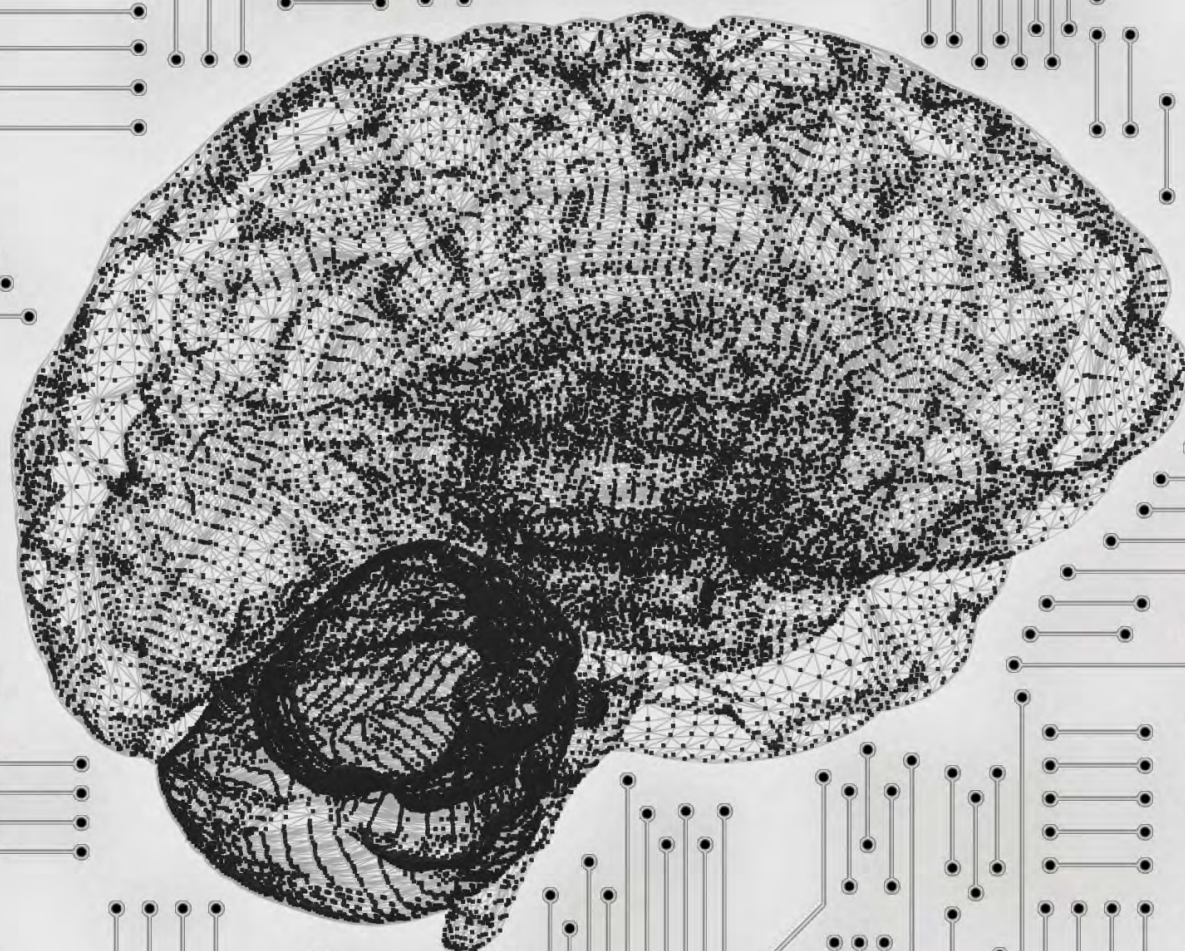
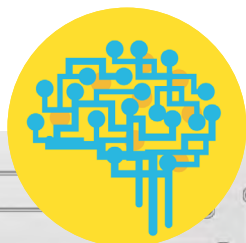


LA GOUVERNANCE DE L'IA :

Risques, incertitudes et avantages

Pour encourager l'innovation dans le domaine de l'intelligence artificielle (IA) tout en réduisant les risques autant que possible, le Canada devrait adopter une approche progressive de la gouvernance de l'IA qui soit axée sur la gestion des risques en s'appuyant sur deux nouveaux organismes consultatifs. [Par Daniel Munro](#)



RÉSUMÉ ET RECOMMANDATIONS

L'intelligence artificielle, ou IA — la capacité qu'ont les machines d'effectuer des tâches intelligentes comme le tri, l'analyse, la prévision et l'apprentissage — promet des avantages considérables pour les Canadiens et les Canadiennes. Les entreprises qui développent et commercialisent l'IA ont un potentiel de croissance et de création d'emplois. Les organisations qui adoptent les technologies de l'IA peuvent quant à elles améliorer leurs opérations, accroître leur productivité et générer des avantages sanitaires, sociaux et économiques dont tout le monde pourra profiter.

Certaines applications de l'IA présentent pourtant des risques pour les individus et les communautés :

- l'automatisation activée par l'IA menace de perturber les marchés du travail et l'emploi;
- les analyses prédictives dans les secteurs de la finance, de l'éducation, des forces de l'ordre, etc. peuvent renforcer les préjugés raciaux, sexistes et de classe; et
- les méthodes de collecte des données utilisées dans le développement et les applications de l'IA enfreignent souvent la vie privée et sont appliquées sans le consentement des intéressés (voir par exemple *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, *Data Governance in the Digital Age*)

Les décideurs politiques en matière d'IA sont tiraillés entre deux choses. D'un côté, ils doivent établir les conditions qui permettront à l'IA de prospérer et de procurer des avantages qui profiteront à tous. De l'autre, ils doivent être conscients des méfaits que certaines applications de l'IA peuvent générer et renforcer, et les combattre. Face à ce dilemme, plusieurs options sont possibles, qui vont du laisser-faire propice au développement et à la diffusion sans limites de l'IA à une approche prudente qui limiterait le développement de l'IA jusqu'à ce que les risques soient mieux compris et que la capacité à les gérer soit en place. L'IA est une technologie de plateforme dont les applications possibles — et donc les profils de risque — sont multiples. Nous recommandons par conséquent de régir l'IA en adoptant une approche contextuelle au cas par cas de la gestion des risques, plutôt que de laisser faire sans réserve ou d'adopter une approche préventive. Cela rendrait possible le développement des technologies et des applications de l'IA tout en surveillant et gérant les risques à mesure que ceux-ci se pointent dans certaines applications spécifiques.

Pour institutionnaliser au Canada une approche de la gouvernance de l'IA axée sur la gestion des risques, nous recommandons au gouvernement du Canada de créer deux nouveaux organismes :

- un conseil de gouvernance des risques associés à l'IA; et
- une agence d'évaluation de l'impact des algorithmes.

Trouver un juste équilibre entre risques et possibilités

L'adoption d'une politique sur l'IA nécessitera de trouver un juste équilibre entre le soutien au développement et à la diffusion de technologies d'IA laissant présager des avantages sociaux, économiques et autres aux Canadiens et Canadiennes, et la garantie de pouvoir réduire les risques que font peser ces technologies pour les droits et le bien-être des Canadiens et Canadiennes. Ce ne sera pas facile.

L'IA est une technologie naissante. La nature et l'étendue réelles des avantages et des risques potentiels sont donc très incertaines. Certains observateurs sont en faveur d'un laisser-faire imposant peu de limites à la recherche sur l'IA et aux applications de l'IA afin d'accélérer la découverte et l'accès aux avantages associés. Ces derniers peuvent être d'ordre économique — comme la croissance et la création d'emplois par les entreprises canadiennes qui développent et commercialisent des technologies d'IA — ou d'ordre social, financier, politique et sanitaire.

La reconnaissance d'images et l'analyse prédictive, par exemple, améliorent le diagnostic des maladies oculaires et cardiovasculaires, du cancer du sein et du mélanome. Par ailleurs, l'analyse prédictive sert notamment à réduire les risques d'accidents du travail, identifier les enfants risquant d'être victimes de violence, prévoir le risque de réhospitalisation de patients et évaluer la solvabilité des personnes n'ayant pas d'antécédents de crédit traditionnels. Les progrès accomplis dans le traitement du langage naturel ont permis la création d'applications de plus en plus courantes, comme les assistants à commande vocale, le support client automatisé, la traduction, les filtres antipourriel et le dialogue

interactif.

D'autres observateurs sont au contraire en faveur d'une approche prudente qui limiterait le développement et l'utilisation de l'IA pour se donner le temps de connaître les risques et de savoir comment les maîtriser. Les premiers progrès et usages de l'IA ont déjà révélé des défis et risques sérieux, notamment dans les domaines suivants :

1. Biais

Les algorithmes, et les données qui les alimentent, peuvent renforcer les inégalités et préjugés existants notamment en matière de race, de sexe et de classe. L'utilisation de données biaisées dans les modèles prédictifs de maintien de l'ordre, par exemple, peut entraîner une hausse de la surveillance policière dans les quartiers à plus forte concentration de résidents appartenant à des minorités, non pas parce que la criminalité y serait plus forte, mais parce que les ensembles de données dont dépendent les modèles prédictifs comprennent plus de rapports d'enquête pour ces quartiers du fait d'une surveillance policière excessive dans le passé.

2. Sécurité

De nombreuses applications de l'IA font peser des risques pour la sécurité, qui vont de modèles algorithmiques utilisés dans le secteur financier fonctionnant mal et générant des pertes financières catastrophiques à la mise au point d'armes létales autonomes activées par l'IA et échappant au contrôle de l'homme.

3. Protection de la vie privée et consentement

L'accès à de vastes ensembles de données pour appuyer l'apprentissage automatique et améliorer



« Les gens s’attendent à ce que les organisations qui prennent des décisions les concernant, par exemple le refus d’un prêt, d’une prestation gouvernementale ou d’un jugement équitable, peuvent les expliquer, mais les systèmes IA avancés produisent des résultats fondés sur des analyses si complexes qu’elles échappent à la compréhension humaine. »

la capacité d’analyse et de prise de décisions est essentiel à la recherche et à l’innovation en IA. Pourtant, les données sont parfois recueillies et utilisées sans le consentement explicite et valable des personnes auprès desquelles elles ont été obtenues, souvent en violation du droit à la vie privée et des attentes en la matière (voir par exemple, “At least two malls are using facial recognition technology to track shoppers’ ages and genders without telling” and “Big other: surveillance capitalism and the prospects of an information civilization”).

4. Explicabilité et responsabilité

Certaines questions restent sans réponse quant à la mesure dans laquelle il importe d’expliquer aux personnes touchées par l’IA les décisions prises et les actions entreprises par celle-ci, et quant à l’identité des personnes ou entités portant la responsabilité ultime de ces décisions et actions. L’une des principales difficultés est que les gens s’attendent à ce que les organisations qui prennent des décisions les concernant, par exemple le refus d’un prêt, d’une prestation gouvernementale ou d’un jugement équitable, peuvent les expliquer, mais que les sys-

tèmes IA avancés produisent des résultats fondés sur des analyses si complexes qu’elles échappent à la compréhension humaine. Certains systèmes recueillent et analysent par exemple l’historique de navigation Internet de demandeurs de prêt et attribuent à ces derniers une cote de crédit (voir *Weapons of Math Destruction*, 143-5) en fonction du degré de similitude de leur historique de navigation avec celui d’emprunteurs ayant fait défaut par le passé. Or, les éléments précis qui posent problème dans l’historique de navigation — et la mesure dans laquelle ils sont liés à d’autres données et modèles — peuvent nécessiter des explications complexes. Cela posera des difficultés, en particulier pour l’innovation dans le secteur public. Il est en effet probable que la prise de décisions reposant sur l’IA devra être justifiée aux fins de la légitimité politique.

Une approche empreinte de précaution serait synonyme de prudence face à ces risques, mais pourrait aussi retarder la découverte et l’accès à des avantages sociaux, économiques et autres pour les Canadiens et Canadiennes. Comment régir l’IA au vu de ce dilemme innovation/risque?

Le Canada a besoin de politiques sur l'éthique et la gouvernance de l'IA

Au Canada, la manière actuelle d'aborder la gouvernance de l'IA favorise l'innovation par rapport à la gestion des risques. Ce faisant, le pays est mal équipé pour faire face aux risques émergents associés à certaines applications de l'IA.

La Stratégie pancanadienne en matière d'intelligence artificielle fait peu de cas de l'éthique et de la gouvernance de l'IA. De même, le ministère de l'Innovation, des Sciences et du Développement économique (ISDE), à l'image des autres organismes fédéraux, ne semble guère avoir élaboré d'approche globale pour cerner et gérer les risques éthiques, sociaux et politiques de l'IA, et leurs répercussions. La stratégie pancanadienne en matière d'intelligence artificielle préconise l'« établissement d'un leadership éclairé d'envergure mondiale relativement aux répercussions économiques, éthiques, politiques et juridiques des percées en intelligence artificielle » et soutient les chercheurs universitaires qui étudient ces questions dans le cadre du Programme IA et société. Les « résultats attendus » de la stratégie ne mentionnent cependant aucunement l'éthique et la gouver-

nance de l'IA et mettent davantage l'accent sur le profil international du Canada sur un triple plan : la recherche et la formation en IA, le développement et l'attraction de talents en IA, et l'amélioration de l'innovation en vue d'obtenir des avantages socio-économiques.

Lorsqu'on lui demande comment l'IA sera réglementée et régie, l'ISDE se contente de répondre que l'élaboration et l'utilisation de l'IA doivent être conformes au « cadre du marché » existant, à la Charte canadienne des droits et libertés et à la Loi sur la protection des renseignements personnels et les documents électroniques. Le Secrétariat du Conseil du Trésor mène des consultations sur l'utilisation responsable de l'IA dans le secteur public. Affaires mondiales Canada a coordonné un symposium étudiant multi-universitaire sur l'IA et les droits de la personne, et certains analystes du gouvernement fédéral envisagent plusieurs approches pour l'évaluation des biais et des impacts des algorithmes. L'ISDE a en outre amorcé des consultations nationales sur le numérique et les données, qui devraient aborder certains des enjeux liés à la collecte et à l'utilisation des données. Peu d'éléments laissent néanmoins penser qu'une stratégie fédérale explicite et complète en matière d'éthique et de

« Au Canada, la manière actuelle d'aborder la gouvernance de l'IA privilégie l'innovation à la gestion des risques. Ce faisant, le pays est mal équipé pour faire face aux nouveaux risques associés à certaines applications de l'IA. Il faut trouver un meilleur équilibre. »

gouvernance de l'IA soit en cours d'élaboration ou soit même envisagée.

Lors d'un sommet du G7 en décembre pour discuter des répercussions de l'intelligence artificielle, il est apparu que le Canada accordera plus d'attention à l'éthique et la gouvernance de l'IA. Le Canada et la France ont annoncé qu'ils envisagent de créer un Groupe international d'experts en intelligence artificielle ayant pour mission de « soutenir et guider une adoption responsable de l'IA centrée sur l'humain et axée sur le respect des droits de la personne, l'inclusion, la diversité, l'innovation et la croissance économique ». Le groupe d'experts entend mobiliser la communauté scientifique, l'industrie, la société civile, les gouvernements et les organisations internationales sur des questions portant sur la collecte de données et la protection de la vie privée, la confiance en l'IA, l'avenir du travail, l'IA responsable et les droits humains, l'équité, la responsabilité et le bien collectif. Si l'éthique et la gouvernance de l'IA semblent clairement être prioritaires pour le Canada, on ignore quel effet concret aura le travail du groupe d'experts sur la gouvernance de l'IA au Canada.

SOLUTIONS ENVISAGEABLES POUR LA GESTION DES RISQUES ASSOCIÉS À L'IA

Le Canada s'est doté d'une stratégie gouvernementale pour soutenir la recherche et l'innovation dans le domaine de l'IA, et certaines provinces investissent des montants considérables dans la recherche sur l'IA, mais les stratégies et mécanismes institutionnels manquent pour cerner, surveiller et atténuer les risques associés à l'IA.

Quels principes et quelles solutions stratégiques pourraient combler les lacunes associées à l'adoption d'une approche de gestion contextuelle au cas par cas des risques?

Les principes

Pour gérer le dilemme entre le soutien à l'innovation et la gestion des risques, l'approche du Canada en matière de gouvernance de l'IA devrait intégrer plusieurs éléments. Elle devrait notamment :

- adopter une politique sur le développement et l'utilisation responsables de l'IA qui accorde la priorité à l'équité, à l'égalité, à la sécurité économique et politique, à la santé et au bien-être de tous; et
- concentrer des mesures de gestion des risques et de réglementation sur les applications d'IA plutôt que sur l'IA en général. Les risques associés à l'IA ne se manifesteront que dans le contexte d'applications et d'utilisations concrètes dans des activités et secteurs particuliers comme le diagnostic de santé, l'évaluation de prêts, la prévention de la criminalité ou l'estimation de l'admissibilité à des prestations. L'évaluation et la gestion des risques devraient mettre l'accent sur ce qui s'avère approprié dans ces contextes.

Les politiques

Pour ce qui est des politiques et des règlements à proprement parler, les gouvernements du Canada devraient tenir compte de ce qui suit :

- Élaborer et adopter une déclaration sur le développement et l'utilisation responsables de l'IA qui annoncerait aux concepteurs et aux adoptants du secteur privé ainsi

qu'aux décideurs et aux fonctionnaires du secteur public l'importance d'accorder la priorité à l'équité, à la sécurité, à la santé et à d'autres valeurs, principes et intérêts dans le développement et l'utilisation de l'IA. La déclaration devrait s'appuyer sur la [Déclaration de Montréal pour le développement responsable de l'IA](#).

- Élaborer une stratégie plus détaillée en matière d'IA qui fournit des directives explicites et apporte un financement pour explorer et gérer les dimensions éthiques, économiques, juridiques et sociales de l'IA, lesquelles sont largement négligées dans la stratégie pancanadienne actuelle axée sur l'innovation. Le Canada s'alignerait ainsi davantage sur d'autres pays qui s'efforcent de tenir compte des volets de l'innovation et de l'éthique de l'IA dans leur contexte national. On peut s'inspirer des rapports français *Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne*, suédois *National Approach for Artificial Intelligence*, britannique *AI in the U.K.: Ready, willing and able?*, etc.
- Exiger la réalisation d'évaluations d'impact des algorithmes avant que l'IA ne soit utilisée dans des domaines sensibles comme les soins de santé, l'éducation, la sécurité publique ou l'octroi de prestations gouvernementales. Elles seraient semblables aux évaluations technologiques en santé et aux études d'impact environnemental, mais porteraient sur [les risques et les avantages](#) de l'IA pour les personnes et les collectivités, tout en assurant la répartition des risques et des avantages pour chaque groupe démographique.
- Envisager d'établir un droit à une explication lorsqu'un système fondé sur l'IA prend des

décisions qui ont des répercussions importantes sur les intérêts fondamentaux financiers, juridiques ou autres d'une personne. [Le Règlement général de l'Union européenne sur la protection des données](#), qui établit sans doute un tel droit, pourrait servir de cadre au débat sur sa pertinence et sa faisabilité. La question de savoir si un tel droit devrait exister — et s'il est techniquement possible d'offrir des explications — devra faire l'objet d'un débat public. À tout le moins, les utilisateurs de l'IA appartenant aux secteurs privé et public devraient être informés du fait qu'ils seront tenus responsables des résultats qui affectent les droits et les intérêts des particuliers.

Dispositifs institutionnels

Pour mettre en œuvre ces principes, politiques et activités d'évaluation et pour fournir des mécanismes de discussion et de gestion des risques liés à l'IA, il convient de mettre en place certains dispositifs institutionnels. Le Canada, les provinces et les territoires devraient envisager de créer ce qui suit :

1 Un conseil dédié à la gouvernance des risques associés à l'intelligence artificielle

L'organe devrait être composé d'experts techniques, juridiques et éthiques pour discuter, évaluer, relater et fournir des conseils au gouvernement et à l'industrie sur l'innovation concernant l'IA et la gestion des risques qui lui sont associés. Plus précisément, le conseil devrait :

- piloter la rédaction d'une déclaration sur le développement et l'utilisation responsables de l'IA et d'une stratégie plus globale pour la gouvernance de l'IA;
- surveiller les tendances de la recherche

sur l'IA et de l'application de l'IA et en faire rapport, et évaluer régulièrement les risques associés aux applications nouvelles, émergentes et proliférantes;

- conseiller le gouvernement et l'industrie quant à la façon de gérer les risques, en s'inspirant des résultats de l'évaluation des risques et des meilleures pratiques d'autres régions; et
- servir d'organisme de coordination pour les discussions canadiennes et internationales sur les risques posés par l'IA dans chaque secteur (p. ex. santé, éducation, innovation, développement économique, droit) et chaque ordre de gouvernement (fédéral, provincial, territorial et municipal).

Le conseil pourrait revêtir la forme d'un organe permanent et autonome du [Conseil des académies canadiennes \(CAC\)](#) l'organisme d'évaluation scientifique indépendant du gouvernement fédéral, et ainsi bénéficier de la capacité opérationnelle et des atouts actuels du CAC pour ce qui est de réunir des experts du milieu universitaire, de l'industrie et des organismes sans but lucratif.

2 Créer une agence d'évaluation de l'impact des algorithmes

L'agence devrait être composée d'experts techniques, juridiques et éthiques qui seraient chargés d'effectuer les évaluations jugées nécessaires ou souhaitables par les ministères et organismes fédéraux, provinciaux et territoriaux, et de veiller à ce que les applications de l'IA et des algorithmes respectent les droits, les intérêts et le bien-être des Canadiens et Canadiennes.

VERS L'INNOVATION ET LA BONNE GOUVERNANCE EN MATIÈRE D'IA

Le Canada a l'occasion d'être un chef de file mondial en matière de recherche et d'innovation sur l'IA et de gouvernance efficace de l'IA. Bien que la création d'avantages sanitaires, économiques et sociaux soit déjà une priorité des gouvernements du Canada, la gestion des potentiels risques sanitaire, juridique, économique et éthique des applications de l'IA a largement été reléguée au second plan. L'expérience acquise en rapport avec d'autres technologies émergentes aurait dû nous apprendre qu'une bonne gestion des risques est une condition préalable pour déterminer et réduire les inconvénients et, ce faisant, pour inspirer suffisamment confiance au public et favoriser la concrétisation des innovations. L'avenir nous dira si ces leçons seront appliquées à la gouvernance de l'IA ou si nous connaissons un avenir de risques d'IA non réglementés et d'innovation à la traîne.

M. Daniel Munro est chercheur invité et directeur de projets de politiques au laboratoire des politiques d'innovation de la Munk School of Global Affairs and Public Policy de l'Université de Toronto. Ses intérêts de recherche comprennent les politiques en matière de science, de technologie et d'innovation ainsi que l'éthique appliquée, dont l'éthique des nouvelles technologies émergentes.

L'auteur tient à remercier Tim Dutton, Sylvia Kingsmill, Maya Medeiros, Aaron Reynolds et Mark Sutcliff pour leurs suggestions utiles et leurs échanges instructifs.



**FORUM DES
POLITIQUES
PUBLIQUES**