

FAIRE PROGRESSER LA CAPACITÉ NATIONALE SOVERAINE DE CALCUL POUR L'IA DU CANADA

EXPOSÉ DE POLITIQUE DU CTIC



Information and
Communications
Technology Council

Conseil des technologies
de l'information
et des communications

Recherche réalisée par



PRÉFACE :

Le Conseil des technologies de l'information et des communications (CTIC) est un centre national d'expertise neutre et sans but lucratif dont la mission est de renforcer l'avantage numérique du Canada dans l'économie mondiale. Depuis plus de 30 ans, le CTIC propose aux particuliers et aux entreprises des recherches prospectives, des conseils stratégiques pratiques et des solutions de développement des capacités. L'objectif de l'organisation est de faire en sorte que la technologie soit utilisée pour stimuler la croissance économique et l'innovation et que la main-d'œuvre canadienne demeure concurrentielle à l'échelle mondiale.

 ictc-ctic.ca/fr

 info@ictc-ctic.ca

POUR CITER CE RAPPORT :

Erik Henningsmoen, Noah Lubendo, Mairead Matthews et Heather McGeer. Faire progresser la capacité nationale souveraine de calcul pour l'IA du Canada : Exposé de politique du CTIC. Conseil des technologies de l'information et des communications (CTIC), juillet 2025. Ottawa, Canada. Les auteurs sont classés par ordre alphabétique.

Recherche et rédaction : Erik Henningsmoen (analyste principal de la recherche et des politiques), Heather McGeer (analyste principale de la recherche et des politiques), Mairead Matthews et Noah Lubendo (analystes de la recherche et des politiques), avec le généreux soutien de Namir Anani et Faun Rice, ainsi que de l'équipe Recherche et politiques du CTIC.

TABLE DES MATIÈRES

Résumé.....	4
-------------	---

Contexte	6
----------------	---

Section I : Accès aux capacités de calcul..... 8

Comment l'industrie canadienne accède actuellement aux capacités de calcul	8
Problèmes les plus urgents auxquels l'industrie canadienne se heurte dans l'accès aux capacités de calcul pour l'IA	10
Propriété étrangère et durabilité environnementale des infrastructures de calcul pour l'IA.....	11
Niveaux et types d'accès aux capacités de calcul pour l'IA requis par l'industrie pour se développer et programmes de soutien gouvernementaux existants	12

Section II : Développer une capacité de calcul souveraine et durable..... 14

Les avantages du développement d'une infrastructure de calcul de pointe détenue et contrôlée par le Canada.....	15
Opportunités et initiatives à court terme pour aider à élargir l'accès aux infrastructures de calcul existantes au Canada.....	17
Possibilités d'intégration du matériel informatique et des logiciels canadiens dans une infrastructure souveraine de calcul pour l'IA	18
Modèles de collaboration pour soutenir les partenariats entre l'industrie et le monde universitaire	18
Étude de cas : Ressources nationales de recherche en IA aux États-Unis.....	19
Développer des relations et collaborer avec des pays partageant les mêmes idées	22 . 19

Section III : Priorités à long terme pour l'infrastructure de calcul..... 20

Éléments prioritaires pour le développement d'une infrastructure nationale de calcul pour l'IA au Canada dans les cinq prochaines années.....	20
Recommandations	23

Conclusion.....	25
-----------------	----

RÉSUMÉ

La technologie de l'intelligence artificielle (IA) repose sur une infrastructure informatique avancée et étendue, comprenant notamment des centres de données équipés de plateformes d'infonuagique hautement performantes et de puces IA spécialisées, souvent coûteuses et rares. Elle nécessite également une panoplie de ressources habilitantes, notamment des ressources techniques telles que des cadres, des bibliothèques, des modèles et des données d'apprentissage pour l'IA ; des ressources humaines, comme des technologues et des techniciens qualifiés, des chercheurs et des scientifiques, ainsi que des ingénieurs pour développer, tester, déployer et exploiter la technologie de l'IA ; ainsi que des infrastructures de soutien, telles que des réseaux de télécommunications à haut débit.

La capacité d'une économie nationale à favoriser et à héberger ces technologies d'IA et les ressources habilitantes nécessaires à leur fonctionnement au niveau national, sans participation ni contrôle étrangers significatifs, constitue la capacité de calcul souveraine d'un pays en matière d'IA. Le développement et l'adaptation rapides et étonnants des technologies d'IA ces dernières années ont attiré une attention croissante sur la question du développement d'une capacité de calcul souveraine en matière d'IA au Canada. Les économies qui sont en mesure de développer et d'adopter efficacement des technologies d'IA de pointe et d'intégrer l'IA dans leur économie, leur écosystème de recherche et d'innovation, leur système d'éducation et de formation et leur secteur public sont bien placées pour bénéficier d'un avantage économique et technologique important dans un ordre économique mondial de plus en plus concurrentiel et incertain.

Au Canada, le débat sur l'élaboration d'une stratégie nationale sur la capacité de calcul souveraine pour l'IA a donné lieu à une consultation publique fédérale visant à définir l'approche future du gouvernement en matière de calcul pour l'IA. Les commentaires recueillis ont abouti à la publication, en décembre 2024, de la Stratégie canadienne sur la capacité de calcul souveraine pour l'IA du gouvernement du Canada et à l'affectation d'un financement de 2 milliards de dollars à ce domaine. En 2025, le gouvernement fédéral a nommé pour la première fois un ministre canadien de l'IA et de l'innovation numérique. Il a également joué un rôle de premier plan sur la scène internationale concernant les questions de gouvernance liées à l'IA lors du 51^e sommet du G7, qui s'est tenu à Kananaskis, en Alberta, en juin 2025.

Les économies comparables, telles celles de la France, de l'Union européenne et des États-Unis, développent

également leurs propres capacités de calcul souveraines pour l'IA, à un rythme soutenu et à grande échelle. Dans le cas des États-Unis, le projet Stargate prévoit d'investir 500 milliards de dollars américains dans les technologies d'IA, y compris les infrastructures de calcul pour l'IA, au cours des quatre prochaines années.

Malgré des années de rôle de chef de file en matière de recherche et de politiques dans le domaine de l'IA, le Canada risque de prendre du retard dans le développement des capacités informatiques souveraines nécessaires pour définir son propre avenir économique et technologique. Le Canada est à la traîne par rapport à ses pairs en matière d'investissement et de développement de capacités informatiques souveraines dans le domaine de l'IA.

Les conséquences d'un échec dans le développement de cette capacité sont importantes. Sans infrastructure informatique souveraine en IA, les chercheurs, les innovateurs, les entrepreneurs, les entreprises et les institutions du secteur public canadiens resteront dépendants de l'accès à des infrastructures informatiques étrangères en IA. Cela mettra l'avenir numérique du pays entre les mains de fournisseurs internationaux et le soumettra aux décisions politiques de gouvernements étrangers. Dans un monde marqué par des tensions géopolitiques croissantes, des différends diplomatiques et commerciaux persistants et une incertitude croissante quant à la coopération mondiale, l'accès continu à des services informatiques liés à l'IA contrôlés par des pays étrangers pour soutenir l'écosystème canadien de l'IA n'est plus acquis d'avance.

À l'heure actuelle, les entreprises canadiennes accèdent à une capacité de calcul pour l'IA par le biais d'une combinaison de postes de travail personnels, de centres

de données internes, de services en nuage et d'outils provenant de fournisseurs. Si l'infonuagique gagne en popularité en raison de sa flexibilité et de ses coûts initiaux moins élevés, son accès reste limité par des coûts élevés, des pénuries de matériel, un manque de talents et l'insuffisance des bandes passantes dans les zones rurales et éloignées. L'infrastructure informatique nationale limitée du Canada en matière d'IA est également source de défis nationaux, et notamment de préoccupations concernant la souveraineté des données, la protection de la propriété intellectuelle (PI) et l'accès à long terme aux ressources informatiques essentielles.

Le développement d'une capacité informatique durable et détenue par des intérêts canadiens peut renforcer la sécurité nationale, soutenir la croissance économique, réduire l'impact environnemental et contribuer à retenir et développer les talents nationaux en IA. Cependant, pour garantir que les entreprises accordent la priorité à la propriété canadienne lorsqu'elles prennent des décisions en matière d'infrastructure, il faudra mettre de l'avant des investissements publics, des incitations et des modèles de collaboration qui soutiennent les partenariats entre l'industrie et le milieu universitaire et s'alignent sur ceux de leurs homologues mondiaux, de manière à renforcer l'écosystème d'innovation en IA du Canada. En outre, une intervention stratégique est nécessaire pour combler les lacunes critiques dans le pipeline de développement de l'IA, notamment en matière de matériel, de modèles, de données d'apprentissage et de talents. L'engagement récent du Canada à respecter les objectifs de dépenses de défense de l'OTAN pourrait offrir une occasion de contribuer à garantir la souveraineté et la position concurrentielle mondiale du Canada dans le domaine de l'IA en renforçant son infrastructure nationale de calcul pour l'IA.

Le présent exposé de politique recommande que cet effort soit guidé par six principes clés : une stratégie industrielle solide, la possession et la conservation de la propriété intellectuelle, la résilience de la chaîne d'approvisionnement, la durabilité environnementale, l'équité régionale et des incitations fiscales et financières ciblées. Des investissements stratégiques dans les infrastructures détenues par des intérêts canadiens, notamment pour le soutien à la fabrication nationale de puces, aux grappes d'IA et au calcul de haute performance, contribueront à conserver la propriété intellectuelle, à réduire la dépendance à l'égard des plateformes étrangères et à préserver la position du Canada dans les secteurs cruciaux de l'IA. Le renforcement de la résilience de la chaîne d'approvisionnement et la priorité accordée à la durabilité atténueront les risques géopolitiques et environnementaux, tandis qu'un accès régional plus large et des mécanismes de co-investissement public-privé garantiront une croissance inclusive dans l'ensemble du Canada.

L'exposé de politique s'appuie sur un document d'orientation présenté en septembre 2024 par le CTIC dans le cadre des consultations publiques de l'automne 2024 sur la puissance de calcul pour l'IA organisées par Innovation, Sciences et Développement économique Canada¹. Sur la base du document d'orientation original du CTIC, l'exposé de politique a été mis à jour et élargi afin de tenir compte des nouveaux développements concernant la politique publique sur la capacité de calcul pour l'IA, des progrès réalisés dans le domaine de l'IA et des technologies numériques, ainsi que de l'évolution des conditions du marché.

¹ Voir : Innovation, Sciences et Développement économique Canada (gouvernement du Canada), « Consultations sur la puissance de calcul pour l'intelligence artificielle (IA) », dernière mise à jour le 5 décembre 2024, <https://ised-isde.canada.ca/site/isde/fr/consultations-publiques/consultations-puissance-calcul-pour-lintelligence-artificielle-ia>.

CONTEXTE

Les innovateurs, les entrepreneurs et les chercheurs canadiens ont besoin d'une grande capacité de stockage de données et d'une puissance de calcul de pointe pour créer, former, tester et déployer des modèles d'intelligence artificielle (IA)². L'AI Now Institute avertit que la rareté des ressources informatiques pour l'IA constitue actuellement un goulot d'étranglement dans la formation et le déploiement des grands modèles de langage (GML), ce qui confère un avantage considérable en matière de développement et d'exploitation des technologies d'IA à ceux qui ont un accès suffisant à des ressources informatiques avancées³.

Par rapport à d'autres économies avancées, le Canada est à la traîne en matière d'infrastructures physiques telles que les serveurs et les processeurs haute performance essentiels à la capacité de calcul pour l'IA⁴. Une étude réalisée en 2023 par le Tony Blair Institute note qu'à l'instar de l'Italie, « le Canada est à la traîne de tous les autres pays du G7 en matière d'investissements dans la capacité de calcul »⁵. Bien qu'il dispose de l'une des capacités globales les plus solides au monde en matière d'IA, grâce à sa position de chef de file en matière de politique d'IA, à la qualité de sa recherche et au développement de ses talents, le Canada pourrait avoir du mal à tirer pleinement parti de ces atouts en raison d'un manque de ressources nationales en matière de capacité de calcul pour l'IA.

Ces défis surviennent dans un contexte d'incertitude géopolitique, de changements technologiques et économiques rapides et de pressions immenses sur l'ordre international fondé sur des règles. Dans un contexte d'alliances et de relations commerciales incertaines à long terme et de menaces à la souveraineté nationale

du Canada, il sera essentiel, dans les années à venir, de développer une capacité de calcul nationale souveraine en matière d'IA afin que le Canada puisse tirer parti des avantages technologiques, économiques et scientifiques de cette technologie.

Conscient de ces défis, le gouvernement du Canada a mené en juin 2024 une consultation publique afin de définir son approche stratégique future en matière de capacité de calcul pour l'IA⁶. Les consultations ont donné lieu à la publication, en décembre 2024, de la Stratégie canadienne sur la capacité de calcul souveraine pour l'IA⁷, ainsi qu'à l'annonce, au début de 2025, d'un programme d'infrastructure souveraine de calcul pour l'IA et d'un Fonds d'accès à une puissance de calcul pour l'IA de 300 millions de dollars⁸. De plus, en novembre 2024, le gouvernement du Canada a lancé l'Institut canadien de la sécurité de l'intelligence artificielle⁹. Ces mesures stratégiques axées sur la capacité de calcul pour l'IA s'appuient sur des politiques gouvernementales tournées vers l'avenir en matière de recherche et d'adoption de l'IA adoptées au cours de la dernière décennie, notamment la Stratégie pancanadienne en matière d'intelligence artificielle de 2017¹⁰.

En outre, le gouvernement fédéral a reconnu le caractère critique de la situation actuelle, notamment en ce qui concerne les investissements dans le développement de l'IA. En mai 2025, le gouvernement libéral nouvellement élu a publié une lettre de mandat qui énonçait les principaux objectifs et priorités du gouvernement fédéral¹¹. Dans cette lettre, le premier ministre Mark Carney a souligné l'importance d'accroître les investissements dans les infrastructures d'IA afin de stimuler la productivité

2 « What are large language models (LLMs)? », IBM, 2 novembre 2023, <https://www.ibm.com/topics/large-language-models>.

3 Jai Vipra et Sarah Myers West, « Computational Power and AI », AI Now Institute, septembre 2023, <https://ainowinstitute.org/publication/policy/compute-and-ai>.

4 Graham Dobbs et Jake Hirsch-Allen, « Qu'en est-il de l'informatique au Canada? Options politiques pour remédier au déficit informatique pour l'IA au Canada », The Dais, Université métropolitaine de Toronto, mars 2024, <https://dais.ca/wp-content/uploads/2024/05/Quen-est-il-de-linformatique-au-Canada.pdf>, 5.

5 Bridget Boakye, et al., State of Compute Access: How to Bridge the New Digital Divide, Tony Blair Institute for Global Change, novembre 2023, <https://www.institute.global/insights/tech-and-digitalisation/state-of-compute-access-how-to-bridge-the-new-digital-divide>.

6 Innovation, Sciences et Développement économique Canada (gouvernement du Canada), « Consultations sur la puissance de calcul pour l'intelligence artificielle (IA) », dernière mise à jour le 5 décembre 2024, <https://ised-isde.canada.ca/site/ised/fr/consultations-publiques/consultations-puissance-calcul-pour-lintelligence-artificielle-ia>.

7 Innovation, Sciences et Développement économique Canada (Gouvernement du Canada), « Stratégie canadienne sur la capacité de calcul souveraine pour l'IA », dernière mise à jour le 7 mars 2025, <https://ised-isde.canada.ca/site/ised/fr/strategie-canada-puissance-calcul-souveraine-pour-lia>.

8 Innovation, Sciences et Développement économique Canada (gouvernement du Canada), « Programme d'infrastructure de calcul souveraine pour l'IA », dernière mise à jour le 27 février 2025, <https://ised-isde.canada.ca/site/ised/fr/programme-dinfrastructure-calcul-souveraine-pour-lia>; Innovation, Sciences et Développement économique Canada (gouvernement du Canada), « Fonds d'accès à une capacité de calcul pour l'IA », dernière mise à jour le 14 mars 2025, <https://ised-isde.canada.ca/site/ised/fr/strategie-canada-puissance-calcul-souveraine-pour-lia/fonds-dacces-capacite-calcul-pour-lia>.

9 Innovation, Sciences et Développement économique Canada (Gouvernement du Canada), « Le Canada lance l'Institut canadien de la sécurité de l'intelligence artificielle », communiqué de presse, 12 novembre 2024, <https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2024/11/le-canada-lance-linstitut-canadien-de-la-securite-de-lintelligence-artificielle.html>.

10 Organisation de coopération et de développement économiques (OCDE), « Pan-Canadian artificial intelligence strategy », Plateforme de l'OCDE sur les politiques numériques, consulté le 29 mai 2025, <https://depp.oecd.org/policies/CAN1230>.

11 Cabinet du premier ministre, « Lettre de mandat », 21 mai 2025, <https://www.pm.gc.ca/fr/lettres-de-mandat/2025/05/21/lettre-de-mandat>.



et l'adoption des technologies¹². En plus de la lettre de mandat, M. Carney a également créé en mai 2025 un nouveau poste de ministre canadien de l'IA et de l'innovation numérique¹³.

Le Canada a également continué à se positionner sur la scène mondiale, notamment en ce qui concerne la promotion de la collaboration internationale dans le domaine de l'IA. En juin 2025, le Canada, qui assurait la présidence du G7 pour 2025, a annoncé le lancement du Grand défi IAgov du G7 afin de trouver des solutions aux obstacles à l'adoption de l'IA, ainsi que son intention de créer le Réseau IA du G7 (RIAG)¹⁴. De plus, dans le cadre de la Feuille de route du G7 pour l'adoption de l'IA, des engagements ont été pris pour favoriser l'accès aux infrastructures informatiques et numériques¹⁵.

Ces développements dans le domaine de la capacité de calcul de l'IA au Canada arrivent à un moment où d'autres économies avancées consacrent d'importants investissements publics et privés aux infrastructures de calcul de l'IA. Par exemple, en janvier 2025, les États-Unis ont annoncé le projet Stargate. Grâce à un partenariat entre SoftBank, OpenAI, Oracle, MGX, Arm, Microsoft et NVIDIA, cette initiative de 500 milliards de dollars américains vise à « garantir le leadership américain dans le domaine de l'IA »¹⁶ en construisant jusqu'à 20 nouveaux centres de données puissants réservés à l'IA et en soutenant les infrastructures électriques dans l'ensemble des États-Unis¹⁷. En juillet 2025, les États-Unis ont dévoilé un plan d'action national pour l'intelligence artificielle¹⁸.

De même, en février 2025, la France a annoncé des investissements de 109 milliards d'euros dans les infrastructures d'IA en prévision du Sommet pour l'action sur l'IA¹⁹. Lors de ce sommet, l'Union européenne a annoncé l'initiative InvestAI visant à « mobiliser » 200 milliards d'euros d'investissements dans l'IA, dont 20 milliards pour la construction de quatre « giga-usines d'IA », consistant en des centres de données à grande

échelle destinés à l'entraînement des modèles d'IA²⁰. En janvier 2025, le Royaume-Uni a publié un plan d'action sur les opportunités de l'IA, qui vise à multiplier par vingt d'ici 2030 ses ressources informatiques utilisées dans la recherche de pointe en IA²¹.

Ces importants investissements nationaux dans les capacités informatiques de l'économie collaborative en matière d'IA surviennent à un moment où les tensions géopolitiques s'intensifient, où la situation internationale en matière de sécurité se détériore et où la souveraineté nationale dans le secteur numérique suscite des inquiétudes. Les préoccupations en matière de souveraineté nationale comprennent le maintien du contrôle national sur les infrastructures informatiques de pointe, la préservation de la souveraineté numérique, ainsi que la protection et la promotion des filières d'innovation nationales. En outre, les engagements pris par le Canada en matière d'augmentation des dépenses de défense offrent la possibilité d'investir dans les infrastructures d'IA, avec les retombées positives que cela apporterait.

Le présent exposé de politique énonce les principales considérations et recommandations politiques pour le développement d'une infrastructure nationale souveraine en matière d'IA et de calcul avancé pour le Canada. Il comprend une évaluation de la manière dont l'industrie canadienne accède actuellement aux ressources informatiques avancées et aux ressources fonduagiques, des problèmes urgents auxquels elle est confrontée et des nouvelles approches adoptées pour accéder à la capacité de calcul pour l'IA ; du soutien dont les entrepreneurs, les innovateurs et les chercheurs universitaires canadiens ont besoin pour accéder à la capacité de calcul pour l'IA, ce qui inclut les modèles de collaboration entre l'industrie et le monde universitaire ; des risques transfrontaliers et des possibilités de collaboration internationale ; ainsi que des opportunités et des éléments prioritaires pour le développement d'une capacité de calcul souveraine pour l'IA au Canada au cours des cinq prochaines années.

¹² Ibid.

¹³ Cabinet du premier ministre, « Le premier ministre Carney présente le nouveau Conseil des ministres », 13 mai 2025, <https://www.pm.gc.ca/fr/nouvelles/communiqués/2025/05/13/premier-ministre-carney-presente-nouveau-conseil-des-ministres>.

¹⁴ G7 2025 Kananaskis, « Déclaration des dirigeants du G7 sur l'IA pour la prospérité », 17 juin 2025 <https://g7.canada.ca/fr/nouvelles-medias/nouvelles/declaration-des-dirigeants-du-g7-sur-lia-pour-la-prosperite/>.

¹⁵ Ibid.

¹⁶ OpenAI, « Announcing the Stargate Project », 21 janvier 2025, <https://openai.com/index/announcing-the-stargate-project/>.

¹⁷ « Trump highlights partnership investing \$500 billion in AI », Associated Press, 22 janvier 2025, <https://apnews.com/article/trump-ai-openai-oracle-softbank-son-altman-ellison-be261f8a8ee07a0623d4170397348c41>.

¹⁸ « White House Unveils America's AI Action Plan », The White House (États-Unis), 23 juillet 2025, <https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/>.

¹⁹ « IA : avec l'annonce de "109 milliards d'euros d'investissements" privés, Emmanuel Macron entend se mesurer aux États-Unis », Le Monde, 10 février 2025, https://www.lemonde.fr/economie/article/2025/02/09/intelligence-artificielle-avec-l-annonce-de-109-milliards-d-euros-d-investissement-emmanuel-macron-entend-se-mesurer-aux-etats-unis_6539184_3234.html.

²⁰ Commission européenne (Union européenne), « EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence », communiqué de presse, 11 février 2025, https://luxembourg.representation.ec.europa.eu/actualites-et-evenements/actualites/eu-launches-investai-initiative-mobilise-eu200-billion-investment-artificial-intelligence-2025-02-11_en.

²¹ Gouvernement de Sa Majesté (Royaume-Uni), « AI Opportunities Action Plan », 13 janvier 2025, <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>.



SECTION I :

ACCÈS AUX CAPACITÉS DE CALCUL

Le CTIC consulte régulièrement l'industrie canadienne au sujet de son infrastructure de technologies de l'information et des communications (TIC), notamment pour savoir si elle acquiert, construit et exploite ses propres centres de données à l'interne ou si elle opte pour des fournisseurs de services infonuagiques externes, comment elle choisit entre différents fournisseurs de services infonuagiques et comment elle décide du type de matériel à utiliser, comme les unités de traitement graphique (GPU). Les informations pertinentes issues des consultations du CTIC auprès de l'industrie sont résumées ci-dessous.

COMMENT L'INDUSTRIE CANADIENNE ACCÈDE ACTUELLEMENT AUX CAPACITÉS DE CALCUL

Les jeunes entreprises et les petites et moyennes entreprises (PME) canadiennes utilisent un large éventail d'infrastructures de calcul pour l'IA afin de créer, former, tester et déployer des modèles d'IA, notamment :



Des postes de travail personnels, tels que des ordinateurs personnels ou des machines ;



Infrastructures informatiques centralisées, notamment des centres de données internes ou des fournisseurs de services infonuagiques externes ;



Des cadres et bibliothèques d'IA ;



Des modèles d'IA ;



Des données d'apprentissage.

Si certaines ressources, telles que les cadres, bibliothèques, modèles et données d'apprentissage IA, sont de source ouverte et librement accessibles, d'autres sont de propriété exclusive et nécessitent un investissement financier. Outre les infrastructures des TIC, l'industrie a besoin de talents possédant les compétences et connaissances nécessaires pour intégrer les ressources d'IA dans les produits, les services et les applications industrielles. Les postes les plus en demande liés au développement de l'IA au sein des entreprises canadiennes sont notamment ceux d'ingénieur en IA, de chercheur en IA, d'ingénieur en vision par ordinateur, d'ingénieur de données, de scientifique de données, d'ingénieur en apprentissage automatique et de développeur de logiciels²².

En 2021, environ la moitié des entreprises canadiennes avaient accès à une infrastructure informatique centralisée par le biais d'un réseau informatique interne à l'échelle de l'entreprise, tandis qu'une proportion similaire recourait à l'infonuagique²³. De plus en plus, les entreprises canadiennes choisissent d'externaliser tout ou partie de leur infrastructure informatique centralisée à des fournisseurs de services en nuage, et la proportion de celles qui utilisent une infrastructure infonuagique devrait continuer à augmenter²⁴. Les services en nuage nécessitent un investissement initial moindre, réduisent

l'ampleur des besoins en talents et compétences internes et permettent aux jeunes entreprises et aux PME d'adapter leur infrastructure des TIC de manière dynamique en fonction de l'évolution de leurs besoins informatiques. De plus, les fournisseurs de services en nuage offrent bien souvent aux jeunes entreprises et aux PME des « crédits informatiques » gratuits afin de réduire les coûts initiaux et de fidéliser leurs clients. La rentabilité à long terme des services en nuage externalisés par comparaison au développement de centres de données internes appartenant à l'entreprise dépend de facteurs tels que la demande globale de ressources informatiques de l'entreprise, le type de ressources nécessaires, ainsi que la stabilité et la prévisibilité de ses charges de travail informatiques.

Les acteurs du secteur ont également accès à des capacités de calcul pour l'IA par l'intermédiaire de leurs fournisseurs de produits et services d'IA. Par exemple, les organisations qui utilisent ChatGPT via son interface web ou son API (interface de programmation d'application) externalisent non seulement l'IA générative, mais aussi les ressources informatiques nécessaires. De même, les entreprises dont les développeurs utilisent des outils tels que GitHub Copilot ou Cursor, qui intègrent des GML, accèdent aux ressources de calcul pour l'IA par l'intermédiaire de leurs fournisseurs.

²² Comme l'indique le récent rapport du CTIC : Todd Legere, Sheldon Lopez et Noah Lubendo, « L'écosystème de l'IA au Canada : coup d'œil sur les compétences recherchées et les tendances », Conseil des technologies de l'information et des communications (CTIC), février 2025, <https://ictc-ctic.ca/fr/rapports/ecosysteme-de-lia-au-canada-coup-doeil-sur-les-competences-recherchees-et-les-tendances>.

²³ En 2021, 45,3 % des entreprises interrogées par Statistique Canada utilisaient l'informatique en nuage, tandis que 53 % utilisaient un réseau informatique à l'échelle de l'entreprise. Voir : « Technologies de l'information et des communications utilisées par industrie et taille d'entreprise », septembre 2022, Statistique Canada, https://www150.statcan.gc.ca/t1/tbl1/fr/tv.action?pid=2210011701&request_locale=fr.

²⁴ « Les organisations canadiennes s'éloignent de plus en plus de leurs déploiements traditionnels de centres de données pour exploiter leur capacité de calcul et de stockage sur des nuages privés et publics exclusifs. » Voir : « Operating a Digital Business in a Hybrid Cloud World », IDC Canada, <https://74388.fs1.hubspotusercontent-na1.net/hubfs/74388/2023%20Cloud%20Report%20-%20En.pdf> ; « 93 % des répondants canadiens déclarent avoir atteint un stade avancé dans l'adoption de la technologie nuagique », tandis que « 59 % des répondants affirment que 40 % ou plus de la charge de travail de leur entreprise se trouve désormais dans le nuage ». Voir : « How to go from migration to modernization by building a strong cloud and automation strategy », mai 2023, KPMG, <https://kpmg.com/ca/en/home/insights/2023/05/building-a-better-cloud.html> ; De 2019 à 2021, le pourcentage d'entreprises canadiennes qui utilisent l'infonuagique a augmenté de 16,5 % ; le pourcentage de celles qui utilisent un réseau informatique à l'échelle de l'entreprise a quant à lui augmenté de 4,5 %. Voir : « Technologies de l'information et des communications utilisées par industrie et taille d'entreprise », septembre 2022, Statistique Canada, https://www150.statcan.gc.ca/t1/tbl1/fr/tv.action?pid=2210011701&request_locale=fr ; IDC estime que les infrastructures non infonuagiques représentaient 34,9 % des infrastructures en 2023 et prévoit que cette proportion diminuera pour atteindre 26,4 % d'ici 2028. Les services infonuagiques partagés et réservés représentaient quant à eux 65,1 % des infrastructures en 2023, proportion qui devrait atteindre 73,6 % d'ici 2028. Voir : « Spending on Shared Cloud Infrastructure Continues to Lead the Way in Enterprise Infrastructure Investments, According to IDC Tracker », mars 2024, IDC, <https://www.idc.com/getdoc.jsp?containerId=prUS52001524>.



PROBLÈMES LES PLUS URGENTS AUXQUELS L'INDUSTRIE CANADIENNE SE HEURTE DANS L'ACCÈS AUX CAPACITÉS DE CALCUL POUR L'IA

Les problèmes les plus urgents auxquels est confrontée l'industrie canadienne dans l'accès aux capacités de calcul pour l'IA au sein de l'écosystème sont actuellement les exigences techniques, le coût, la situation géographique, la disponibilité de talents et la disponibilité du matériel.

Exigences techniques :

Les exigences techniques sont la principale considération pour l'industrie canadienne lorsqu'elle prend des décisions en matière d'infrastructure des TIC. Par exemple, les organisations soumises à des exigences en matière de résidence des données doivent s'assurer que le stockage et le traitement des données se font au Canada. Les organisations opérant dans des secteurs d'infrastructure critiques doivent intégrer une redondance dans leur infrastructure des TIC en dupliquant le traitement et le stockage des données entre de multiples fournisseurs et emplacements géographiques. De même, les entreprises qui intègrent des modèles d'IA dans leurs produits et services doivent utiliser du matériel informatique suffisamment avancé, tel que des GPU, pour faire fonctionner leurs applications de manière efficace.

Coût :

Le coût est un facteur important pour les industries canadiennes lorsqu'elles prennent des décisions en matière d'infrastructure des TIC, par exemple lorsqu'elles doivent choisir entre construire des centres de données internes ou externaliser leur infrastructure informatique centralisée à des fournisseurs de services en nuage et décider quel fournisseur de services en nuage choisir et quel type de GPU utiliser pour l'apprentissage, les tests et le déploiement de modèles d'IA. Le coût est souvent pris en compte au même titre que les exigences techniques. Par exemple, une jeune entreprise spécialisée dans l'IA est susceptible de choisir le GPU le moins cher tout en étant capable d'exécuter son application dans un délai raisonnable, en fonction des besoins des utilisateurs. Quelle que soit la méthode d'accès, les coûts associés aux capacités de calcul pour l'IA peuvent être considérables. Les petites organisations, telles que les jeunes entreprises et les chercheurs universitaires disposant de ressources financières limitées, peuvent être exclues de l'écosystème d'innovation en matière d'IA en raison de leur manque d'accès à une capacité de calcul pour l'IA, même si elles ont des idées prometteuses et une expertise technique.

Emplacement géographique :

L'emplacement géographique des infrastructures informatiques est une considération courante pour les entreprises opérant dans des secteurs soumis à des exigences en matière de résidence des données, tels que les soins de santé. Pour les entreprises qui ne sont pas soumises à ces exigences, l'emplacement géographique est généralement moins important. Cependant, il peut tout de même constituer un facteur déterminant en raison de pratiques telles que l'informatique en périphérie, où le traitement et le stockage des données se font à proximité des utilisateurs finaux afin de réduire la latence et d'améliorer l'efficacité de la bande passante, ou pour des raisons liées à la durabilité environnementale, comme l'implantation d'infrastructures dans des régions qui utilisent des énergies propres.

Disponibilité de talents :

La disponibilité des talents est une contrainte courante qui limite la possibilité pour les organisations canadiennes d'accéder aux capacités de calcul pour l'IA. La création, la formation, la mise à l'essai, le déploiement et l'adoption de modèles d'IA nécessitent des compétences spécialisées. Dans ses commentaires à l'intention du CTIC, l'industrie souligne souvent les difficultés d'accès aux talents nécessaires pour fournir des capacités de calcul pour l'IA ainsi que pour concevoir, construire et mettre en œuvre des solutions d'IA.

Disponibilité du matériel :

La disponibilité du matériel peut limiter l'accès aux capacités de calcul pour l'IA, en particulier dans le cas des ressources informatiques hautement spécialisées ou rares. Par exemple, le CTIC a récemment interviewé un représentant d'une grande entreprise technologique canadienne, qui a déclaré avoir été placé sur de longues listes d'attente lorsqu'il a essayé de louer des processeurs graphiques H100 auprès de son fournisseur de services en nuage pour former un modèle d'IA.

Internet haut débit :

L'utilisation d'infrastructures d'IA à distance nécessite des réseaux à haut débit et à faible latence, essentiels au transfert et au traitement rapides des données dans les applications d'IA. De nombreuses régions rurales, éloignées et nordiques du Canada ne disposent pas actuellement d'une couverture haut débit à large bande²⁵. Sans accès à l'Internet haut débit, les organisations sont limitées dans leur capacité de tirer parti des infrastructures d'IA à distance, telles celles offertes par les fournisseurs de services en nuage, ce qui limite leur capacité de développer, déployer et adopter des produits et services d'IA.

PROPRIÉTÉ ÉTRANGÈRE ET DURABILITÉ ENVIRONNEMENTALE DES INFRASTRUCTURES DE CALCUL POUR L'IA

La question de savoir si l'infrastructure informatique est détenue et contrôlée par des intérêts canadiens n'est généralement pas une considération primordiale pour l'industrie canadienne lorsqu'elle prend des décisions concernant son infrastructure des TIC. De même, la durabilité environnementale n'est généralement pas une priorité dans les choix d'approvisionnement en infrastructure des TIC. Lorsque la durabilité est prise en compte, l'accent est mis généralement sur la

consommation d'énergie et les émissions de carbone plutôt que sur des impacts plus larges tels que l'utilisation de l'eau, l'utilisation des ressources naturelles ou la biodiversité²⁶. Néanmoins, la durabilité est une considération essentielle pour les investissements canadiens dans les capacités de calcul pour l'IA, car les sources d'électricité régionales et la renouvelabilité jouent un rôle important dans les dépenses énergétiques des centres de données, en fonction de leur emplacement²⁷.

²⁵ Des vitesses de téléchargement de 50 mégabits par seconde (Mbps) et des vitesses de chargement de 10 Mbps sont nécessaires pour prendre en charge les applications logicielles infonuagiques. En 2017, seuls 37 % des ménages ruraux avaient accès à des débits de 50/10 Mbps, contre 97 % des ménages urbains. Voir : « La haute vitesse pour tous : la stratégie canadienne pour la connectivité », Innovation, Sciences et Développement économique Canada, mars 2025, <https://ised-isde.canada.ca/site/accés-internet-haute-vitesse-canada/fr/strategie-canadienne-pour-connectivite/haute-vitesse-pour-tous-strategie-canadienne-pour-connectivite>.

²⁶ Kaitlyn Carr, Allison Clarke et Mairead Matthews, « Créer un écosystème de TIC durables : Stratégies et pratiques exemplaires pour réduire les effets néfastes sur l'environnement dans un monde numérique », Conseil des technologies de l'information et des communications (CTIC), janvier 2024, <https://ictc-ctic.ca/fr/rapports/creer-un-ecosysteme-de-tic-durables>.

²⁷ Victor Schmidt, et al., « Machine Learning CO2 Impact », consulté le 3 juillet 2025, <https://mlco2.github.io/impact/>.



NIVEAUX ET TYPES D'ACCÈS AUX CAPACITÉS DE CALCUL POUR L'IA REQUIS PAR L'INDUSTRIE POUR SE DÉVELOPPER ET PROGRAMMES DE SOUTIEN GOUVERNEMENTAUX EXISTANTS

Les organisations canadiennes qui adoptent des produits et services d'IA prêts à l'emploi peuvent souvent accéder à des capacités de calcul pour l'IA par l'intermédiaire de leurs fournisseurs de produits et services d'IA. Toutefois, si elles prévoient entraîner, affiner et distiller en interne des modèles d'IA prêts à l'emploi, elles devront avoir accès à leur propre infrastructure de calcul pour l'IA.

Les organisations qui développent leurs propres produits et services d'IA en interne ou en collaboration avec un fournisseur ont besoin d'une large gamme d'infrastructures de calcul pour l'IA afin de créer, former, tester et déployer des modèles d'IA. Cela comprend des

postes de travail personnels, tels que des ordinateurs ou des machines, une infrastructure informatique centralisée dotée de matériel de pointe, des cadres et bibliothèques d'IA, des modèles d'IA et des données d'apprentissage. Outre l'infrastructure, le secteur a besoin d'accéder à des talents qualifiés capables d'intégrer ces ressources dans des produits, des services et des applications industrielles.

Le tableau 1 présente un aperçu des programmes fédéraux existants qui soutiennent les intrants en IA et pourraient être mis à court terme pour améliorer l'accès à ces intrants.

Programme fédéral	Postes de travail personnels	Capacité de calcul centralisée (sur place)	Capacité de calcul centralisée (en nuage)	R&D en calcul avancé	Modèles d'IA	Ensembles de données	Talents en IA
<u>Fonds d'accès à une capacité de calcul pour l'IA</u>			✓				
<u>IA au service de la logistique</u>					✓	✓	✓
<u>Programme d'infrastructure de calcul souveraine pour l'IA</u>				✓			
<u>Projet pilote de connexion au nuage de l'Alliance</u>			✓				
<u>Défi « Informatique quantique appliquée »</u>				✓	✓		
<u>Défi « L'intelligence artificielle au service de la conception »</u>					✓	✓	✓
<u>Projet sur la base de données techniques sur la mer de Beaufort</u>						✓	
<u>Dépenses d'entreprise</u>	✓	✓	✓		✓	✓	✓
<u>Chaires de recherche du Canada</u>	✓	✓	✓		✓	✓	✓
<u>Réseau canadien des centres de données de recherche</u>						✓	
<u>CodeCan</u>							✓
<u>Subventions des IRSC, du CRSNG et du CRSH</u>	✓	✓	✓		✓	✓	✓
<u>Renouvellement du nuage communautaire</u>			✓				
<u>Ordinateurs pour les écoles et plus</u>	✓						

Programme fédéral	Postes de travail personnels	Capacité de calcul centralisée (sur place)	Capacité de calcul centralisée (en nuage)	R&D en calcul avancé	Modèles d'IA	Ensembles de données	Talents en IA
<u>Mise à niveau des centres de données</u>		✓	✓				
<u>Stratégie pour l'infrastructure de recherche numérique</u>		✓	✓				
<u>Compétences numériques pour les jeunes</u>							✓
<u>Renouvellement des capacités de calcul de haute performance</u>		✓					
<u>Programmes du Conseil national de recherches du Canada</u>				✓	✓	✓	✓
<u>Programme d'aide à la recherche industrielle du CNRC</u>							✓
<u>Programme emploi jeunesse du PARI CNRC</u>							✓
<u>Bourses d'études du CRSNG</u>							✓
<u>Portail du gouvernement ouvert</u>						✓	
<u>Fonds d'investissement stratégique pour les établissements postsecondaires</u>		✓					
<u>Centre de recherche en quantique et en nanotechnologies</u>				✓	✓		
<u>Fonds de soutien à la recherche</u>	✓	✓	✓	✓	✓	✓	✓
<u>Scale AI</u>					✓	✓	✓
<u>Programme de stages pratiques pour étudiants</u> (comme le programme AIT numérique du CTIC)							✓
<u>Soutien des chercheuses et chercheurs qui utilisent des ordinateurs quantiques</u>		✓					
<u>Initiative de perfectionnement des compétences pour l'industrie</u> (comme le programme Au-delà du nuage du CTIC)							✓

Tableau 1. Programmes fédéraux existants qui fournissent un soutien aux intrants en IA et qui pourraient être mis à profit pour améliorer l'accès à ces intrants à court terme. Analyse du CTIC basée sur des informations publiques.

SECTION II :

DÉVELOPPER UNE CAPACITÉ DE CALCUL SOUVERAINE ET DURABLE

Les questions relatives à la propriété et au contrôle nationaux ne sont généralement pas prises en considération par les entreprises individuelles lorsqu'elles prennent des décisions en matière d'infrastructure des TIC. Toutefois, le développement d'une capacité de calcul souveraine et durable peut être considéré comme un enjeu de sécurité économique et de résilience nationale.

L'entraînement de grands modèles d'IA nécessite bien souvent l'accès à de vastes ressources informatiques qui sont généralement hébergées par des fournisseurs étrangers de services en nuage (p. ex. AWS, Google Cloud, Microsoft Azure). Cela signifie que des données canadiennes sensibles pourraient transiter par des juridictions situées à l'extérieur du Canada ou y être stockées, ce qui les exposerait à des lois étrangères et à une surveillance potentielle. Lorsque les modèles sont entraînés ou affinés à l'aide de ressources de calcul étrangères, il existe un risque que les modèles, les données sur leur performance ou même leur architecture soient accessibles à des fournisseurs de services tiers, ce qui porterait atteinte à la propriété intellectuelle et à la confidentialité des données canadiennes.

Bien que le Canada soit un chef de file dans la recherche fondamentale et la formation de talents en IA, il accuse un retard en matière de capacité de calcul nationale pour l'IA. Selon l'Observatoire OCDE des politiques de l'IA, « les capacités de recherche et de calcul prêtes à l'emploi du Canada sont à la traîne non seulement par rapport à son voisin du sud, les États-Unis, mais aussi par rapport à tous les autres pays du G7 »²⁸. L'indice mondial de l'IA 2024 de Tortoise révèle que, si le Canada se classe au huitième rang mondial dans l'écosystème de l'IA, il est à la traîne en matière d'infrastructures par rapport aux leaders de l'IA tels que les États-Unis, la Chine, Singapour, le Japon et la Corée du Sud²⁹.

²⁸ Graham Dobbs et Jake Hirsch-Allen, « Canada's plans to bridge the AI compute gap and how it can make industry policy inclusive and sustainable », Observatoire OCDE des politiques de l'IA, 16 avril 2024, <https://oecd.ai/en/work/canadas-ai-compute-gap>.

²⁹ Tortoise, « Global AI Index », 2024, <https://www.tortoisemedia.com/intelligence/global-ai/>.

LES AVANTAGES DU DÉVELOPPEMENT D'UNE INFRASTRUCTURE DE CALCUL DE POINTE DÉTENUE ET CONTRÔLÉE PAR LE CANADA

Bien que la propriété canadienne et le contrôle national ne soient pas souvent pris en considération par les parties prenantes individuelles dans les décisions relatives aux infrastructures des TIC, ils peuvent offrir plusieurs avantages à l'échelle nationale :

Avantages économiques :

Les dépenses mondiales en infrastructures des TIC augmentent rapidement, les services en nuage devant à eux seuls atteindre 1,35 billion de dollars d'ici 2027, sous l'impulsion de l'informatique haute performance et de l'IA³⁰. Les pays qui possèdent et contrôlent les infrastructures de calcul et de stockage peuvent bénéficier d'une augmentation de leurs revenus, d'une croissance du PIB et de la création de nouveaux emplois. En 2023, le marché informatique canadien a généré environ 3,8 milliards de dollars US, soit 96 \$ US par habitant, contre 39,1 milliards de dollars US, soit 114 \$ US par habitant, pour le marché américain³¹. L'expansion des infrastructures détenues par des intérêts canadiens et l'utilisation de matériel et de logiciels canadiens dans les infrastructures nationales du Canada pourraient contribuer à équilibrer ces chiffres et à créer des opportunités économiques pour les entreprises informatiques canadiennes.

Sécurité nationale et sécurité publique :

Le secteur des TIC est l'un des dix secteurs d'infrastructure essentiels du Canada, vital pour la santé, la sécurité et le bien-être économique des Canadiens et pour le bon fonctionnement du gouvernement³². Les infrastructures informatiques sont au cœur du secteur des TIC et sont essentielles, au même titre que d'autres infrastructures telles que les services publics, les transports, l'alimentation et les finances. Une infrastructure informatique détenue et contrôlée par des intérêts canadiens renforce la capacité du Canada à fonctionner de manière indépendante, à réduire sa dépendance à l'égard des infrastructures détenues par des acteurs étrangers, à sécuriser les données et à développer des applications d'IA névralgiques dans un environnement souverain.

Durabilité environnementale :

Les infrastructures des TIC ont de nombreuses répercussions sur l'environnement tout au long de leur cycle de vie, notamment l'extraction des matières premières, la consommation d'énergie, les émissions, la consommation d'eau, la production de déchets, le changement d'affectation des terres et la pollution³³. Le développement d'infrastructures appartenant à des intérêts canadiens offre la possibilité de construire des installations informatiques et de stockage plus durables en tirant parti des sources d'énergie propres et en adhérant aux meilleures pratiques, telles que la norme L.1300 de l'Union internationale des télécommunications, intitulée Bonnes pratiques pour les centres de traitement de données écologiques³⁴.

30 « Worldwide Spending on Public Cloud Services is Forecast to Reach \$1.35 Trillion in 2027, According to New IDC Spending Guide, » August 2023 », août 2023, IDC, <https://www.idc.com/getdoc.jsp?containerId=prUS51179523>.

31 « Revenue of the computing market worldwide by country in 2023 », 2024 Statista, <https://www.statista.com/forecasts/758685/revenue-of-the-computing-market-worldwide-by-country>.

32 « Stratégie nationale sur les infrastructures essentielles », gouvernement du Canada, 2009, <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/srtg-crtcl-nfrstrctr/srtg-crtcl-nfrstrctr-fra.pdf>.

33 Kaitlyn Carr, Allison Clarke et Mairead Matthews, « Créer un écosystème de TIC durables : Stratégies et pratiques exemplaires pour réduire les effets néfastes sur l'environnement dans un monde numérique », Conseil des technologies de l'information et des communications (CTIC), janvier 2024, <https://ictc-ctic.ca/fr/rapports/creer-un-ecosysteme-de-tic-durables>.

34 « L.1300 : Bonnes pratiques pour les centres de traitement de données écologiques », UIT, juin 2014, <https://www.itu.int/rec/T-REC-L.1300-201406-l/fr>.



**Inclusion
numérique :**

Si les infrastructures informatiques détenues par des intérêts canadiens étaient plus rentables ou subventionnées, elles pourraient améliorer l'accès des personnes et des organisations qui en sont actuellement privées en raison de contraintes financières.

**Favoriser le
développement
d'un pipeline
d'innovation
national :**

Lorsque les ressources de calcul pour l'IA sont entièrement concentrées à l'étranger, les talents canadiens en IA peuvent migrer vers des pays et des entreprises offrant un meilleur accès aux ressources informatiques. La perte de talents hautement qualifiés en IA, notamment des technologues, des chercheurs et des entrepreneurs, équivaut à une fuite des cerveaux et met en péril le pipeline d'innovation du Canada en matière d'IA. Le fait de développer au Canada les ressources de calcul nécessaires pour l'IA peut être favorable au pipeline d'innovation national du pays.

**Fiabilité de
l'accès aux
capacités
de calcul et
atténuation
des risques
géopolitiques :**

Les différends commerciaux mondiaux et l'incertitude géopolitique actuels donnent à penser que l'accès transfrontalier fiable aux ressources en nuage, aux capacités de calculs pour l'IA et à d'autres infrastructures numériques pourrait ne plus aller de soi dans l'avenir. Le fournisseur de services financiers BlackRock note que le « découplage technologique mondial » et le « protectionnisme commercial mondial » devraient être deux facteurs de risque géopolitique jusqu'en 2025³⁵. Les infrastructures informatiques avancées détenues et contrôlées au niveau national peuvent contribuer à atténuer les risques de limitation ou de restriction d'accès aux infrastructures informatiques – tels que le blocage de l'accès aux services infonuagiques ou les contrôles à l'exportation de matériel informatique comme les processeurs graphiques – au-delà des frontières lors de futurs conflits commerciaux et autres crises géopolitiques. Cela permettrait de réduire la vulnérabilité économique nationale et d'atténuer les risques pour la cohésion économique si de telles situations venaient à se produire.

Malgré ces avantages, les jeunes entreprises et les PME canadiennes ont tendance à accorder plus d'importance aux exigences techniques, aux coûts, à l'emplacement géographique et à la disponibilité qu'à la propriété canadienne lorsqu'elles choisissent des infrastructures de TIC. Si les avantages nationaux d'une infrastructure détenue par des intérêts canadiens sont jugés importants pour la souveraineté nationale, des investissements publics ou d'autres incitations économiques pourraient être nécessaires pour promouvoir son utilisation.

Par exemple, le Canada a déjà mis en place certaines exigences légales en matière de résidence des données dans les appels d'offre³⁶. Bien que les exigences en matière de résidence des données au Canada ne soient pas généralisées et s'appliquent principalement aux activités du secteur public, elles suggèrent que des incitations ou des mandats ciblés pourraient être nécessaires pour encourager une adoption plus large des solutions basées au Canada.

³⁵ BlackRock, « Geopolitical risk dashboard », consulté le 23 mai 2025, <https://www.blackrock.com/corporate/insights/blackrock-investment-institute/interactive-charts/geopolitical-risk-dashboard>.

³⁶ Secrétariat du Conseil du Trésor du Canada, « Gouvernement du Canada livre blanc : souveraineté des données et nuage public », 28 juillet 2020, <https://publications.gc.ca/site/fra/9.858958/publication.html>.

OPPORTUNITÉS ET INITIATIVES À COURT TERME POUR AIDER À ÉLARGIR L'ACCÈS AUX INFRASTRUCTURES DE CALCUL EXISTANTES AU CANADA

Si l'investissement dans l'accès aux capacités de calcul pour l'IA pour les chercheurs et les entreprises canadiennes constitue un effort à long terme continu, des initiatives à court terme peuvent également influencer l'écosystème de calcul pour l'IA du Canada :

Soutien aux centres de données internes :

Certaines organisations peuvent préférer construire des centres de données internes plutôt que de faire appel à des fournisseurs de services en nuage en raison d'exigences techniques, d'avantages financiers ou d'autres considérations. Par exemple, les organisations dont les charges de travail en IA sont stables et prévisibles pourraient trouver plus rentable de construire des centres de données internes plutôt que de louer des processeurs graphiques spécialisés. Le soutien à ces organisations pourrait contribuer à étendre l'infrastructure informatique du Canada à long terme.

Encourager l'utilisation d'infrastructures appartenant à des entités canadiennes :

Étant donné que la propriété canadienne n'est généralement pas un facteur déterminant dans les décisions relatives aux TIC, les organisations peuvent ne pas choisir des options canadiennes en matière de calcul, même si celles-ci sont largement disponibles. Afin de garantir la viabilité à long terme de l'infrastructure de calcul pour l'IA, il doit exister une demande spécifique pour les installations appartenant à des entités canadiennes. Si cette demande ne se manifeste pas naturellement, des mesures incitatives peuvent être nécessaires, telles que des avantages financiers ou environnementaux. Les avantages environnementaux pourraient inclure l'accès à des sources d'énergie à faible émission de carbone comme l'hydroélectricité, la réduction des émissions ou des normes environnementales plus strictes, ce qui rendrait les installations canadiennes plus attrayantes pour les organisations ayant des objectifs de durabilité. Bien que la durabilité ne soit pas encore un facteur déterminant dans la plupart des décisions d'achat en matière de TIC, les avantages environnementaux pourraient gagner en importance s'ils étaient liés à des exigences réglementaires ou à des obligations en matière de rapports ESG (facteurs environnementaux, sociaux et de gouvernance)³⁷.

³⁷ Kaitlyn Carr, Allison Clark et Mairead Matthews, « Créer un écosystème de TIC durables : Stratégies et pratiques exemplaires pour réduire les effets néfastes sur l'environnement dans un monde numérique », CTIC, janvier 2024, <https://ictc-ctic.ca/fr/rapports/creer-un-ecosysteme-de-tic-durables>.

POSSIBILITÉS D'INTÉGRATION DU MATÉRIEL INFORMATIQUE ET DES LOGICIELS CANADIENS DANS UNE INFRASTRUCTURE SOUVERAINE DE CALCUL POUR L'IA

Le développement d'une infrastructure informatique nationale de calcul pour l'IA offre l'occasion de tirer parti de la stratégie industrielle. Au moment d'élaborer de nouvelles politiques en matière de capacité de calcul pour l'IA, le Canada devrait adopter une perspective stratégique industrielle en encourageant l'utilisation d'infrastructures détenues et contrôlées par des intérêts canadiens et en créant des possibilités d'approvisionnement pour les fournisseurs canadiens. Par exemple, les recherches en

cours menées par le CTIC soulignent que les entreprises canadiennes du secteur des semi-conducteurs ont développé de solides capacités de recherche et développement dans le domaine des technologies de semi-conducteurs de nouvelle génération pertinentes pour la production de puces spécialisées destinées à des applications informatiques avancées, telles que l'IA et l'informatique quantique³⁸.

MODÈLES DE COLLABORATION POUR SOUTENIR LES PARTENARIATS ENTRE L'INDUSTRIE ET LE MONDE UNIVERSITAIRE

L'industrie devance le milieu universitaire dans la recherche fondamentale sur l'IA, en partie en raison des disparités en matière de capacités de calcul³⁹. Si les établissements universitaires continuent de se heurter à des obstacles pour accéder aux capacités de calcul pour l'IA, il y a un risque que les acteurs privés dominent de plus en plus le programme de recherche en IA. Les partenariats de recherche publics-privés peuvent combler ce fossé en permettant aux chercheurs universitaires d'accéder aux ressources de calcul nécessaires pour l'IA tout en fournissant à l'industrie une expertise spécialisée en matière de recherche. De tels modèles de collaboration peuvent créer des occasions de « valeur partagée »⁴⁰ pour les établissements universitaires, l'industrie et les organisations à but non lucratif, chaque partenaire apportant ses propres atouts⁴¹. Afin de renforcer les capacités nationales de calcul pour l'IA, les partenariats de recherche publics-privés devraient donner la priorité à la coopération entre les partenaires du secteur privé national et les établissements universitaires canadiens.

Les modèles collaboratifs pourraient inclure des partenariats avec les principaux fournisseurs de services en nuage. Par ailleurs, les organisations du secteur privé pourraient fournir leurs propres ressources de calcul pour l'IA pour des projets collaboratifs publics-privés ou fournir des ressources financières pour l'acquisition de nouvelles capacités de calcul pour l'IA. Les établissements universitaires disposant d'une puissance de calcul spécialisée, comme les Sites d'hébergement nationaux de supercalculateurs de l'Alliance de recherche numérique du Canada, peuvent également contribuer à des projets conjoints entre le monde universitaire et l'industrie, en particulier pour les jeunes entreprises d'IA qui ont besoin d'accéder à des ressources informatiques puissantes afin de développer des preuves de concept pour de nouvelles technologies d'IA en vue d'attirer des investissements⁴².

³⁸ Voir « ICTC Partners with Georgetown's CSET to Map Semiconductor Industries in Canada and the United States », communiqué de presse, Conseil des technologies de l'information et des communications (CTIC), 14 novembre 2024, <https://ictc-ctic.ca/news-and-events/news-articles/ictc-partners-georgetown-cset-map-semiconductor-industries-canada>.

³⁹ Brian Eastwood, « Study: Industry now dominates AI research », MIT Sloan School of Management, 18 mai 2023, <https://mitsloan.mit.edu/ideas-made-to-matter/study-industry-now-dominates-ai-research>.

⁴⁰ Michael E. Porter et Mark R. Kramer, « Creating Shared Value », Harvard Business Review, janvier-février 2011, <https://hbr.org/2011/01/the-big-idea-creating-shared-value>.

⁴¹ Mark R. Kramer et Marc W. Pfitzer, « The Ecosystem of Shared Value », Harvard Business Review, octobre 2016, <https://hbr.org/2016/10/the-ecosystem-of-shared-value>.

⁴² « Sites d'hébergement nationaux », Alliance de recherche numérique du Canada, consulté le 16 août 2024, <https://alliancecan.ca/fr/services/calcul-informatique-de-pointe/la-federation/sites-dhebergement-nationaux>.

Étude de cas : Ressources nationales de recherche en IA aux États-Unis

Le National AI Research Resource (NAIRR) des États-Unis est un programme pilote de la National Science Foundation qui permet aux chercheurs universitaires et aux enseignants d'accéder à des ressources informatiques avancées, à des ensembles de données, à des modèles, à des logiciels et à un soutien pour les applications d'IA⁴³. Le programme s'est associé à 10 agences fédérales américaines et 25 entreprises privées, dont Nvidia, Microsoft, Meta et OpenAI, afin de développer les capacités nationales des États-Unis en matière d'IA, d'établir des partenariats de recherche et développement entre l'industrie, les établissements universitaires et la société civile, et de soutenir une recherche ouverte, sécurisée et axée sur l'enseignement dans le domaine de l'IA⁴⁴. Il comprend des volets consacrés à la recherche ouverte en IA (NAIRR Open), à la recherche sécurisée en IA avec des mesures de protection de la confidentialité (NAIRR Secure), aux logiciels d'IA et aux outils d'interopérabilité (NAIRR Software) ainsi qu'à l'IA en classe (NAIRR Classroom)⁴⁵.

Le NAIRR a été très bien accueilli par les établissements universitaires, ce qui met en évidence la forte demande en matière d'accès aux ressources de calcul pour l'IA. Cependant, il apporte aussi certains enseignements utiles. Un rapport publié en 2024 par la Brookings Institution recommande de créer des sources de financement distinctes pour les petits établissements universitaires et d'encourager les partenariats entre petites et grandes entités afin de remédier aux déséquilibres dans les ressources que ces deux groupes peuvent consacrer à la rédaction de propositions⁴⁶. Les bailleurs de fonds du programme doivent être conscients de ce déséquilibre des ressources et concevoir leurs appels à propositions en conséquence.

Les universités canadiennes ont également créé des écosystèmes d'innovation, de commercialisation et d'incubation d'entreprises, tels que le MaRS Discovery District à Toronto et le Creative Destruction Lab⁴⁷, afin d'aider les étudiants et les enseignants entrepreneurs à commercialiser de nouveaux produits et technologies. Faciliter l'accès aux capacités de calcul pour l'IA par l'entremise de fournisseurs privés ou d'un programme national public de capacités de calcul pour l'IA est un service précieux que ces écosystèmes pourraient offrir dans l'avenir.

DÉVELOPPER DES RELATIONS ET COLLABORER AVEC DES PAYS PARTAGEANT LES MÊMES IDÉES

Les partenaires internationaux du Canada sont confrontés à des défis similaires en matière de capacité de calcul pour l'IA, notamment la concentration des chaînes d'approvisionnement des TIC et de l'IA, les obstacles à l'accès à une capacité de calcul pour l'IA et les impacts environnementaux croissants des TIC. L'un des principaux partenaires commerciaux du Canada a récemment adopté une loi qui, pour la première fois, lie les systèmes d'IA à leurs impacts sur la durabilité environnementale en exigeant des fournisseurs de modèles d'IA à usage général qu'ils calculent et divulguent la consommation énergétique de leurs modèles et en encourageant le développement et l'utilisation des systèmes d'IA d'une manière durable sur le plan environnemental⁴⁸.

Une capacité de calcul souveraine canadienne pour l'IA pourrait être mise à profit dans les relations du Canada avec des pays partageant les mêmes valeurs afin d'aider à relever ces défis, en offrant à ses pairs un moyen de diversifier et d'accroître la résilience de leurs chaînes

d'approvisionnement et de leurs secteurs des TIC, de lever les obstacles à l'accès aux capacités de calcul pour l'IA et d'adopter des infrastructures des TIC conformes aux normes en matière de TIC durables. Comme le souligne le rapport conjoint du CTIC et de l'Association pour l'IA, les données et la robotique (Adra), intitulé Souveraineté de l'IA et croissance économique, une collaboration plus étroite entre le Canada et l'Union européenne pourrait tirer parti des atouts complémentaires des deux parties en matière de gouvernance, de recherche, de réglementation et d'applications industrielles. De tels partenariats sont essentiels pour renforcer à la fois la souveraineté en matière d'IA et la compétitivité à l'échelle mondiale, tout en favorisant des voies de développement responsables permettant des investissements conjoints dans l'informatique, le matériel de calcul, les infrastructures infonuagiques souveraines et l'innovation en matière d'énergie propre afin de répondre de manière durable aux besoins des systèmes d'IA à grande échelle⁴⁹.

⁴³ « Democratizing the future of AI R&D: NSF to launch National AI Research Resource pilot (communiqué de presse) », Fondation nationale pour la science des États-Unis, 24 janvier 2024, <https://new.nsf.gov/news/democratizing-future-ai-rd-nsf-launch-national-ai>.

⁴⁴ « The U.S. Just Took a Crucial Step Toward Democratizing AI Access », Time, 26 janvier 2024, <https://time.com/6589134/nairr-ai-resource-access/>.

⁴⁵ Fondation nationale pour la science des États-Unis, 2024, op. cit.

⁴⁶ Jennifer Wang et Mark Muro, « How the National Artificial Intelligence Research Resource can pilot inclusive AI », Brookings Institution, 9 juillet 2024, <https://www.brookings.edu/articles/how-the-national-artificial-intelligence-research-resource-can-pilot-inclusive-ai/>.

⁴⁷ « MaRS Discovery District », <https://www.marsdd.com/>; « Creative Destruction Lab (CDL) », <https://creativedestructionlab.com/>.

⁴⁸ « RÉGLEMENT (UE) 2024/1689 DU PARLEMENT EUROPÉEN ET DU CONSEIL », Journal officiel de l'Union européenne, 2024, https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:L_202401689.

⁴⁹ Amaya Garmendia, Todd Legere et Noah Lubendo, « Souveraineté de l'IA et croissance économique : renforcer le leadership transatlantique entre l'UE et le Canada », Association pour l'IA, les données et la robotique (Adra) / Conseil des technologies de l'information et des communications (CTIC), mai 2025, <https://ictc-ctic.ca/fr/rapports/souverainete-de-lia-et-croissance-economique>.



SECTION III :

PRIORITÉS À LONG TERME POUR L'INFRASTRUCTURE DE CALCUL

Investir dans l'infrastructure de calcul nationale nécessitera d'équilibrer les priorités et de répondre aux divers besoins des parties prenantes de l'écosystème de l'IA. Il sera essentiel de comprendre les besoins particuliers des différents acteurs de l'industrie et de la communauté scientifique afin de garantir que les solutions auront un impact positif et durable.

ÉLÉMENTS PRIORITAIRES POUR LE DÉVELOPPEMENT D'UNE INFRASTRUCTURE NATIONALE DE CALCUL POUR L'IA AU CANADA DANS LES CINQ PROCHAINES ANNÉES

La stratégie canadienne en matière d'infrastructure nationale de calcul pour l'IA devrait se concentrer sur les intrants pour l'IA qui peuvent être difficiles d'accès pour les entreprises, les organisations et les particuliers par d'autres moyens, que ce soit en raison de leur coût, de leur disponibilité ou d'autres contraintes. Le tableau 2 identifie les intrants de l'IA qui sont de sources ouvertes et librement accessibles, par opposition à ceux qui sont

de propriété exclusive et payants. Il donne à penser que, étant donné que les jeunes entreprises, les entreprises, les chercheurs et les étudiants canadiens ont généralement un bon accès aux cadres et bibliothèques d'IA, la stratégie devrait se concentrer sur le renforcement de la disponibilité et de l'accessibilité du matériel, des modèles, des données d'entraînement et des talents en matière d'IA. Ces éléments sont examinés plus en détail ci-dessous.

Intrant		Description	Modèle d'octroi de licences	Coût
Matériel d'IA	Postes de travail personnels	Postes de travail personnels, tels que des ordinateurs personnels dotés d'un matériel suffisamment puissant pour développer, tester et entraîner des modèles d'IA à petite échelle.	Propriété exclusive	Accès payant
	Infrastructure informatique centralisée	Centres de données privés ou plateformes d'infonuagique accessibles au public, équipés de processeurs GPU, ASIC, TPU ou FPGA haute performance pour l'entraînement et le déploiement de l'IA à des fins d'inférence, etc.	Propriété exclusive	Accès payant
Cadres et bibliothèques d'IA		Cadres (par exemple, TensorFlow, PyTorch, Keras) et bibliothèques (par exemple, Nvidia CUDA, Intel ONEAPI, Caffe) utilisés pour développer et déployer des modèles d'IA.	Source ouverte	Accès gratuit
Modèles d'IA		Modèles pré-entraînés et personnalisés pouvant être utilisés pour accomplir différents types de tâches.	Source ouverte et propriété exclusive	Accès gratuit et payant
Données d'entraînement de l'IA		Ensembles de données utilisés pour entraîner les modèles d'IA, qui peuvent être de propriété exclusive ou de source ouverte selon la source et le cas d'utilisation	Source ouverte et propriété exclusive	Accès gratuit et payant
Talents en IA		Personnes possédant les compétences nécessaires pour développer, tester, entraîner et déployer des modèles d'IA	Sans objet	Accès payant

Tableau 2. Caractéristiques des intrants de l'IA.

Composante I : Matériel d'IA

Les organisations qui développent et adoptent des solutions d'IA ont besoin d'un matériel sophistiqué suffisamment puissant pour développer, tester, entraîner et déployer des modèles d'IA. Les machines personnelles doivent être capables de traiter les modèles d'IA localement et à petite échelle avant de passer à la production. Une infrastructure informatique centralisée et puissante est nécessaire pour développer, tester, entraîner et déployer rapidement des modèles d'IA à grande échelle. Si les infrastructures de calcul pour l'IA détenues et contrôlées par des entreprises étrangères répondent aux besoins de bon nombre de jeunes entreprises, entreprises, chercheurs et étudiants canadiens, elles peuvent toutefois être coûteuses et inaccessibles pour certains en raison de contraintes budgétaires. Afin de garantir l'accès au matériel nécessaire, le Canada devrait explorer les moyens de réduire les obstacles tels que le coût, la disponibilité, l'emplacement géographique et l'accès à Internet. Les efforts visant à diversifier l'approvisionnement et à renforcer les capacités locales contribueraient à atténuer les risques liés aux goulets d'étranglement mondiaux dans le domaine des semi-conducteurs et à soutenir les investissements dans du matériel d'IA spécialisé au Canada.

Le Canada devrait également prioriser et encourager le développement d'une industrie locale des semi-conducteurs. S'il n'est peut-être pas viable de développer une industrie des semi-conducteurs aussi vaste que celle de Taïwan, de la Corée du Sud ou des États-Unis, le Canada peut renforcer ses capacités de fabrication de certains matériels, tels que les puces spécifiques à l'IA⁵⁰, en se concentrant sur la conception et la fabrication de circuits intégrés spécifiques à des applications (ASIC) adaptés aux tâches d'apprentissage automatique et d'apprentissage profond, ainsi qu'en finançant la recherche et les efforts commerciaux visant à développer des accélérateurs d'IA ou des technologies de semi-conducteurs de niche au Canada. Le Canada devrait également encourager les partenariats avec des chefs de file mondiaux de la conception de puces pour l'IA, tels que Nvidia et AMD, qui favorisent la production de puces haute performance pour les applications d'IA.

Composante II : Modèles d'IA

Les modèles d'IA sont à la base du développement de produits et de services d'IA. Les professionnels des TIC ont à leur disposition un vaste éventail de modèles d'IA capables d'effectuer des tâches telles que la détection d'événements, la reconnaissance, l'identification et la catégorisation de données, la prévision, la personnalisation, l'optimisation de systèmes, le raisonnement ainsi que la génération de contenu, y compris des sorties textuelles, audio et visuelles⁵¹.

Les modèles d'IA peuvent être de source ouverte, c'est-à-dire accessibles au public et gratuits, ou de propriété exclusive, c'est-à-dire payants pour être intégrés dans des produits et services. Si certains analystes considèrent les modèles de source ouverte comme risqués en raison de leurs mesures de protection limitées, d'autres voient des avantages dans leur possibilité de personnalisation, leur adaptabilité et leur capacité de fonctionner dans des environnements fermés sans partage de données⁵². De plus, les modèles de source ouverte permettent aux chercheurs non affiliés à de grandes entreprises technologiques de mener des recherches indépendantes sur l'IA, y compris sur la sécurité de l'IA⁵³.

La répartition entre modèles de source ouverte et de propriété exclusive varie selon les types d'IA. Jusqu'à récemment, la plupart des grands modèles de langage (GML) avancés étaient de propriété exclusive, comme les modèles GPT-4 d'OpenAI, Claude d'Anthropic et Gemini de Google DeepMind⁵⁴. Cependant, en juillet 2024, Meta AI a lancé Llama 3.1 405B, le premier GML de source ouverte capable de rivaliser avec les modèles fermés⁵⁵. En janvier 2025, la jeune entreprise DeepSeek a lancé DeepSeek-R1, un nouveau GML de source ouverte qui prétend utiliser moins de données d'apprentissage et donc moins de puissance de calcul que les autres applications d'IA de pointe actuelles⁵⁶.

Afin de favoriser la disponibilité des modèles d'IA pour les chercheurs et l'industrie canadiens, le gouvernement devrait continuer à financer la R&D axée sur l'IA par l'intermédiaire d'organismes tels que le Conseil national de recherches du Canada, le Conseil de recherches en sciences naturelles et en génie du Canada, les Chaires de recherche du Canada et le Pôle d'innovation mondial canadien Scale AI⁵⁷.

50 Voir : Saif M. Khan et Alexander Mann, « AI Chips: What They Are and Why They Matter », Center for Security and Emerging Technology (CSET), avril 2020, <https://cset.georgetown.edu/publication/ai-chips-what-they-are-and-why-they-matter/>.

51 « OECD Framework for the Classification of AI systems », OCDE, février 2022, https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html.

52 Will Knight, « Meta's New Llama 3.1 AI Model Is Free, Powerful, and Risky », WIRED, juillet 2024, <https://www.wired.com/story/meta-ai-llama-3/> ; « Introducing Llama 3.1: Our most capable models to date », Meta, juillet 2024, <https://ai.meta.com/blog/meta-llama-3-1/>.

53 « Introducing Llama 3.1: Our most capable models to date », Meta, juillet 2024, <https://ai.meta.com/blog/meta-llama-3-1/>.

54 Ibid.

55 Ibid.

56 « What is DeepSeek? The Chinese OpenAI rival sparking chaos in tech markets », CBC News, 27 janvier 2025, <https://www.cbc.ca/news/business/deepseek-ai-startup-1.7442382>.

57 Voir : Who we are », Scale AI, <https://www.scaleai.ca/about-us/>.

Composante III : Données d'apprentissage pour l'IA

Les données d'apprentissage sont essentielles pour permettre aux professionnels des TIC d'affiner les modèles d'IA destinés à des applications industrielles. Tout comme les modèles d'IA, les données d'apprentissage peuvent être de source ouverte ou de propriété exclusive. De récents litiges juridiques entre des entreprises d'IA et des fournisseurs de données de propriété exclusive soulignent l'importance de l'accessibilité des données d'apprentissage⁵⁸. Afin de soutenir l'innovation en matière d'IA, le gouvernement devrait continuer à développer des ensembles de données accessibles grâce à des initiatives telles que le Portail du gouvernement ouvert, le Réseau canadien des centres de données de recherche et le Projet sur la base de données techniques sur la mer de Beaufort.

Composante IV : Talents en IA

Les organisations qui développent et adoptent des solutions d'IA ont besoin d'accéder à des talents qualifiés possédant les connaissances et compétences nécessaires pour transformer les données d'IA en nouveaux produits, services et applications industrielles. Les organisations canadiennes ont souvent des difficultés à accéder aux compétences nécessaires pour travailler avec l'IA, qu'il s'agisse de concevoir, de développer ou d'adopter des solutions d'IA. Pour remédier à cette situation, le gouvernement devrait continuer à soutenir les programmes axés sur le développement des talents et des compétences en IA⁵⁹. Les compétences clés requises sont les suivantes :

- › Mise en place d'une infrastructure informatique centralisée
- › Fournir des ressources matérielles telles que des processeurs graphiques (GPU) à partir de centres de données internes et de fournisseurs de services en nuage
- › Assurer la gestion et l'ingénierie des données (par exemple, collecte, nettoyage, traitement et préparation des données)
- › Sélectionner des modèles d'IA en fonction de caractéristiques comme la qualité, la vitesse et le prix
- › Construire, entraîner, affiner et distiller des modèles d'IA
- › Prendre en compte les considérations liées à l'éthique, l'équité et la réduction des biais de l'IA afin de garantir un développement responsable
- › Sécuriser les systèmes d'IA et assurer de la confidentialité des données
- › Déployer des modèles d'IA et surveiller les performances
- › Utiliser des cadres et des bibliothèques d'IA pour intégrer l'IA dans des applications
- › Collaborer entre disciplines pour intégrer l'expertise sectorielle dans les solutions d'IA

Le développement de talents en IA possédant une expertise en calcul peut être soutenu par la création, dans les collèges et universités canadiens, de programmes de formation spécialisés axés sur l'infrastructure de l'IA, l'ingénierie des systèmes et l'informatique distribuée. Afin d'atténuer le problème de la fuite des cerveaux et de garantir que les chercheurs et ingénieurs canadiens en IA ne soient pas attirés par des écosystèmes de calcul riches situés à l'étranger, le gouvernement devrait offrir un financement compétitif et des incitations fiscales aux entreprises.

⁵⁸ « Getty Images v. Stability AI », 2024, BakerHostetler, <https://www.bakerlaw.com/getty-images-v-stability-ai/>; Dylan Walsh, « The Legal Issues Presented by Generative AI », MIT Sloan School of Management, août 2023, <https://mitsloan.mit.edu/ideas-made-to-matter/legal-issues-presented-by-generative-ai>.

⁵⁹ Maryna Ivus, Akshay Kotak et Ryan McLaughlin, « La nouvelle normalité axée sur le numérique », Conseil des technologies de l'information et des communications (CTIC), août 2020, <https://ictc-ctic.ca/fr/rapports/la-nouvelle-normalite-axee-sur-le-numerique>; Rosina Hamoni, Olivia Lin, Mairead Matthews et Peter Taillon, « Construire la future main-d'œuvre canadienne dans le domaine de l'intelligence artificielle », mars 2021, Conseil des technologies de l'information et des communications (CTIC), <https://ictc-ctic.ca/fr/rapports/construire-la-future-main-doeuv-re-canadienne-dans-le-domaine-de-lintelligence-artificielle>.

RECOMMANDATIONS

Outre les éléments susmentionnés, le développement de l'infrastructure nationale de calcul pour l'IA du Canada devrait être guidé par six principes : stratégie industrielle, possession et conservation de la propriété intellectuelle (PI), résilience de la chaîne d'approvisionnement, durabilité environnementale, régionalité et incitations fiscales et financières.

A. Stratégie industrielle :

Le Canada devrait tirer parti de sa stratégie industrielle lors de l'élaboration de son infrastructure nationale de calcul pour l'IA. Cela comprend la mise en œuvre de politiques axées sur la demande pour encourager le développement et l'utilisation d'infrastructures appartenant à des entreprises canadiennes ; la priorité aux fournisseurs canadiens lors du développement de la capacité de calcul nationale pour l'IA ; le renforcement des capacités de fabrication locales pour les puces spécifiques à l'IA ; la création de grappes nationales ou régionales en IA qui intègrent les ressources de calcul avec les laboratoires de recherche, les universités et les partenaires industriels ; la poursuite du soutien à la R&D en calcul pour l'IA ; et l'encouragement d'investissements substantiels dans le calcul haute performance (HPC), l'infrastructure infonuagique et les capacités de calcul quantique afin de soutenir l'avantage concurrentiel du Canada dans la course au développement de capacités de calcul pour l'IA. En outre, afin de préserver la compétitivité du Canada sur la scène mondiale, le Canada devrait investir dans la capacité de calcul nationale pour l'IA de manière proportionnelle par rapport aux économies comparables. Les engagements récents du Canada à augmenter ses dépenses pour la défense nationale pourraient être l'occasion de tirer parti des fonds alloués à la défense pour financer des infrastructures d'IA, ce qui aurait des retombées positives pour le reste de l'économie.

B. Possession et conservation de la propriété intellectuelle :

Malgré d'importants investissements publics dans l'IA, les inventeurs canadiens détiennent relativement peu de propriété intellectuelle dans ce domaine par rapport à ce qu'ils créent, et les jeunes entreprises technologiques de grande valeur sont souvent rachetées par des acheteurs étrangers ou quittent l'écosystème canadien⁶⁰. Le Canada risque de devenir un consommateur plutôt qu'un créateur de technologies fondamentales en IA, ce qui pourrait freiner le développement de la propriété intellectuelle nationale et limiter l'avantage concurrentiel dans les secteurs à forte intensité d'IA comme la santé, la technologie financière, la fabrication et les technologies climatiques.

Sans infrastructure nationale de calcul, les entreprises et les chercheurs canadiens spécialisés dans l'IA pourraient être contraints de s'appuyer sur des plateformes étrangères pour développer des modèles de pointe. Cela pourrait entraîner un transfert hors du Canada de la chaîne de valeur, notamment la propriété des modèles, l'expertise en optimisation et les technologies dérivées. Compte tenu des nouveaux investissements publics dans l'IA, il est essentiel d'intégrer les considérations relatives à la propriété intellectuelle dans la stratégie nationale afin d'éviter de perdre un avantage stratégique en la matière. En outre, l'adoption de lois plus strictes en matière de résidence des données pourrait contribuer à garantir le traitement à l'intérieur des frontières des données sensibles du Canada.

⁶⁰ Mairead Matthews et Faun Rice, « Le contexte est important : Renforcement de l'impact de l'investissement étranger sur l'innovation au Canada », 2022, Conseil des technologies de l'information et des communications (CTIC), <https://ictc-ctic.ca/fr/rapports/le-contexte-est-important>.

C. Résilience de la chaîne d'approvisionnement :

La fabrication et la production liées aux TIC reposent sur une chaîne d'approvisionnement complexe et hautement mondialisée⁶¹. Le matériel et les appareils de TIC sont fabriqués à partir d'une longue liste de matières premières, telles que l'indium, le lithium, le tantale, le gallium, le cuivre, l'argent, l'or et les terres rares, ainsi que de composants, notamment des cartes de circuits imprimés, des fonds de panier, des boîtiers, des câbles, des composants de machines de précision et des modules optiques⁶². L'extraction et la production de ces intrants se font à l'échelle mondiale, dont une grande partie dans des régions très éloignées géographiquement du Canada, comme la Chine et la République démocratique du Congo⁶³.

Au cours de la phase d'utilisation du cycle de vie des TIC, le matériel informatique situé dans les centres de données dépend également d'un approvisionnement énergétique important et fiable pour alimenter les appareils et assurer le fonctionnement des bâtiments, ainsi que de grandes quantités d'eau pour évacuer la chaleur dégagée par les équipements informatiques⁶⁴. Une étude de 2019 estime que les centres de données d'une capacité de TI de 15 MW consomment entre 0,8 et 1,3 million de litres d'eau par jour. Une autre étude a estimé qu'en 2014, rien qu'aux États-Unis, les centres de données ont consommé collectivement 165 milliards de gallons d'eau⁶⁵. En raison de ces dépendances, il est essentiel de renforcer les capacités nationales de production des intrants pour les TIC afin de réduire les risques liés à la chaîne d'approvisionnement.

D. Durabilité environnementale :

Le gouvernement devrait donner la priorité à la durabilité lors du développement d'infrastructures de calcul pour l'IA, en suivant des normes telles que la norme L. 1300 de l'Union internationale des télécommunications intitulée Bonnes pratiques pour les centres de traitement de données écologiques⁶⁶.

E. Régionalité :

La plupart des centres de données du Canada sont concentrés dans quelques provinces, soit l'Ontario, le Québec, la Colombie-Britannique et l'Alberta, et dans de grandes villes comme Toronto, Montréal, Vancouver, Mississauga, Winnipeg, Ottawa et Halifax. L'expansion des infrastructures devrait tenir compte des besoins des zones urbaines et rurales dans toutes les régions.

⁶¹ « Assessment of the Critical Supply Chains Supporting the US ICT Industry », février 2022, Département américain du commerce et Département américain de la sécurité intérieure, <https://www.commerce.gov/sites/default/files/2022-02/Assessment-Critical-Supply-Chains-Supporting-US-ICT-Industry.pdf>.

⁶² « Assessment of the Critical Supply Chains Supporting the US ICT Industry », Département américain du commerce et Département américain de la sécurité intérieure, février 2022, <https://www.commerce.gov/sites/default/files/2022-02/Assessment-Critical-Supply-Chains-Supporting-US-ICT-Industry.pdf>; « Digital Economy Growth And Mineral Resources : Implications for Developing Countries », décembre 2020, CNUCED, https://unctad.org/system/files/official-document/tn_unctad_ict4d16_en.pdf.

⁶³ Conférence des Nations Unies sur le commerce et le développement (CNUCED), « Digital Economy Growth And Mineral Resources : Implications for Developing Countries », Notes techniques de la CNUCED sur les TIC au service du développement n° 16, décembre 2020, https://unctad.org/system/files/official-document/tn_unctad_ict4d16_en.pdf.

⁶⁴ Kaitlyn Carr, Allison Clarke et Mairead Matthews, « Créer un écosystème de TIC durables : Stratégies et pratiques exemplaires pour réduire les effets néfastes sur l'environnement dans un monde numérique », Conseil des technologies de l'information et des communications (CTIC), janvier 2024, <https://ictc-ctic.ca/fr/rapports/creer-un-ecosysteme-de-tic-durables>.

⁶⁵ Ibid.

⁶⁶ Union internationale des télécommunications (UIT), « L.1300 : Bonnes pratiques pour les centres de traitement de données écologiques », Secteur de la normalisation des télécommunications de l'UIT, juin 2014, <https://www.itu.int/rec/T-REC-L.1300-201406-I/fr>.

F. Incitations fiscales et financières :

Des instruments fiscaux et d'autres incitations financières, telles que des subventions, pourraient être utilisés pour aligner les investissements privés nationaux dans les infrastructures de calcul pour l'IA sur la politique industrielle canadienne au sens large. Les grandes entreprises nationales, comme les institutions financières, les sociétés énergétiques et les entreprises de télécommunications, pourraient bénéficier d'incitations fiscales afin d'encourager les co-investissements publics-privés dans les infrastructures de calcul pour l'IA.

Des mécanismes fiscaux innovants, tels que les actions accréditatives, pourraient être adoptés pour les jeunes entreprises et les fournisseurs d'infrastructures d'IA canadiens afin d'encourager les investissements nationaux dans les infrastructures d'IA. Les actions accréditatives sont un instrument fiscal couramment utilisé dans les secteurs de l'énergie et des mines au Canada⁶⁷. D'autres instruments, tels que des remboursements d'impôt, pourraient être accordés aux jeunes entreprises canadiennes et à d'autres petites et moyennes entreprises afin de subventionner l'accès à des capacités de calcul pour l'IA fournies au niveau national. De telles politiques pourraient contribuer à la création d'une masse critique d'infrastructures de calcul pour l'IA basées au Canada.

CONCLUSION

Les jeunes entreprises et les PME canadiennes utilisent un large éventail d'infrastructures de calcul pour l'IA, notamment des postes de travail personnels, des centres de données centralisés, des services en nuage, des cadres, des modèles, des données et des talents pour développer et déployer des technologies d'IA novatrices. Malgré ses atouts initiaux en matière de recherche, de talents et d'adoption de l'IA, le Canada accuse un retard quant aux capacités de calcul pour l'IA, ce qui pourrait nuire à son potentiel de leadership technologique et mettre en péril sa prospérité future. Des économies comparables, tels les États-Unis, la France et l'Union européenne, dépassent de loin les investissements canadiens dans la capacité de

calcul souveraine. L'élargissement de la capacité de calcul du Canada pour entraîner et déployer des technologies d'IA contribuera à renforcer la souveraineté numérique du pays dans un environnement géopolitique incertain et en constante évolution.

Au moment où le Canada élabore des programmes d'accès public aux capacités de calcul pour l'IA et développe ses futures capacités souveraines de calcul pour l'IA, il est crucial de consulter les entreprises privées, les établissements universitaires et les organisations sans but lucratif afin de bien comprendre les besoins du pays en matière de calcul pour l'IA et ses liens à l'échelle mondiale.

⁶⁷ Voir : Actions accréditatives, gouvernement du Canada, dernière mise à jour le 28 avril 2008, <https://www.canada.ca/fr/agence-revenu/services/impot/entreprises/sujets/actions-accreditives.html>.

ICTC  **CTIC**
www.ictc-ctic.ca