ICTC CTIC

Renforcer la confiance dans l'intelligence artificielle grâce à une transparence accrue

Mémoire du Conseil des technologies de l'information et des communications présenté dans le cadre de la consultation publique du gouvernement du Canada sur la transparence de l'IA

23 septembre 2026

À propos du CTIC

Le Conseil des technologies de l'information et des communications (CTIC) est un centre d'expertise national, neutre et à but non lucratif dont la mission est de renforcer l'avantage numérique du Canada dans l'économie mondiale. Depuis plus de 30 ans, nous fournissons aux particuliers, aux entreprises et au secteur public des recherches tournées vers l'avenir, des conseils pratiques en matière de politique et des solutions novatrices de développement des capacités fondées sur l'expérience de l'industrie. L'objectif de notre équipe nationale d'experts répartis d'un océan à l'autre est de veiller à ce que la technologie soit utilisée pour stimuler la croissance économique et l'innovation et à ce que la main-d'œuvre canadienne reste compétitive à l'échelle mondiale.

Résumé

Le CTIC accueille favorablement la consultation du gouvernement du Canada sur les mesures visant à améliorer, à renforcer et à accroître la transparence en matière d'IA. Des exigences de transparence claires et proportionnées peuvent renforcer la confiance du public, améliorer l'adoption par les entreprises, soutenir la surveillance réglementaire et donner aux Canadiens et Canadiennes de véritables moyens de comprendre les utilisations de l'IA aux conséquences importantes et d'y réagir.

Afin d'appuyer le gouvernement du Canada dans ses efforts pour accroître la transparence de l'IA et établir les cadres de réglementation, de politiques et de gouvernance nécessaires au renforcement de la confiance envers l'IA, le CTIC a rédigé la présente réponse, qui porte sur quatre domaines précis de la consultation du gouvernement : l'interaction avec l'IA, les renseignements sur les systèmes d'IA, les incidents liés à l'IA et les agents d'IA.

La réponse du CTIC repose sur le principe selon lequel la transparence est la plus efficace lorsqu'elle est conçue en fonction d'un destinataire et d'un objectif définis. Une divulgation techniquement disponible, mais difficile à comprendre ou à exploiter, contribuera peu à réduire les préjudices, à améliorer la responsabilisation ou à renforcer la confiance des Canadiens et Canadiennes envers l'IA. Pour atteindre les objectifs du gouvernement du Canada, soit une meilleure compréhension de l'IA et une confiance accrue à son égard au sein de la population canadienne, le CTIC formule dix recommandations qui se renforcent mutuellement.

Le Canada devrait adopter un cadre fondé sur les risques et les rôles et adapté au contexte pour régir les exigences en matière de transparence de l'IA. Ces exigences devraient augmenter en fonction des conséquences possibles d'un système d'IA et du degré de pouvoir et de responsabilité délégué aux systèmes d'IA. Les développeurs, les déployeurs, les exploitants et les plateformes d'IA auront donc des responsabilités différentes, mais complémentaires. Le Canada devrait associer des exigences contraignantes pour les utilisations aux conséquences importantes avec des normes, des lignes directrices, des conditions d'approvisionnement, des travaux de recherche et le renforcement des capacités des utilisatrices et utilisateurs finaux, afin d'ancrer la transparence dans l'ensemble de l'industrie et de l'écosystème canadiens de l'IA. Le cadre devrait également distinguer la divulgation publique de l'information fournie aux personnes concernées, aux entreprises clientes, aux vérificateurs et vérificatrices et aux organismes de réglementation.

La position du CTIC s'appuie sur ses recommandations antérieures, qui préconisaient une gouvernance prévisible fondée sur les risques, la divulgation proactive, l'interopérabilité réglementaire, des mécanismes d'assurance accrédités, des bacs à sable réglementaires, des investissements dans la littératie en matière d'IA et un soutien concret aux petites et moyennes entreprises pour l'adoption, l'intégration et la gouvernance de l'IA. Pour les agents d'IA, le CTIC propose un cadre d'Autonomie et agentivité graduelles.

Un agent d'IA ne devrait se voir confier un pouvoir opérationnel ou économique accru qu'après avoir fait la preuve de sa compétence, de sa fiabilité, de sa sécurité et de sa responsabilisation dans l'environnement où

il fonctionne. Les agents actifs devraient disposer d'identités vérifiables, d'autorisations circonscrites, de journaux d'activité proportionnés, d'une délégation traçable et d'un mécanisme de révocation efficace, et relever d'une personne ou d'une organisation identifiable qui en demeure responsable.

En adoptant ces recommandations et en les déployant au moyen d'une combinaison coordonnée d'instruments de politique publique, le Canada peut accroître la confiance envers les systèmes et les agents d'IA tout en favorisant l'adoption par les entreprises, en stimulant la productivité et en améliorant la compétitivité et l'innovation des entreprises à l'échelle mondiale.

Résumé des recommandations

1. **Adopter un cadre fondé sur les risques et les rôles et adapté au contexte.** Les obligations en matière de transparence devraient refléter la fonction du système d'IA, son contexte de déploiement, le degré de pouvoir qui lui est confié et les conséquences possibles pour les personnes et les organisations.
2. **Créer un véritable droit d'être informé.** Les Canadiens et Canadiennes devraient recevoir un avis clair en cas d'interaction directe avec l'IA ou lorsque l'IA influe de façon substantielle sur une décision, un service ou un processus aux conséquences importantes les concernant.
3. **Normaliser les renseignements à plusieurs niveaux sur les systèmes d'IA.** Le Canada devrait élaborer une documentation adaptable qui réponde aux besoins des particuliers, des entreprises qui adoptent l'IA, des vérificatrices et vérificateurs, des chercheuses et chercheurs et des organismes de réglementation, sans exiger la divulgation publique sans discernement de renseignements sensibles.
4. **Établir un cadre coordonné de déclaration des incidents liés à l'IA et d'apprentissage.** Les incidents graves devraient être déclarés au moyen d'un cadre commun qui complète les régimes sectoriels existants, protège la confidentialité et publie des leçons et des divulgations anonymisées afin de renforcer la confiance.
5. **Mettre en place un cadre d'Autonomie et agentivité graduelles pour les agents d'IA.** Le pouvoir délégué aux agents d'IA devrait s'élargir ou se restreindre en fonction de la performance démontrée, du risque, du contexte et des conséquences des actions que l'agent d'IA est autorisé à entreprendre.
6. **Bâtir l'infrastructure de confiance de l'économie des agents d'IA.** La création et l'adoption à grande échelle d'agents d'IA ouvrent de nouveaux territoires. Le Canada devrait jouer un rôle de chef de file en matière d'identité et de mécanismes d'identification des agents d'IA, d'autorisation progressive, de protocoles de délégation, de preuves d'activité, d'assurance, d'interopérabilité et de révocation.
7. **Recourir aux normes, à l'assurance, aux bacs à sable réglementaires et à l'approvisionnement.** Les mesures législatives devraient s'appuyer sur des normes techniques, une vérification indépendante, une expérimentation supervisée et des exigences fédérales en matière d'approvisionnement, afin que les exigences de transparence soient proportionnées, pertinentes et suffisamment souples pour suivre le développement rapide de l'IA.
8. **Offrir aux PME des voies proportionnées pour stimuler l'adoption.** Des modèles de gouvernance, une assistance technique, des outils et des ressources partagés, des exigences échelonnées et des services-conseils de confiance devraient réduire les coûts de conformité tout en soutenant l'adoption responsable de l'IA.
9. **Investir dans la littératie en matière d'IA et la capacité de gouvernance.** Les travailleurs et travailleuses, les consommateurs et consommatrices, les entreprises et les institutions publiques doivent acquérir les compétences nécessaires pour interpréter les divulgations et y donner suite. Ces groupes ont également besoin d'une littératie de base en matière d'IA pour comprendre le fonctionnement de l'IA, le traitement de leurs données sensibles et privées et les mesures à prendre pour protéger leur vie privée tout en utilisant l'IA de façon sûre et responsable.
10. **Assurer la coordination entre les administrations et évaluer les résultats.** Les mesures fédérales devraient s'harmoniser avec les cadres provinciaux, territoriaux et internationaux et faire l'objet d'une évaluation portant sur la compréhension, la possibilité de contestation, la réduction des risques et la responsabilisation. Le Canada devrait se tourner vers des partenaires de confiance pour harmoniser les cadres réglementaires, afin d'améliorer l'interopérabilité et de réduire les obstacles pour les entreprises.

Principes pour une transparence efficace de l'IA

Le gouvernement du Canada reconnaît que la transparence constitue un point d'appui pour une prise de décisions éclairée et la responsabilisation. L'objectif de politique publique devrait donc être de rendre le rôle, les limites, les risques et la gouvernance de l'IA visibles, utiles et pertinents pour les personnes et les

organisations qui ont besoin de cette information, tout en préservant et en respectant la diversité des systèmes d'IA et des contextes de déploiement.

À cette fin, le CTIC propose plusieurs principes directeurs pour aider le gouvernement du Canada à élaborer des cadres de transparence de l'IA adaptés aux rôles et aux contextes.

Principe	Application
Proportionnalité	Les exigences en matière de transparence de l'IA devraient être proportionnées aux risques, aux contextes et aux rôles. Des exigences plus strictes devraient s'appliquer lorsque l'IA a une incidence sur les droits, la sécurité, l'emploi, les finances, l'accès aux services essentiels, les données sensibles ou les systèmes critiques.
Responsabilité selon le rôle	Chaque acteur de la chaîne de valeur de l'IA devrait être responsable de l'information dont il a la maîtrise. Les développeurs expliquent le système ; les déployeurs expliquent son utilisation ; les exploitants rendent compte des décisions prises et des résultats obtenus dans la pratique.
Divulgation valable	L'information devrait être communiquée en temps opportun, compréhensible, accessible et liée à une décision, à un droit, à un processus d'examen ou à une mesure corrective.
Accès à plusieurs niveaux	Les résumés publics, les avis aux utilisateurs et utilisatrices, la documentation destinée aux entreprises et les dossiers réglementaires confidentiels devraient présenter différents niveaux de détail.
Responsabilisation	La transparence devrait indiquer qui est responsable et comment une personne concernée peut obtenir un examen ou un recours.
Interopérabilité	Les mesures canadiennes devraient s'harmoniser avec les normes internationales et les grands marchés tout en protégeant les droits et les intérêts canadiens.
Adaptabilité	Les normes et les lignes directrices devraient être mises à jour en fonction de l'évolution des technologies, des données probantes et des pratiques de déploiement.

Ces principes concordent avec la recommandation antérieure du CTIC voulant que le Canada établisse une gouvernance de l'IA prévisible et proportionnée, appuyée par la divulgation proactive, une surveillance indépendante, des bacs à sable réglementaires et l'interopérabilité internationale. Ils reflètent aussi le point de vue du CTIC selon lequel la confiance du public et l'adoption de l'IA se renforcent mutuellement. Les entreprises sont plus susceptibles d'investir dans des technologies transformatrices lorsque les obligations sont claires, que les renseignements communiqués par les fournisseurs sont comparables et fiables et que la responsabilité est définie tout au long de la chaîne de valeur.

Interaction avec l'IA

Situations nécessitant une divulgation

Les Canadiens et Canadiennes devraient savoir lorsqu'ils communiquent avec un système d'IA susceptible d'être raisonnablement confondu avec une personne. Ils devraient également être informés lorsque l'IA influe de façon substantielle sur une décision ou un processus aux conséquences importantes, même en l'absence d'interaction directe avec le système. Cela comprend les décisions concernant l'emploi, le crédit et l'accès aux services financiers, l'assurance, les soins de santé, l'éducation, les prestations gouvernementales, les services juridiques et l'accès aux services essentiels.

Une exigence universelle de divulgation de chaque utilisation de l'IA à faible incidence entraînerait des coûts administratifs et une lassitude à l'égard des divulgations, sans avantage correspondant pour les utilisatrices et utilisateurs finaux. L'automatisation courante des processus internes ne devrait pas nécessiter d'avis individuel, à moins qu'elle ait une incidence substantielle sur une personne, qu'elle traite des renseignements sensibles ou qu'elle engendre un risque important. Le facteur déclencheur devrait reposer sur la nature et les conséquences de l'utilisation plutôt que sur la simple présence de l'IA ; par exemple, si un système d'IA joue un rôle important dans la prise d'une décision ayant une incidence sur une personne, la divulgation devrait être obligatoire afin que cette personne soit pleinement consciente de l'incidence de l'IA.

Le tableau ci-dessous présente plusieurs types d'interactions entre l'humain et l'IA, ainsi que l'exigence de divulgation recommandée par le CTIC pour chacun. Sans être exhaustifs, ces exemples correspondent à des cas d'usage courants d'interaction entre l'humain et l'IA.

Type d'interaction	Exigence recommandée
Communication directe s'apparentant à un échange humain, comme les agents conversationnels ou les assistants vocaux de service à la clientèle.	Indiquer au début de l'interaction que l'utilisateur ou l'utilisatrice communique avec l'IA. Laisser la divulgation accessible pendant toute l'interaction et offrir la possibilité de communiquer avec un être humain à la place.
Décision ou processus aux conséquences importantes, comme la présélection de candidatures à un emploi.	Indiquer le recours à l'IA, son rôle, le degré d'examen humain, les principaux renseignements pris en compte et la voie à suivre pour contester le résultat ou demander sa révision à un niveau supérieur.
Utilisation de renseignements personnels sensibles	Expliquer les modalités de collecte, d'utilisation, de conservation, de correction et de suppression des renseignements, sous réserve de la législation applicable sur la protection de la vie privée.
Utilisation interne à faible incidence	Aucun avis général, à moins que l'utilisation ait une incidence substantielle sur une personne ou crée un risque important en matière de vie privée, de sûreté, de droits ou de sécurité.

Responsabilités au sein de la chaîne de valeur de l'IA

Les déployeurs de systèmes d'IA assument la responsabilité première de la divulgation auprès des utilisateurs et utilisatrices, car ils déterminent l'objectif, le contexte, le public, les pratiques en matière de données et le degré de supervision humaine de chaque cas d'usage. Les développeurs d'IA devraient fournir aux déployeurs des renseignements exacts sur les capacités, les limites, les modes de défaillance et les méthodes de divulgation appropriées. Les employeurs et employeuses ainsi que les autres exploitants devraient également divulguer les utilisations organisationnelles aux conséquences importantes. Les plateformes devraient soutenir des normes communes, techniques et de présentation, lorsque les interactions passent par leur infrastructure.

Définir la divulgation valable

- Le gouvernement du Canada devrait envisager de définir la divulgation valable afin d'établir des normes applicables à l'ensemble des cas d'usage, des systèmes et des contextes. Une divulgation valable devrait donner à la personne qui la reçoit les moyens d'adapter son comportement, de vérifier et de corriger l'information, d'obtenir l'aide d'un être humain ou de contester un résultat. Une déclaration générale selon laquelle une organisation utilise l'IA quelque part dans ses activités ne répond pas à cette norme. Une divulgation valable peut notamment consister à fournir un avis clair avant ou au début de l'interaction avec l'IA, plutôt que de l'enfouir dans les conditions d'utilisation ;
- employer un langage clair et un format adapté à l'interface, y compris un avis oral pour les assistants vocaux ;
- indiquer le rôle du système d'IA et préciser s'il peut prendre des engagements ou des décisions au nom de l'organisation ;
- préciser si des renseignements personnels sont utilisés et comment ils sont protégés ;
- décrire le niveau de supervision humaine et offrir une voie concrète d'examen humain lorsque cela est justifié ;
- rendre les avis accessibles, les offrir dans les deux langues officielles et les mettre à l'essai auprès des utilisatrices et utilisateurs visés.

Les organisations devraient demeurer responsables des systèmes d'IA qu'elles mettent en place pour communiquer avec le public. Une entreprise ne devrait pas pouvoir se soustraire à sa responsabilité à l'égard des déclarations faites par son agent conversationnel autorisé au seul motif que le système a généré une réponse erronée. Une responsabilisation claire incitera également à définir et à délimiter le pouvoir du système, à surveiller sa performance et à prévoir un transfert à un être humain en temps opportun.

Mesures recommandées

- Établir, à titre de règle de base, le droit de toute personne de savoir qu'elle communique avec un système d'IA susceptible d'être raisonnablement confondu avec un être humain.
- Exiger des avis renforcés et des mécanismes d'examen humain pour les utilisations aux conséquences importantes.
- Élaborer un modèle de divulgation à plusieurs niveaux comprenant un court avis, une explication concise du rôle du système, l'accès à des renseignements détaillés et, s'il y a lieu, une voie de révision à un niveau supérieur.
- Publier des lignes directrices sectorielles et des modèles réutilisables, y compris des outils simplifiés pour les PME.
- Définir une norme de divulgation valable afin d'assurer la cohérence entre les cas d'usage.

Renseignements sur les systèmes d'IA

Les renseignements devraient refléter les besoins du destinataire et le cas d'usage

Le Canada devrait promouvoir des renseignements sur les systèmes d'IA qui sont normalisés, à plusieurs niveaux et adaptés aux différents publics. L'objectif devrait être de fournir aux utilisateurs et utilisatrices les renseignements pertinents dont ils ont besoin pour évaluer l'adéquation du système, les risques et la responsabilisation, plutôt que de maximiser le volume de renseignements divulgués.

Comme l'indique le document de travail du gouvernement du Canada, les renseignements sur les systèmes d'IA comprennent deux considérations générales : ce qu'un système d'IA peut et ne peut pas faire, et la façon dont il a été conçu, y compris les données sur lesquelles il a été entraîné. Afin de soutenir la transparence de l'IA et de bâtir la confiance, les renseignements sur ces deux considérations devraient être

mis à la disposition des destinataires et s'appuyer sur un cadre de documentation commun qui en assure la cohérence, l'accessibilité et la pertinence.

Un cadre de documentation partagé

Une documentation normalisée faciliterait la comparaison des systèmes, réduirait les coûts d'approvisionnement et aiderait les petites organisations à exercer leur vérification diligente. Les modèles devraient demeurer adaptables aux différentes technologies et aux différents secteurs. Le Canada devrait aussi envisager d'élaborer des modèles de documentation communs avec des partenaires de confiance, comme l'Union européenne, afin d'assurer l'interopérabilité et de réduire les obstacles au commerce.

Le cadre de documentation peut être structuré par niveaux en fonction de l'incidence relative des systèmes d'IA, des cas d'usage et des besoins du destinataire visé. Pour les systèmes à incidence élevée, la documentation devrait être approfondie et porter sur les éléments suivants :

- Identité du système, version, fournisseur et organisation responsable ;
- Usages prévus, usages appropriés et restrictions substantielles ;
- Capacités, performance, limites et modes de défaillance connus ;
- Performance auprès des populations pertinentes et dans les conditions de déploiement pertinentes ;
- Sources de données, pratiques en matière de provenance et lacunes substantielles de représentativité ;
- Traitement des données personnelles, exclusives, protégées par le droit d'auteur et autochtones ;
- Mise à l'essai en matière de sûreté, de sécurité, de protection de la vie privée et de biais ;
- Exigences en matière de supervision humaine, de surveillance et de mise à jour ;
- Personnes-ressources et procédures de déclaration des incidents ;
- Changements importants après le déploiement ;
- État de l'évaluation indépendante, de la certification ou de l'assurance.

Données d'entraînement et renseignements confidentiels

Une véritable transparence des données d'entrée ne nécessite pas la divulgation de chaque élément d'un ensemble de données d'entraînement. Les développeurs d'IA devraient toutefois divulguer les principales catégories et sources de données d'entraînement, les périodes de collecte et la couverture géographique pertinentes, la présence ou non de renseignements personnels ou de contenu protégé par le droit d'auteur, les limites substantielles en matière de représentativité et les processus offerts aux personnes et aux titulaires de droits pour faire part de leurs préoccupations. Les développeurs d'IA devraient également indiquer si les données des utilisatrices et utilisateurs finaux serviront à poursuivre l'entraînement des systèmes et des modèles d'IA. Les organisations en aval, comme les déployeurs d'IA, devraient recevoir suffisamment de renseignements pour évaluer l'adéquation du système et le risque juridique.

Le Canada devrait faire la distinction entre la confidentialité légitime et l'information nécessaire pour protéger les droits, évaluer la sûreté ou soutenir la surveillance réglementaire. La documentation publique peut demeurer générale au besoin, tandis que les organismes de réglementation et les vérificatrices et vérificateurs accrédités reçoivent des renseignements plus détaillés, assortis de protections appropriées en matière de confidentialité et de sécurité.

Mesures recommandées

- Élaborer des modèles de documentation normalisés et adaptables pour les modèles et les systèmes d'IA, harmonisés avec les normes internationales ;
- Fixer des exigences de documentation plus strictes pour les systèmes à incidence élevée et les systèmes qui traitent des renseignements sensibles ;
- Établir des droits minimaux à l'information pour les entreprises qui se procurent des systèmes d'IA tiers ;
- Exiger un avis lorsqu'une mise à jour importante modifie le risque, la performance ou l'usage approprié d'un système ;

- Élaborer des voies d'accréditation et d'assurance indépendante pour les applications à risque plus élevé ;
- Recourir à l'approvisionnement fédéral pour exiger de la documentation, des preuves de mise à l'essai, des avis de mise à jour et une clarté contractuelle ;
- Offrir aux PME et aux organismes à but non lucratif des outils partagés de vérification diligente, des lignes directrices en matière d'approvisionnement et des services-conseils de confiance.

La documentation ne suffit pas à elle seule ; elle doit s'accompagner d'un renforcement des capacités organisationnelles afin que les entreprises, les personnes et les organisations soient en mesure d'interpréter la documentation et de comprendre comment elle s'applique à l'usage qu'elles prévoient. Les entreprises ont besoin de personnel capable d'évaluer les affirmations des fournisseurs, d'intégrer les systèmes de façon responsable, de surveiller leur performance et de réagir aux incidents. L'intégration de formations en IA propres à chaque secteur, le soutien à l'adoption par les PME et le renforcement de la littératie en matière d'IA au sein de la main-d'œuvre peuvent faciliter l'utilisation de la documentation.

Incidents liés à l'IA

Un cadre coordonné de déclaration et d'apprentissage

Le Canada ne dispose pas encore d'une vue d'ensemble systématique des incidents liés à l'IA et des tendances émergentes en matière de risques. Les régimes de déclaration des incidents en place pour les instruments médicaux, les transports, les produits de consommation, la protection de la vie privée et la cybersécurité recensent certains événements, mais ils reposent sur des définitions différentes et n'ont pas été conçus pour offrir une compréhension intersectorielle des risques liés à l'IA.

Le Canada devrait établir un cadre coordonné qui complète ces régimes et évite les déclarations en double. Ce cadre devrait favoriser la responsabilisation à l'égard des incidents graves et l'apprentissage collectif à partir d'événements de moindre gravité. La déclaration obligatoire convient lorsque le préjudice est grave ou imminent. Une déclaration volontaire protégée peut encourager les organisations à faire part des incidents évités de justesse et des défaillances maîtrisées qui demeureraient autrement invisibles, tout en améliorant la compréhension et la confiance dans l'ensemble de l'écosystème.

Définition d'un incident grave lié à l'IA

Un incident grave lié à l'IA ou un incident évité de justesse devrait se définir par ses conséquences réelles et possibles. Les événements à déclarer devraient comprendre toute défaillance, utilisation abusive, tout acte malveillant, toute atteinte à la sécurité ou défaillance de gouvernance liés à l'IA qui causent l'un des éléments suivants, y contribuent de façon substantielle ou créent un risque imminent à cet égard :

- Décès ou atteinte grave à la santé ou à la sécurité ;
- Atteinte importante aux droits de la personne ou aux droits constitutionnels ;
- Discrimination importante ;
- Perte financière ou fraude importante ;
- Accès non autorisé à des renseignements sensibles ou divulgation non autorisée de ceux-ci ;
- Atteinte à des infrastructures critiques ou à des services essentiels ;
- Perte d'un véritable contrôle humain sur un système aux conséquences importantes ;
- Dommages graves à l'environnement ou aux biens ;
- Manipulation ou perturbation à grande échelle des processus démocratiques et civiques ;
- Préjudice systémique touchant une population importante ;
- Exercice ou délégation non autorisés du pouvoir d'un agent d'IA.

La déclaration des incidents liés à l'IA et des incidents évités de justesse devrait être structurée par niveaux selon la gravité, la probabilité et l'ampleur des conséquences. Par exemple, le Canada pourrait adopter un système à plusieurs niveaux qui classe les incidents liés à l'IA selon les catégories suivantes : critique, grave, incident évité de justesse et tendance opérationnelle. Le tableau ci-dessous illustre la façon dont ces niveaux pourraient être différenciés et l'exigence de déclaration correspondante.

Niveau	Événement représentatif	Approche en matière de rapports
Critique	Décès, blessure grave, perturbation d'infrastructures critiques, violation majeure de droits ou atteinte importante à la sécurité	Déclaration obligatoire, rapide et confidentielle à l'autorité compétente
Grave	Préjudice financier substantiel, résultat discriminatoire, atteinte majeure à la vie privée ou action non autorisée aux conséquences importantes	Déclaration obligatoire dans un délai déterminé
Incident évité de justesse	Défaillance maîtrisée, tentative d'exploitation ou erreur détectée avant qu'un préjudice survienne	Déclaration sans mise en cause de responsabilité
Tendance opérationnelle	Erreurs récurrentes, tendances dans les plaintes ou variations de la fiabilité	Déclaration périodique de données agrégées

Information et responsabilité

Le gouvernement du Canada devrait également établir des normes régissant l'information divulguée dans les déclarations d'incidents liés à l'IA, selon la gravité de l'incident. Les déclarations devraient préciser le système et sa version, le développeur, le déployeur, l'exploitant, le contexte de déploiement, les parties touchées, le moment et la portée de l'incident, la nature du préjudice, les modalités de supervision, les mesures de confinement, la cause sous-jacente, lorsqu'elle est connue, et les mesures correctives. Les informations sensibles sur le plan de la sécurité devraient être divulguées au moyen de processus coordonnés qui n'exposent pas une vulnérabilité non corrigée.

La responsabilité de la déclaration devrait dépendre de la connaissance et du contrôle de chaque acteur. Les déployeurs et les exploitants sont souvent les mieux placés pour repérer les incidents dans la pratique. Les développeurs sont mieux placés pour repérer les défaillances ou les vulnérabilités au niveau du modèle susceptibles de toucher plusieurs déploiements. Tous deux devraient faire une déclaration lorsque chacun détient de l'information substantielle ; les contrats devraient exiger le partage d'information en temps opportun tout au long de la chaîne de valeur.

Conception institutionnelle

Le Canada devrait établir une fonction fédérale de coordination ou un organisme de confiance chargé des déclarations, qui achemine les cas vers l'organisme de réglementation approprié et collabore avec l'Institut canadien de la sécurité de l'intelligence artificielle, les autorités de protection de la vie privée et le Centre canadien pour la cybersécurité. Le cadre devrait reposer sur une taxonomie commune et donner lieu à la publication de rapports anonymisés sur les tendances, d'alertes de sécurité et de leçons tirées. Un organisme consultatif multipartite pourrait contribuer à la mise à jour des définitions, des seuils et des normes de déclaration à mesure que les données probantes évoluent.

Des protections appropriées en matière de confidentialité et d'exonération de responsabilité devraient encourager la déclaration de bonne foi, sans toutefois mettre à l'abri la négligence, la dissimulation ou l'inconduite. Le Canada devrait aussi harmoniser sa taxonomie et ses formats de déclaration avec des systèmes internationaux de confiance afin de réduire le dédoublement des efforts de conformité et d'améliorer l'apprentissage transfrontalier, tout en prévoyant des mesures de confidentialité limitées pour réduire le risque de révéler des vulnérabilités systémiques à des acteurs hostiles.

Mesures recommandées

- Recenser les régimes de déclaration fédéraux, provinciaux et territoriaux existants avant d'introduire de nouvelles exigences ;
- Élaborer une taxonomie canadienne commune des incidents liés à l'IA et un formulaire de déclaration harmonisé ;
- Exiger la déclaration des incidents graves et soutenir la déclaration protégée des incidents évités de justesse ;
- Clarifier les protections en matière de confidentialité, de responsabilité et de déclaration de bonne foi ;
- Exiger des mesures correctives et un avis lorsqu'une vulnérabilité peut toucher d'autres déploiements ;
- Publier des constatations anonymisées et des tendances en matière de risques à l'attention de l'industrie, des chercheurs et chercheuses et du public ;
- Mettre à l'essai les processus de déclaration dans des bacs à sable réglementaires et certains secteurs à incidence élevée ;
- Fournir des outils simplifiés et des lignes directrices aux PME.

Agents d'IA

Autonomie et agentivité graduelles

Les agents d'IA nécessitent une approche distincte au sein des cadres canadiens de transparence de l'IA, car ils peuvent agir, utiliser des outils, accéder à des données, effectuer des transactions et déléguer des tâches avec une plus grande autonomie que d'autres systèmes d'IA. Le risque ne dépend pas uniquement de la sophistication du modèle sous-jacent. Il découle du pouvoir accordé à l'agent, des actifs et des systèmes auxquels il a accès, du degré d'action indépendante permis et des conséquences d'une erreur ou d'une utilisation abusive.

Le CTIC propose un cadre d'Autonomie et agentivité graduelles comme principe organisateur de la gouvernance des agents d'IA. L'octroi d'un pouvoir opérationnel ou économique accru devrait découler de la compétence, de la fiabilité, de la sécurité et de la responsabilisation démontrées dans l'environnement d'exploitation de l'agent.

Exemples d'étapes d'attribution du pouvoir

Étape	Exemple de pouvoir	Contrôles nécessaires
Observation	Récupérer ou résumer de l'information	Contrôles de base en matière d'identité, de vie privée, de sécurité et d'accès
Recommandation	Proposer une action	Divulgation, renseignements sur les sources et examen humain
Agir avec confirmation	Préparer une transaction nécessitant une approbation	Authentification forte, confirmation claire et preuves d'activité

Étape	Exemple de pouvoir	Contrôles nécessaires
Action déléguée limitée	Acheter auprès de fournisseurs approuvés en deçà d'un certain seuil	Identité vérifiée, mécanismes d'identification à portée limitée, limites, surveillance et révocation
Action organisationnelle élargie	Exécuter des transactions plus importantes ou modifier des systèmes approuvés	Assurance indépendante, surveillance continue, séparation des tâches et déclaration des incidents
Délégation multiagent	Déléguer une tâche définie à un autre agent approuvé	Registres de la chaîne de délégation, limites de délégation, règles relatives aux mécanismes d'identification et propagation de la révocation

Le passage aux étapes suivantes ne devrait pas être automatique. Le pouvoir devrait plutôt s'élargir, se restreindre, expirer ou être révoqué lorsque le contexte, la performance, le comportement du système, les mécanismes d'identification ou le risque changent. De plus, une mise à jour substantielle du modèle devrait entraîner une réévaluation lorsqu'elle est susceptible de modifier le comportement ou la fiabilité de l'agent.

Le cadre de confiance pour les agents actifs sur le plan économique

Identité vérifiable : Un agent actif devrait disposer d'une identité vérifiable qui établit ce qu'est l'agent, son fournisseur et sa version, la personne ou l'organisation au nom de laquelle il agit et l'état de ses mécanismes d'identification et de son pouvoir. Les exigences devraient être plus strictes lorsque les agents exercent un contrôle direct sur des tâches, des données ou de l'information de nature sensible, par exemple pour accéder à des renseignements sensibles, conclure des ententes, modifier des systèmes critiques ou interagir avec d'autres agents.

Pouvoir progressif : Les autorisations devraient être explicites, limitées à une fin précise et, s'il y a lieu, limitées dans le temps, et être encadrées selon la valeur de la transaction, le type de données, le fournisseur, le système ou l'environnement d'exploitation. Le pouvoir devrait pouvoir faire l'objet d'un examen, d'un élargissement, d'une restriction, d'une suspension et d'une expiration. La capacité technique d'un agent à effectuer une action ne devrait pas être considérée comme une autorisation de l'effectuer.

Preuve d'activité : Les organisations devraient conserver suffisamment de preuves pour constituer un historique d'activité et établir quel agent a agi, au nom de qui, en vertu de quelle autorisation, quelle action aux conséquences importantes il a entreprise, à quel moment il a agi, avec quel système ou quel agent il a interagi, si une approbation était nécessaire et a été obtenue et quel a été le résultat. Cela ne devrait pas exiger la conservation sans discernement des requêtes, des renseignements personnels, de l'information exclusive ou du raisonnement interne du modèle. L'objectif est d'assurer la traçabilité des actions aux conséquences importantes aux fins de surveillance continue, d'amélioration et de déclaration en cas d'incident.

Chaîne de délégation : Lorsqu'un agent délègue à un autre, les registres devraient indiquer la personne ou l'organisation responsable à l'origine, l'agent d'origine, le pouvoir initialement accordé, chaque délégation subséquente, toute réduction ou extension de la fin ou de la durée, ainsi que l'agent ayant effectué l'action finale. Un agent délégué ne devrait jamais recevoir un pouvoir plus étendu que celui de l'agent délégant.

Révocation : Les mécanismes d'identification et les autorisations doivent pouvoir faire l'objet d'une restriction, d'une suspension, d'une expiration ou d'une révocation. La révocation devrait se propager dans les chaînes de délégation et être possible lorsque le comportement change, que les mécanismes d'identification expirent, qu'un système fait l'objet d'une mise à jour substantielle, qu'une atteinte à la sécurité survient ou qu'un incident révèle que le pouvoir existant n'est pas sûr.

Responsabilisation humaine et organisationnelle : Chaque agent actif aux conséquences importantes devrait être rattaché à une personne ou à une organisation responsable identifiable. La partie responsable devrait définir le pouvoir de l'agent, surveiller sa performance, réagir aux incidents, offrir des recours, gérer la délégation et révoquer l'accès le cas échéant. Les développeurs d'IA devraient demeurer responsables des défauts, des déclarations trompeuses et des défaillances relevant de leur contrôle, tandis que les déployeurs d'IA devraient demeurer responsables des choix de déploiement, des autorisations, de la surveillance et de l'utilisation organisationnelle. Fait important, l'accroissement de l'autonomie d'un agent d'IA ne devrait pas faire disparaître la responsabilisation humaine.

Un modèle de risque fondé sur le pouvoir et les conséquences

La supervision devrait être adaptée en fonction de quatre facteurs :

- **Pouvoir** : les actions que l'agent est autorisé à entreprendre
- **Actifs** : l'argent, les données, les systèmes, les infrastructures ou les droits auxquels il peut accéder
- **Autonomie** : sa capacité d'agir sans confirmation humaine au moment de l'action
- **Conséquences** : le préjudice pouvant résulter d'une erreur, d'une atteinte à la sécurité ou d'une utilisation abusive

Un agent qui résume des documents devrait être soumis à des exigences moins strictes qu'un agent autorisé à dépenser un montant limité auprès de fournisseurs approuvés ; ce dernier devrait disposer d'une identité vérifiée, de registres de transactions, de limites définies et d'un mécanisme de révocation simple. Un agent pouvant transférer des fonds importants, accéder à des renseignements sur la santé, conclure des contrats, modifier des systèmes critiques ou déléguer son pouvoir devrait faire l'objet d'exigences plus strictes en matière d'assurance, de surveillance, d'approbation et de déclaration des incidents, avec une supervision humaine intégrée à des points de contrôle réguliers.

L'occasion pour le Canada : l'infrastructure de confiance des agents

Le développement des agents d'IA et leur utilisation croissante nécessitent l'élaboration de normes et de cadres réglementaires cohérents et communs. Le Canada peut chercher à jouer un rôle de chef de file dans le développement de l'infrastructure réglementaire qui donnera aux organisations et aux marchés les moyens de faire confiance aux agents d'IA. Cela peut comprendre l'identité et les mécanismes d'identification des agents, l'autorisation progressive, les protocoles de délégation, les journaux d'activité, l'assurance indépendante, l'interopérabilité et la révocation. Ces capacités seront nécessaires, que les modèles sous-jacents soient développés au Canada ou à l'étranger.

Une infrastructure de confiance canadienne pour les agents serait conforme aux objectifs du Canada en matière de souveraineté numérique, de cybersécurité, de commercialisation, de gouvernance responsable et de participation à l'élaboration de normes internationales. Elle pourrait aussi créer des occasions pour les entreprises canadiennes dans les domaines de l'identité numérique, des technologies d'amélioration de la confidentialité, des paiements, de la cybersécurité, des logiciels de gouvernance, de l'observabilité, de la mise à l'essai, de la certification, de l'assurance et de l'infrastructure sécurisée d'informatique en nuage.

Mesures recommandées

- Élaborer un cadre national de confiance pour les agents d'IA actifs sur les plans économique et opérationnel ;
- Établir des normes interopérables pour l'identité et les mécanismes d'identification des agents ;
- Définir des protocoles d'autorisation progressive et de délégation ;
- Exiger des preuves proportionnées des actions aux conséquences importantes et des chaînes de délégation ;
- Exiger l'expiration, la suspension et la révocation efficaces des mécanismes d'identification ;

- Préciser que les organisations qui déploient ces systèmes demeurent responsables du pouvoir accordé aux agents ;
- Élaborer des principes de responsabilité et de recours pour les actions non autorisées ou erronées des agents ;
- Recourir à des bacs à sable réglementaires pour mettre à l'essai le commerce agentique, la délégation en entreprise et les transactions entre machines ;
- Promouvoir l'assurance indépendante pour les agents dotés d'un pouvoir économique ou opérationnel important ;
- Faire de l'approvisionnement fédéral un premier terrain d'essai pour l'identité des agents, les contrôles d'autorisation, les preuves d'activité et la révocation ;
- Soutenir les PME au moyen d'une infrastructure partagée d'identité, d'assurance et de conformité.

Mise en œuvre

Recourir à une combinaison coordonnée d'instruments de politique publique

Aucun instrument ne peut à lui seul répondre à l'éventail des besoins en matière de transparence relevés dans la consultation. Le Canada devrait recourir à la législation lorsque des droits exécutoires, des préjudices graves ou la déclaration obligatoire sont en jeu. Les normes et les lignes directrices peuvent définir des pratiques techniques et opérationnelles. L'approvisionnement peut créer une demande précoce de produits responsables. Les bacs à sable réglementaires peuvent produire des données probantes lorsque la technologie et les normes manquent encore de maturité. Les programmes de littératie et de services-conseils peuvent aider les utilisateurs et utilisatrices ainsi que les organisations à donner suite à l'information reçue.

Instrument	Meilleure utilisation
Législation et réglementation	Droits de base, utilisations aux conséquences importantes, déclaration des incidents graves, responsabilisation et recours
Normes et lignes directrices	Formats de documentation, identité, mécanismes d'identification, journalisation, délégation, assurance et interopérabilité
Bacs à sable réglementaires	Mise à l'essai des nouvelles exigences pour les agents et d'autres applications en évolution rapide
Approvisionnement public	Création d'une demande pour la documentation, la supervision humaine, la préparation aux incidents et les contrôles des agents
Recherche et évaluation	Mise à l'essai de la compréhension, de l'efficacité, des effets involontaires et de la faisabilité technique
Littératie et assistance technique	Aide aux personnes et aux organisations pour interpréter les divulgations et adopter des pratiques responsables

Soutenir les petites et moyennes entreprises

Les PME ne devraient pas être exemptées de l'obligation de protéger les personnes contre les préjudices graves. Cependant, les voies de conformité devraient refléter leur capacité et leur rôle dans la chaîne de valeur. On ne devrait pas s'attendre à ce qu'une petite entreprise qui déploie un système tiers produise des renseignements techniques dont seul le développeur dispose. À l'inverse, les fournisseurs ne devraient pas

transférer la responsabilité de l'information au niveau du modèle à une clientèle qui ne peut raisonnablement l'obtenir.

Le gouvernement devrait soutenir les PME au moyen des éléments suivants :

- Avis types, modèles de documentation et clauses contractuelles types ;
- Guides de mise en place sectoriels et outils de diagnostic de l'adoption ;
- Services-conseils de confiance et assistance technique ;
- Mise en place échelonnée et tenue de dossiers proportionnée ;
- Infrastructure partagée de mise à l'essai, d'assurance, d'identité et de déclaration des incidents ;
- Bacs à sable réglementaires et programmes pilotes supervisés ;
- Formation en approvisionnement, en gouvernance, en cybersécurité et en intervention en cas d'incident.

Assurer la coordination entre les administrations

Les mesures fédérales devraient reposer autant que possible sur des définitions et des modèles communs et être élaborées avec les provinces, les territoires, les organismes de réglementation sectoriels, les organismes de normalisation, les partenaires autochtones, l'industrie, le milieu syndical, les chercheurs et chercheuses et la société civile. L'harmonisation internationale est particulièrement importante pour la documentation, les taxonomies d'incidents, les mécanismes d'identification des agents, l'assurance et la délégation transfrontalière. Le Canada devrait rechercher l'interopérabilité avec des partenaires de confiance tout en conservant la capacité de protéger les droits, les données, les infrastructures et les intérêts économiques du Canada.

Mesurer l'efficacité de la transparence

La conformité ne devrait pas se mesurer uniquement à l'existence d'un avis ou d'un document. Le gouvernement devrait évaluer si les destinataires visés comprennent l'information, si celle-ci améliore les décisions d'approvisionnement ou d'utilisation, si les personnes concernées peuvent obtenir un examen et un recours, si la déclaration des incidents donne lieu à des mesures préventives et si les contrôles des agents préviennent les activités non autorisées. Les exigences devraient être révisées lorsque les données probantes montrent qu'elles engendrent de la lassitude, de la confusion, des coûts inutiles ou une conformité inefficace.

Conclusion

Une véritable transparence de l'IA peut renforcer la confiance du public, protéger les Canadiens et Canadiennes, encourager une utilisation responsable de l'IA et donner aux entreprises une plus grande certitude pour adopter l'IA de façon responsable. Le CTIC recommande un cadre fondé sur les risques et les rôles et adapté au contexte, qui rend l'information compréhensible et exploitable pour ses destinataires visés, assorti de responsabilités claires tout au long de la chaîne de valeur de l'IA. Ce cadre devrait établir un véritable droit d'être informé des interactions directes avec l'IA et des utilisations aux conséquences importantes, normaliser les renseignements à plusieurs niveaux sur les systèmes d'IA et coordonner la déclaration des incidents graves et la déclaration protégée des incidents évités de justesse afin de soutenir la responsabilisation et l'apprentissage collectif. Pour les agents d'IA, un cadre d'Autonomie et agentivité graduelles devrait lier le pouvoir délégué à la compétence, à la fiabilité, à la sécurité et à la responsabilisation démontrées. Des identités vérifiables, des autorisations circonscrites, la traçabilité des actions et de la délégation et une révocation efficace devraient sous-tendre un cadre de confiance canadien pour les agents, une personne ou une organisation identifiable demeurant responsable.

La mise en place nécessitera une combinaison coordonnée de mesures législatives, de normes techniques, d'assurance indépendante, de bacs à sable réglementaires et d'approvisionnement public, appuyée par des investissements dans le développement de la littératie en matière d'IA et des capacités de gouvernance des Canadiens et Canadiennes. Des outils concrets, une infrastructure partagée et des voies de conformité proportionnées pour les PME aideront à traduire ces exigences en une adoption responsable qui soutient la productivité, l'innovation et la compétitivité des entreprises sans compromettre la sécurité ni la responsabilisation. La coordination entre les administrations canadiennes et l'harmonisation avec des partenaires internationaux de confiance peuvent améliorer l'interopérabilité tout en protégeant les droits et les intérêts canadiens.

La réussite du cadre canadien de transparence de l'IA devrait se mesurer à sa capacité d'améliorer la compréhension et la prise de décisions, de rendre possibles l'examen et le recours et de réduire les préjudices. Ensemble, ces mesures peuvent bâtir la confiance envers l'IA tout en soutenant la productivité, l'innovation, la souveraineté numérique et la compétitivité du Canada dans l'économie mondiale.

Publications du CTIC mentionnées

Anani, Namir et Erik Henningsmoen. *Une feuille de route pour l'économie numérique du Canada à l'horizon 2030.* Ottawa : Conseil des technologies de l'information et des communications, avril 2025. https://ictc-ctic.ca/fr/rapports/feuille-de-route-pour-leconomie-numerique-du-canada-a-lhorizon-2030.

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

Henningsmoen, Erik, 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*. (Ottawa : Conseil des technologies de l'information et des communications, juillet 2025). https://ictc-ctic.ca/fr/notre-impact/notes-dinformation/faire-progresser-la-capacite-nationale-souveraine-de-calcul-pour-lia-du-canada.

Conseil des technologies de l'information et des communications. *De l'IA pour tous à l'IA en action : la mise en œuvre de la stratégie nationale en matière d'intelligence artificielle du Canada pour une prospérité partagée.* (Ottawa : Conseil des technologies de l'information et des communications, juin 2026). https://ictc-ctic.ca/fr/rapports/de-lia-pour-tous-a-lia-en-action.

Conseil des technologies de l'information et des communications. *Sept recommandations clés pour le renouvellement de la stratégie canadienne en matière d'IA.* (Ottawa : Conseil des technologies de l'information et des communications, octobre 2025). https://ictc-ctic.ca/fr/rapports/sept-recommandations-cles-pour-le-renouvellement-de-la-strategie-canadienne-en-matiere-dia.