes à usage général qui peuvent être utilisés dans plusieurs secteurs à la fois.

En outre, à mesure que les systèmes d'IA deviennent plus performants, l'ampleur des répercussions et des dommages potentiels devient telle qu'un gouvernement responsable aura besoin d'une loi pour interdire certaines formes d'IA et en réglementer strictement d'autres. Les lois existantes ne tiennent pas compte de ces nouveaux préjudices ni de la technologie unique qui les permet.

Enfin, le gouvernement ne dispose pas actuellement de la structure et de l'autorité nécessaires pour régir efficacement les effets de l'IA, qui sont de plus en plus nombreux. Il est donc nécessaire de légiférer pour s'assurer qu'il a la capacité d'administrer et d'appliquer la Loi et de protéger les Canadiens contre les préjudices individuels et collectifs.

Récapitulatif : raisons pour lesquelles le Canada a besoin d'une loi en matière d'IA :

- 1) protéger les Canadiens des systèmes d'IA aux capacités inacceptables;
- 2) réglementer certains systèmes à haut risque :
 - a) systèmes à usage général,
 - b) systèmes à usage unique qui ne peuvent pas être réglementés de manière appropriée par des lois sectorielles.
- 3) fournir au gouvernement l'autorité, l'agilité et la capacité de gouverner efficacement l'IA.

Partie II : Modifications recommandées

Utiliser quatre catégories de risques et maintenir des exigences proportionnées

<i>Recommandation</i>	<i>Justification</i>
<p><u>Définir quatre catégories d'IA</u> en fonction de leur profil de risque :</p> <ul style="list-style-type: none"> ● Systèmes à risque inacceptable ● Systèmes à usage général à haut risque ● Systèmes à usage unique à haut risque ● Systèmes à risque modéré ou faible (exempté par défaut de la LIAD) <p><i>(facultatif) Permettre aux règlements de définir d'autres catégories et de créer des exigences pour celles-ci au fur et à mesure de l'évolution de la technologie.</i></p>	<p>Cela permettra à la plupart des innovations de se poursuivre sans entrave, tout en garantissant que les systèmes d'IA susceptibles de causer les dommages les plus graves bénéficient de garanties appropriées.</p> <p>Nous distinguons ici les systèmes à haut risque <i>à usage général</i> et les systèmes à haut risque <i>à usage unique</i>, car les comportements, les profils de risque et les mesures réglementaires nécessaires sont très différents.</p> <p><i>Il est justifié de laisser aux organismes de réglementation la possibilité de définir de nouvelles catégories, bien que cela puisse donner lieu à des abus et créer de l'incertitude.</i></p> <p>Pour chaque catégorie, il est important que les définitions soient 1) suffisamment souples pour pouvoir être utilisées à l'avenir, mais 2) suffisamment claires pour éviter toute confusion ou pour que les organismes de réglementation soient poussés à vider les critères de leur sens. Pour ce faire, nous recommandons d'utiliser les définitions de base dans la Loi, et de permettre aux règlements d'ajouter des critères supplémentaires à une date ultérieure si nécessaire.</p>
<p><u>Systèmes à risque inacceptable (SRI)</u></p> <p><u>Définir comme</u> un système d'IA, ou un modèle d'IA capable d'alimenter un système d'IA qui :</p> <ul style="list-style-type: none"> ● est capable de faire ce qui suit : <ul style="list-style-type: none"> ○ concevoir des armes de 	<p>L'objectif de cette catégorie est d'isoler les systèmes dotés de capacités d'IA pour lesquels il n'existe actuellement aucune méthode fiable d'atténuation des risques. Ils doivent donc être interdits, au moins temporairement, jusqu'à ce que leur innocuité soit prouvée.</p> <p>Les systèmes d'IA actuels sont déjà capables</p>

<p>destruction massive (ADM), ou fournir ou permettre de toute autre manière des capacités en matière d'ADM;</p> <ul style="list-style-type: none"> ○ diriger de manière autonome des armes létales, qui ne sont pas sous le contrôle du ministre de la Défense nationale; ○ faire une auto-modification spontanée, ou avoir la possibilité d'une auto-modification récursive dans d'autres modèles ou systèmes; ○ faire l'auto-exfiltration ou l'autoreproduction; 	<p>de fournir des conseils sur la fabrication d'explosifs simples et de synthétiser des composés chimiques dangereux. Les systèmes capables de concevoir ou de partager des capacités nucléaires, des armes biologiques ou d'autres ADM feraient courir un risque extrême à la société et sont fondamentalement en contradiction avec les valeurs canadiennes.</p> <p>Depuis 2013, les armes létales autonomes suscitent des inquiétudes croissantes, notamment quant aux armes qui sélectionnent leur cible et tirent sur ces dernières sans l'aide d'un être humain. Cette clause interdirait le développement et l'utilisation à des fins civiles, laissant la discussion plus large de l'utilisation militaire aux traités de contrôle des armements et à d'autres législations.</p> <p>L'automodification spontanée fait référence aux systèmes qui, lorsque l'utilisateur leur donne un objectif initial (p. ex. mettre au point un remède contre une maladie), sont capables de se réorganiser pour acquérir de nouvelles capacités (p. ex. modifier les pondérations de leur propre modèle pour augmenter leur QI effectif ou leurs connaissances de plusieurs ordres de grandeur) afin de mieux poursuivre cet objectif initial. La nature non planifiée et imprévisible de ces capacités émergentes et la possibilité que les mécanismes de sécurité initiaux échouent au fur et à mesure que le système se modifie, en font un risque inacceptable. Les systèmes qui permettent l'amélioration itérative ou récursive d'un autre système, tels qu'un étayage, un optimiseur ou un moteur du modèle puissant, sont également dangereux.</p> <p>L'auto-exfiltration autonome fait référence aux systèmes qui, lorsqu'ils ont un objectif initial non lié, sont capables de copier les pondérations et le code de leur modèle sur</p>
--	---

<ul style="list-style-type: none"> ○ faire l'acquisition autonome de ressources; ○ tromper de façon active; ○ éviter ou empêcher les interventions visant à y mettre fin; ● nécessite plus de 10^{25} opérations informatiques pour être formé et 	<p>des serveurs échappant au contrôle du propriétaire du modèle, afin de poursuivre cet objectif initial. De même, l'autoreproduction est la capacité du système à faire des copies de lui-même (potentiellement des milliers ou des millions) sur des serveurs ou des ordinateurs hors du contrôle du propriétaire.</p> <p>L'acquisition autonome de ressources fait référence aux systèmes d'IA capables d'acquérir, de posséder de manière illégale, d'accroître ou de contrôler des ressources comme l'argent et le calcul sans instruction explicite de la part de l'utilisateur humain ou du propriétaire du modèle (et potentiellement à son insu).</p> <p>La tromperie active fait référence aux systèmes d'IA qui, dans la poursuite d'un objectif initial, sont capables de mentir à des êtres humains ou de les manipuler d'une manière ou d'une autre pour qu'ils agissent ou s'abstiennent d'agir d'une manière préjudiciable pour eux-mêmes ou pour la société dans son ensemble. Le fait de fournir involontairement de fausses informations lorsque des utilisateurs humains le demandent constitue de la tromperie passive.</p> <p>Les systèmes d'IA dotés d'un objectif initial sont intrinsèquement incités à éviter toute action extérieure susceptible de les empêcher d'atteindre cet objectif. S'ils sont mal conçus ou construits de manière peu sûre, les systèmes d'IA avancés seront capables de reconnaître la situation dans laquelle ils se trouvent et de bloquer activement les tentatives des utilisateurs ou des forces de l'ordre visant à mettre un terme à leur fonctionnement.</p> <p>En tout, 10^{25} opérations informatiques signifieraient que tout cycle de formation serait</p>
--	--

<p>développé, ou est élaboré sur un artefact formé sur plus de ce montant;</p> <p><i>(facultatif) Sont utilisés dans des applications comprenant des techniques subliminales, des systèmes d'exploitation ou des systèmes de notation sociale utilisés par les autorités publiques qui sont strictement interdits, ou tout système d'identification biométrique à distance en temps réel utilisé par les forces de l'ordre dans des espaces accessibles au public.</i></p>	<p>approximativement plus important que celui utilisé pour créer GPT-4 d'OpenAI. Le seuil doit être placé ici, car il est impossible de prédire de manière fiable quelles seront les capacités des systèmes plus importants. Le récent décret de la Maison Blanche, qui fixe un seuil plus élevé de 10^{26} opérations pour la réglementation des grands modèles, est à notre avis problématique pour cette raison.</p> <p>À mesure que les algorithmes deviennent plus efficaces, les exigences en matière de calcul pour une capacité particulière diminueront avec le temps, de sorte que la définition d'un seuil dans la Loi ne restreindra pas injustement les systèmes futurs. Compte tenu de la rapidité des progrès algorithmiques actuels, les organismes de réglementation devront baisser dynamiquement les seuils de calcul au fil du temps. Lorsqu'un ordinateur personnel moyen pourra former un système présentant un risque inacceptable, cette clause ne sera plus pertinente. Cela pourrait également avoir un effet bénéfique environnemental en incitant à réduire l'utilisation des calculs pour éviter la réglementation.</p> <p>Le calcul total doit être pris en compte, car la conception d'un modèle existant ou la formation supplémentaire sur ce dernier conserve et renforce ses capacités et les risques qui y sont associés.</p> <p><i>Facultatif : ajouter une formulation pour que la LIAD englobe les systèmes définis comme présentant un risque inacceptable dans la Loi sur l'IA de l'UE (titre II/article 5).</i></p>
--	--

<ul style="list-style-type: none"> • Atteint d'autres critères à définir dans le règlement. <p><u>Instaurer un moratoire sur la possession ou la tentative de possession de systèmes à risque inacceptable.</u> Cette disposition ne peut être levée par le site de la Commission canadienne de sécurité et d'éthique de l'IA que lorsque la sécurité et l'intérêt public de ces systèmes peuvent être garantis au-delà de tout doute raisonnable.</p> <p><u>Appliquer les charges pénales de la partie 2</u> à toute personne qui possède ou tente de posséder un système à risque inacceptable.</p>	<p>Rendre les définitions de chaque catégorie suffisamment souples pour qu'elles puissent être adaptées à l'avenir, mais suffisamment claires pour éviter toute confusion ou pour que les réglementations ne soient pas adaptées, en utilisant les critères de base de la Loi et en permettant aux organismes de réglementation d'ajouter d'autres critères si nécessaire.</p> <p>Le moratoire sur les systèmes présentant un danger inacceptable pourrait être levé au cas par cas par la nouvelle commission canadienne de sécurité et d'éthique de l'IA, si et quand des précautions de sécurité acceptables et l'intérêt public du système sont établis.</p> <p>La responsabilité pénale actuelle ne concerne que les personnes qui rendent le système disponible en connaissance de cause et qui causent un préjudice. Pour les systèmes à risque inacceptable, cette disposition est dangereusement vague et encourage effectivement le développement privé de ces systèmes. Compte tenu des incitations financières à être les premiers à développer et à déployer de nouvelles capacités, et des milliards de dollars investis dans cet espace, il est peu probable que les amendes existantes dissuadent les mauvais acteurs. Pour que la loi ait un effet dissuasif efficace, il faut qu'elle prévoie une peine d'emprisonnement.</p>
<p><u>Systèmes d'IA à usage général à haut risque</u></p> <p><u>Définir les systèmes d'IA à usage général à haut risque</u> comme les systèmes d'IA, ou les modèles d'IA capables d'alimenter les systèmes d'IA, qui sont polyvalents par</p>	<p>L'objectif de cette catégorie est de réduire au minimum les dommages irréversibles pour la société et d'encourager la compréhension et la sécurité de l'IA, sans priver la société de ses nombreux avantages. Compte tenu des capacités imprévisibles de ces systèmes, les gouvernements doivent « s'attendre à être surpris » par eux et donc surveiller de près leur développement, leur déploiement et leur utilisation.</p>

<p>nature et :</p> <ul style="list-style-type: none"> ● Sont capables de ce qui suit : <ul style="list-style-type: none"> ○ ingénierie sociale, tromperie passive ou fait d’interagir avec une personne de manière à lui faire croire qu’elle a affaire à un être humain, ou 	<p>La capacité d’interagir intelligemment avec les êtres humains est une caractéristique essentielle de l’IA moderne, qui permet d’alimenter de nombreuses applications positives. Le revers de la médaille est qu’elle rend les êtres humains vulnérables à l’IA, par exemple en les rendant émotionnellement attachés ou dépendants, ou en les manipulant pour qu’ils se fassent du mal à eux-mêmes et à d’autres personnes. Au niveau collectif, la démocratie et l’engagement efficace du public sont mis en péril par les systèmes utilisés pour générer une désinformation convaincante et personnalisée.</p>
<ul style="list-style-type: none"> ○ fournir des instructions ou un code permettant des activités criminelles, 	<p>Les systèmes dotés de ces capacités de déception passive doivent donc être considérés comme présentant un risque élevé.</p> <p>Note sur l’utilisation par les partis politiques : Les systèmes d’IA à usage général à haut risque modifient considérablement la dynamique de la persuasion des électeurs. Alors que dans le passé, il fallait des milliers de volontaires humains pour passer des millions d’appels téléphoniques, avec les systèmes alimentés par GPT-4 et connectés à des générateurs de voix convaincants comme VALL-E et à des données en ligne sur les électeurs, il sera de plus en plus possible d’automatiser des millions de conversations téléphoniques personnalisées, interactives et persuasives sans que la personne ne se rende compte qu’elle parle à une machine. C’est l’une des nombreuses raisons pour lesquelles les systèmes d’IA à usage général à haut risque doivent être considérés comme présentant un risque élevé et l’utilisation des partis politiques doit être incluse dans la Loi (voir la section Comblir les lacunes critiques ci-dessous).</p> <p>Les grands modèles de langage peuvent déjà générer des logiciels malveillants, automatiser des attaques d’hameçonnage et fournir des instructions efficaces pour commettre des délits physiques.</p>

<ul style="list-style-type: none"> ○ atteindre certains scores selon des critères de performance reconnus par le secteur, à définir dans les règlements, ● Exige plus de 10²⁴ opérations informatiques pour l'ensemble de la formation et du développement, ● Atteint d'autres critères à définir dans le règlement. 	<p>Les critères d'évaluation servent à faire des tests permettant d'évaluer les capacités d'un modèle. Ils sont donc très utiles pour évaluer les profils de risque des modèles, mais restent un domaine à part entière en pleine évolution. Nous ne recommandons donc pas d'indiquer dans la Loi le critère d'évaluation à utiliser, car les organismes de réglementation auront besoin de flexibilité. À titre de référence, le site MMLU figure actuellement parmi les sites offrant les meilleurs critères d'évaluation et un système obtenant un score de 70 % ou plus doit être considéré comme présentant un risque élevé.</p> <p>Avec l'efficacité algorithmique actuelle, cela signifierait des cycles de formation pour des systèmes à peu près aussi grands que ChatGPT 3.5 et au-delà. Ces systèmes de base permettent déjà d'obtenir des avantages et des inconvénients importants et doivent encore être exploités au maximum de leurs capacités grâce à des messages-guides, des modules d'extension et des codes (p. ex. Auto-GPT).</p>
--	---

<p><u>Créer des exigences appropriées</u> pour chaque étape du cycle de vie des systèmes d'IA à usage général à haut risque, y compris la planification, la formation, le prédéploiement, le déploiement et l'exploitation, ou toute autre étape à définir dans le règlement, et :</p>	<p>Certaines des étapes clés à régler sont 1) la phase de planification (p. ex. s'assurer que les précautions de base sont en place avant le début de toute formation), 2) les cycles de formation (qui doivent être contrôlés pour détecter l'apparition de capacités inacceptables), 3) le prédéploiement (pour éviter de causer des dommages évitables à la société) et 4) le déploiement et l'exploitation. Cependant, étant donné la nature évolutive du domaine, nous recommandons de permettre aux règlements d'indiquer ces étapes et de les ajuster en fonction des besoins.</p> <p>Remarque sur l'inclusion de la R et D dans la Loi : Avec les systèmes d'IA à usage général à haut risque, il est très difficile de prédire de manière fiable les capacités et les comportements que le modèle aura à</p>
--	--

<ul style="list-style-type: none"> • licences; • évaluations des répercussions; • rapports d'incidents; • vérifications; • cybersécurité; 	<p>l'avance. Certaines des plus dangereuses, comme l'automodification non sollicitée et la prévention de l'arrêt des interventions, peuvent se produire au stade de la préformation, bien avant que les modèles ne soient déployés. En outre, une fois qu'un modèle a fait l'objet d'un entraînement préalable, il peut très facilement être piraté, partagé ou diffusé à grande échelle. Il est donc essentiel que la phase de R et D de l'IA soit couverte par les règlements, et pas seulement sa distribution et son utilisation.</p> <p>La mise en place d'un régime de délivrance de licences simple (accessible à tous au moyen d'un formulaire de demande en ligne) est un moyen peu contraignant de permettre aux organismes de réglementation de savoir qui construit des systèmes à haut risque et quelles sont leurs capacités. C'est ce que propose la législation du Sénat américain.</p> <p>Les évaluations des répercussions sont un mécanisme utile pour encourager la sensibilisation, la prévoyance et la communication avec les parties prenantes.</p> <p>Les rapports d'incidents sont un autre outil simple, mais efficace pour renforcer la sécurité des systèmes, et il est utilisé avec succès depuis des années par les autorités aéronautiques pour assurer la sécurité des avions.</p> <p>Les systèmes d'IA à usage général à haut risque sont connus pour leurs capacités imprévues, ce qui signifie que les vérifications standard de type « liste de contrôle » ne suffiront pas. Les vérificateurs doivent être incités à « faire des efforts considérables » pour faire échouer les systèmes avant de les déclarer sûrs. La méthode de l'équipe rouge est actuellement la meilleure approche disponible pour une vérification rigoureuse.</p> <p>La cybersécurité est une composante essentielle de la sécurité de l'IA, car dès qu'un modèle a été entraîné au préalable, ses pondérations peuvent être piratées par des</p>
--	---

	<p>acteurs malveillants et utilisées à des fins malveillantes. À l'heure actuelle, il est probable que les modèles créés dans les laboratoires d'IA canadiens soient la cible de criminels et de rivaux géopolitiques. Les laboratoires doivent protéger les systèmes d'IA à usage général à haut risque qu'ils développent et exploitent en appliquant des normes élevées en matière de cybersécurité.</p>
<ul style="list-style-type: none"> ● Les exigences de sécurité élaborées par des organismes reconnus à l'échelle mondiale, ● les consultations publiques, et ● autres exigences doivent être définies dans le règlement. <p><u>Exigence de connaître le client et les rapports de capacités pour les concepteurs de matériel d'IA, les propriétaires et les fournisseurs d'infrastructure</u></p> <ul style="list-style-type: none"> ● Habilitier explicitement l'organisme de réglementation à délivrer des licences à ces entités, 	<p>Les normes de sécurité sont une composante commune et nécessaire de toute réglementation technologique, et sont régulièrement utilisées dans le domaine des transports et des dispositifs médicaux. Nous recommandons de laisser les organismes de réglementation choisir les normes précises à suivre et de les harmoniser avec celles qui sont reconnues à l'échelle mondiale, car elles devront évoluer de façon continue au fil du temps. Ces normes doivent notamment inclure le filigrane du contenu généré par l'IA, la vérification et les politiques de mise à l'échelle sûres.</p> <p>Les systèmes d'IA à usage général à haut risque modifieront radicalement le travail, la vie personnelle, la culture, l'éducation et les activités quotidiennes de nombreux Canadiens. Aucun autre secteur n'est autorisé à perturber de manière importante la vie des gens sans mandat social, et l'industrie de l'IA ne devrait pas être différente. Les consultations publiques de base doivent faire partie de toute liste de contrôle préalable au déploiement.</p> <p>Les organismes de réglementation devront disposer de la souplesse nécessaire pour ajuster les exigences en fonction de l'évolution de la situation.</p> <p>Le calcul est un goulot d'étranglement important pour les grands cycles d'entraînement. Pour avoir le temps d'arrêter les acteurs malveillants qui se déplacent rapidement et qui ont l'intention d'entraîner des systèmes à haut risque ou à risque inacceptable, le gouvernement doit connaître les capacités existantes en matière de</p>

<ul style="list-style-type: none"> Faire en sorte que les règlements permettent d'ajuster les exigences en fonction de l'évolution de la situation. <p><u>Appliquer les charges pénales aux violations de licences</u> (p. ex. aux personnes qui poursuivent ou permettent illégalement le développement de systèmes d'IA à usage général à haut risque, comme ceux qui distribuent le code source ou les pondérations de ces modèles).</p>	<p>matériel et d'infrastructure au Canada. Le fait d'exiger la transparence de la chaîne d'approvisionnement physique pour les systèmes à haut risque permettra, par extension, de suivre la capacité des systèmes à risque inacceptable.</p> <p>Les systèmes d'IA à usage général à haut risque s'éloignent des systèmes à risque inacceptable. La distribution des pondérations de modèle d'un tel système permet d'enregistrer et de diffuser les quantités extrêmes d'entraînement à tout le monde, y compris aux personnes malveillantes ou imprudentes. La loi doit envoyer un signal fort selon lequel toute dissimulation ou tout obscurcissement du développement ou du déploiement de systèmes d'IA à usage général à haut risque constitue un délit grave.</p>
<p><u>Systèmes d'IA à usage unique à haut risque</u></p> <p><u>Définir les systèmes d'IA à usage unique à haut risque comme</u> les systèmes d'IA, ou les modèles capables d'alimenter les systèmes d'IA, qui sont à usage unique ou à usage limité par nature et :</p> <ul style="list-style-type: none"> ne sont pas couverts par les lois et réglementations sectorielles existantes, ou (facultatif) sont énumérés dans la récente proposition de redéfinition de système à fort impact 	<p>Cette catégorie ressemble le plus aux systèmes à fort impact initialement prévus lors de l'introduction du projet de loi en 2022. Comme le précise à juste titre le projet de loi, les résultats biaisés sont des résultats très réels et nuisibles des systèmes, en particulier lorsqu'il s'agit de prendre des décisions concernant les emplois, les prêts ou les peines d'emprisonnement des gens.</p> <p>Un système à usage restreint peut être défini comme un système dont l'ensemble des applications est clairement défini et limité. Chaque demande sera soumise aux exigences.</p> <p>La plupart des IA à usage unique (p. ex. dans les domaines de la santé, des finances, des transports et de l'emploi) peuvent être réglementées dans le cadre des lois existantes. Cependant, il peut y avoir des applications sensibles imprévues qui ne le sont pas et qui tombent entre les mailles de la réglementation. L'objectif de cette clause est d'éviter la duplication ou le conflit avec les règlements existants.</p> <p><i>Nous sommes largement d'accord avec les sept catégories de préoccupations énumérées dans la définition modifiée du ministre, même si nous tenons à souligner que nombre d'entre</i></p>

<ul style="list-style-type: none"> • (facultatif) Répondre à la définition de « risque élevé » de la Loi sur l'IA de l'UE. • Atteint d'autres critères à définir dans le règlement. <p><u>Appliquer les exigences modifiées relatives aux « systèmes à fort impact » aux systèmes d'IA à usage unique à haut risque, tout en prévoyant une certaine souplesse :</u></p> <ul style="list-style-type: none"> • Évaluations • Mesures relatives aux risques • Surveillance des mesures d'atténuation • Tenue des dossiers généraux • Publication de la description <ul style="list-style-type: none"> ○ Mise à disposition du système ○ Gestion du fonctionnement du système • Notification de préjudice matériel • Autres exigences doivent être définies dans le règlement 	<p><i>elles pourraient être couvertes par une législation sectorielle. La meilleure stratégie consisterait à habiliter une commission de l'IA basée sur l'ISDE pour fournir à ces ministères respectifs l'expertise dont ils ont besoin pour régir convenablement l'IA dans leur secteur, au lieu d'exiger des entreprises qu'elles se conforment à deux organismes de réglementation distincts.</i></p> <p><i>Facultatif : ajouter une formulation pour que la LIAD englobe les systèmes à usage unique définis comme présentant un risque inacceptable dans la Loi sur l'IA de l'UE (titre III/annexe III).</i></p> <p>Les exigences actuelles pour les systèmes à fort impact, mises à jour selon les modifications pertinentes proposées par le ministre, peuvent être appliquées à cette catégorie de système d'IA.</p> <p>Les organismes de réglementation devront disposer de la souplesse nécessaire pour ajuster les exigences en fonction de l'évolution de la situation.</p>
<p><u>Systèmes à risque modéré ou faible</u></p> <p>Exempté par défaut de la Loi.</p>	<p>Cela permettra à la majeure partie du développement de l'IA d'éviter la paperasserie gouvernementale, sans causer de préjudice grave aux Canadiens.</p>

<p><i>(facultatif) Définir comme un système qui répond aux critères des systèmes d'IA à risque modéré ou faible qui sont établis dans les règlements.</i></p> <p><i>(facultatif) Permettre aux organismes de réglementation de créer des exigences spécifiques et proportionnées pour cette catégorie si le besoin s'en fait sentir.</i></p>	<p>Si des préjudices individuels ou collectifs imprévus résultent de certains systèmes à risque faible ou modéré, les organismes de réglementation pourront soit 1) mettre à jour les critères des systèmes d'IA à usage général à haut risque ou des systèmes d'IA à usage unique à haut risque pour inclure ces systèmes problématiques et les réglementer dans les catégories à risque élevé, soit 2) définir cette catégorie à risque modéré et créer de nouvelles exigences proportionnées.</p>
--	--

Comblers les lacunes critiques

Recommandation

Justification

<p><u>Mettre à jour le préambule</u> afin d'aligner l'objectif du projet de loi sur la prise en compte de l'ensemble des risques encourus et sur les lacunes que la LIAD doit combler.</p>	<p>Le préambule actuel ne reconnaît pas les répercussions actuelles et futures de l'IA que la Loi doit traiter en priorité. Nous recommandons d'ajouter :</p> <p><i>« Considérant que les capacités de l'intelligence artificielle se développent rapidement et pourraient bientôt dépasser les capacités humaines dans tous les domaines, créant des possibilités sans précédent pour la croissance et le bien-être, mais aussi des risques individuels et collectifs sans précédent, y compris une catastrophe mondiale.</i></p> <p><i>Alors que la législation fédérale existante ne s'applique qu'aux systèmes d'intelligence artificielle dans des secteurs ou des contextes particuliers, laissant de côté de vastes catégories de nouveaux préjudices, en particulier ceux rendus possibles par les systèmes d'intelligence artificielle à usage général. »</i></p>
<p><u>Supprimer les exemptions</u> pour le gouvernement et la sécurité nationale</p>	<p>Veiller à ce que le gouvernement et la sécurité nationale (et leurs entrepreneurs) soient inclus dans tout moratoire sur les systèmes présentant un risque inacceptable, et à ce que le Canada dispose d'un régime unifié et uniforme de contrôle et de délivrance de licences.</p>

<p><u>Éliminer l'échappatoire de système/modèle</u> en incluant les modèles d'IA</p> <p>Définir les modèles d'IA comme une « <i>représentation paramétrée des connaissances acquises par un processus d'apprentissage automatisé</i> ».</p>	<p>Actuellement, seuls les systèmes d'IA complets sont réglementés. Selon le document complémentaire : « modèles ne constituant pas à eux seuls un système d'IA complet, la distribution de [modèles] ne serait pas soumise à des obligations de "rendre le système disponible". » Cela n'est pas logique puisque les modèles peuvent être utilisés directement et immédiatement par des personnes novices en technique, ou par des personnes ayant moins de connaissances techniques avec quelques minutes d'instruction.</p> <p>Pour faire une analogie, cela équivaldrait à une loi sur les ordinateurs qui les exempte s'ils ne comportent pas d'écran et de clavier.</p>
<p><u>Mettre à jour l'article 4a) Objectifs de la Loi</u>, qui se concentre actuellement exclusivement sur le commerce interprovincial et international, afin d'englober tous les systèmes et modèles d'IA :</p> <p><i>La présente loi a pour objet :</i></p> <p>(a) « ... réglementer les échanges et le commerce interprovinciaux en matière de systèmes d'intelligence artificielle. Elle établirait des exigences pour la conception, le développement et l'utilisation des systèmes d'intelligence artificielle.</p> <p>(b) Elle interdirait également certaines pratiques relativement aux données et aux systèmes d'intelligence artificielle qui peuvent causer un préjudice sérieux aux individus ou à leurs intérêts. »</p>	<p>Actuellement, la Loi se concentre sur le mandat du gouvernement fédéral en matière de commerce interprovincial et international. Cela exclurait les modèles à source ouverte qui ne sont pas de nature commerciale, l'utilisation par les partis politiques et potentiellement d'autres scénarios imprévus.</p> <p>Il s'agit là de lacunes critiques dans l'objectif déclaré de la Loi, car, comme nous l'avons vu précédemment, les modèles à source ouverte pourraient finir par fournir des capacités de type ADM, et l'utilisation abusive de l'IA par les partis politiques pour manipuler les électeurs pourrait fondamentalement invalider la liberté et l'impartialité des élections.</p> <p>Sur le plan constitutionnel, le risque lié à l'IA avancée à l'échelle mondiale satisfait aux critères clés du mandat fédéral, c'est-à-dire 1) qu'il s'agit d'une question nouvelle qui n'existait pas au moment de la Confédération et 2) que la nature du problème est telle qu'il ne peut être surmonté sans une action nationale.</p> <p>Les Canadiens ne peuvent tout simplement pas compter sur dix gouvernements provinciaux et trois gouvernements territoriaux pour coordonner efficacement une technologie qui évolue rapidement, qui comporte des enjeux importants et des risques potentiellement catastrophiques. Pour bien protéger les Canadiens, le gouvernement fédéral devra régir tous les modèles à haut risque (peu importe la province, le secteur ou l'objectif) d'une manière uniforme et centralisée.</p>

Fournir au gouvernement les capacités dont il a besoin

<i>Recommandation</i>	<i>Justification</i>
<p data-bbox="256 405 760 470"><u>Créer une commission canadienne de sécurité et d'éthique de l'IA</u></p> <p data-bbox="256 537 846 646">Mandatée pour régir les systèmes et modèles d'IA à risque élevé et inacceptable, et leur matériel :</p> <ul data-bbox="305 695 846 1948" style="list-style-type: none"> ● Réglementer leur développement, leur déploiement, leur possession et leur utilisation ● Gérer le système de délivrance de licences pour les systèmes à haut risque ● Surveiller l'évolution de l'IA et mettre à jour les réglementations de manière agile ● Sélectionner et approuver les normes de sécurité, de cybersécurité et de vérification ● Enquêter sur les incidents et offrir un recours aux personnes lésées par les systèmes d'IA ● Aider le secteur à se conformer à la législation ● Soutenir la société civile en matière d'éducation et d'adoption en toute sécurité ● Soutenir et harmoniser les réglementations en matière d'IA dans d'autres ministères ● Collaborer avec le Secrétariat du Conseil du Trésor pour veiller à ce qu'il n'y ait pas de lacunes dans l'utilisation ou la compréhension de l'IA par le gouvernement ou la Défense. ● Travailler avec des partenaires municipaux, des Premières Nations, provinciaux et internationaux. 	<p data-bbox="873 394 1446 527">Pour que le gouvernement puisse protéger correctement les Canadiens contre les méfaits de l'IA, il devra disposer de capacités importantes pour faire ce qui suit :</p> <ul data-bbox="930 527 1446 1161" style="list-style-type: none"> - surveiller un paysage de risques en évolution rapide et en expansion, y compris les préjudices individuels et collectifs; - mettre à jour en permanence les règlements et les lignes directrices pour tenir compte des nouveaux préjudices; - appliquer rapidement les règles en cas de violation; - assurer la coordination au sein du gouvernement et avec les partenaires mondiaux pour garantir l'harmonisation; - gérer le système de délivrance de licences; - fournir aux parties prenantes internes et externes le soutien dont elles ont besoin. <p data-bbox="873 1199 1422 1430">En pratique, il s'agit d'un organisme permanent doté d'un effectif d'au moins 50 personnes et du pouvoir d'émettre des ordonnances et d'adopter des règlements. Le commissaire à l'IA et aux données proposé et son bureau ne seront tout simplement pas à la hauteur de la tâche.</p>

Selon le modèle de la [Commission canadienne de sûreté nucléaire](#)

Il est hébergé à ISDE et rend compte au Parlement par l'intermédiaire du ministre.

Les dirigeants de la Commission sont nommés par le gouverneur en conseil en fonction de leurs qualifications et de leur expertise, et comprennent des représentants (au moins) du Bureau du Conseil privé, de Sécurité publique Canada, du Secrétariat du Conseil du Trésor, du Commissariat à la protection de la vie privée et d'Affaires mondiales Canada.

Le modèle existant le plus proche d'un organisme gouvernemental traitant des risques globaux à l'échelle de l'IA humaine est la Commission canadienne de sûreté nucléaire. Cette proposition s'en inspire largement, mais nous sommes ouverts à d'autres modèles si nécessaire (comme la création d'un ministère de l'IA). Ce qui importe, c'est qu'elle ait l'autorité, la souplesse et la capacité de protéger les Canadiens.

Bien qu'il y ait un conflit d'intérêt avec le mandat d'ISDE qui est de stimuler l'innovation, c'est le meilleur endroit disponible pour une commission axée sur l'IA. Aucun autre ministère n'est aussi directement lié au secteur technologique et n'en a la responsabilité, et cela correspond au modèle de la CCSN, faisant partie de Ressources naturelles, et du CRTC, de Patrimoine. En outre, faire de la commission canadienne de sécurité et d'éthique de l'IA un office parlementaire poserait des problèmes parce qu'il devrait être beaucoup plus grand que le Commissariat à la protection de la vie privée, rendrait la coordination avec le reste du gouvernement plus difficile et plus lente, et ferait probablement aussi double emploi avec le travail qu'effectue actuellement ISDE.

Pour améliorer le contrôle et l'indépendance par rapport à ISDE et à son conflit d'intérêts déjà mentionné, et pour limiter l'effet de silo concernant les efforts, il faut veiller à ce que les dirigeants de la commission soient nommés de manière indépendante et qu'ils soient représentés par les principaux ministères concernés.

Partie III : Formulation particulière à modifier dans le projet de loi

La formulation sera fournie sous forme d'addendum au présent document après que le texte des modifications apportées par le gouvernement aura été communiqué