

GUIDE DES BONNES PRATIQUES D'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

APPLICABLE AUX OUTILS D'INTELLIGENCE ARTIFICIELLE
GÉNÉRATIVE PUBLICS

MINISTÈRE DE LA CYBERSÉCURITÉ

ET DU NUMÉRIQUE

RÉDACTION

Inscrire l'auteur (ex. : La Direction des politiques et des affaires institutionnelles de la Direction générale des politiques et de la gouvernance du Sous-ministériat adjoint à la gouvernance et au financement du ministère de la Cybersécurité et du Numérique)

Si vous éprouvez des difficultés techniques ou si vous souhaitez obtenir une version adaptée du document, veuillez communiquer avec la Direction des communications :

Direction des communications
Ministère de la Cybersécurité et du Numérique
900, place D'Youville, 2^e étage
Québec (Québec) G1R 3P7
Courriel : information@mcn.gouv.qc.ca

Tous droits réservés pour tous pays. La reproduction, par quelque procédé que ce soit, la traduction ou la diffusion de ce document, même partielles, sont interdites sans l'autorisation des Publications du Québec. Cependant, la reproduction de ce document ou son utilisation à des fins personnelles, d'étude privée ou de recherche scientifique, mais non commerciales, est permise à condition d'en mentionner la source.

TABLE DES MATIÈRES

Historique des changements	i
Abréviations.....	ii
Public cible	1
Objectifs	1
Cadre légal et réglementaire	2
Stratégie d'intégration de l'ia dans l'administration publique	3
Contexte	3
Qu'est-ce que l'IA générative ?	4
Puis-je saisir des renseignements confidentiels dans un outil d'ia générative public ?	5
Cas d'utilisation.....	6
Risques	8
Catégorie 1 : Risques liés à l'utilisation d'outils d'IA générative publics par le personnel de l'administration publique	9
Catégorie 2 : Risques liés à l'utilisation d'outils d'IA générative publics par les cybercriminels.	11
Catégorie 3 : Risques découlant des vulnérabilités des modèles d'IA générative publics exploitées par les cybercriminels	12
Recommandations à l'échelle organisationnelle.....	13
Bonnes pratiques individuelles	14
Qui contacter?	16
Annexe – Résumé des bonnes pratiques individuelles	17

HISTORIQUE DES CHANGEMENTS

Version	Date de publication	Modifications
• 1.0	• Octobre 2024	• Publication de la première édition
• 1.1	• Mars 2026	• Ajustements pour l'arrimage avec l'Indication d'application contenant les mesures applicables lors de l'utilisation de l'intelligence artificielle générative (IAG) par les organismes publics et la mise à jour de l'Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics : - Mise à jour du titre - Mise à jour de la section <i>Abréviations</i> (anciennement <i>Liste d'acronymes</i>) - Mise à jour de la section <i>Public cible</i> - Mise à jour de la section <i>Objectifs</i> - Mise à jour de la section <i>Cadre légal et réglementaire</i> - Mise à jour de la section <i>Stratégie d'intégration de l'intelligence artificielle dans l'administration publique</i> - Mise à jour de la section <i>Contexte</i> - Mise à jour de la section <i>Qu'est-ce que l'IA générative ?</i> - Mise à jour de la section <i>Cas d'utilisation</i> - Mise à jour de la section <i>Risques</i> - Mise à jour de la section <i>Recommandations à l'échelle organisationnelle</i> - Mise à jour de la section <i>Les bonnes pratiques individuelles</i> - Mise à jour de la section <i>Qui contacter ?</i> - Mise à jour des références dans le texte et les notes de bas de page (la section <i>Références bibliographiques</i> est enlevée) - Mise à jour de l'annexe <i>Résumé des bonnes pratiques individuelles</i> - Ajout de la section <i>Puis-je saisir des renseignements confidentiels dans un outil d'IA générative public ?</i>

La version en vigueur est disponible à l'adresse suivante :

<https://www.quebec.ca/gouvernement/numerique/intelligence-artificielle-administration-publique/utilisation-responsable-intelligence-artificielle>

ABRÉVIATIONS

GPT	Transformeur génératif préentraîné (<i>generative pre-trained transformer</i>)
IA	Intelligence artificielle
IAG	Intelligence artificielle générative
MCN	Ministère de la Cybersécurité et du Numérique
OP	Organismes publics
OQLF	Office québécois de la langue française
PI	Propriété intellectuelle
RLRQ	Recueil des lois et des règlements du Québec

PUBLIC CIBLE

Ce guide est d'abord destiné aux instances de gouvernance des organismes publics (OP) de l'administration publique québécoise dans le cadre de la gestion et du soutien du personnel dans l'utilisation d'outils d'intelligence artificielle (IA) générative publics.

Le guide s'adresse également au personnel de l'administration publique québécoise susceptible d'utiliser, dans le cadre de ses activités professionnelles, des solutions d'IA générative publiques qui ne sont pas déployées officiellement par les OP.

Ce guide porte sur les outils d'IA générative publics qui sont librement accessibles à tous (qu'il s'agisse de version gratuite ou payante) et offerts par un service externe, et ce, sans avoir été développés ou acquis par une organisation.

Pour les règles et pratiques à adopter pour les outils officiellement acquis ou développés par un OP, veuillez vous référer à l'[Indication d'application IA-RI-2025-003-OP \(Mesures applicables lors de l'utilisation de l'intelligence artificielle générative\)](#) et au [Guide d'application des mesures applicables lors de l'utilisation de l'intelligence artificielle générative](#).

OBJECTIFS

Les objectifs de ce guide sont de :

- guider les OP vers une utilisation responsable et adéquate de cette technologie;
- sensibiliser le personnel de l'administration publique aux possibilités et aux risques de l'utilisation d'outils d'IA générative publics.

Les moyens pour y parvenir sont de :

- préconiser la prudence lors de l'utilisation de l'IA générative en général;
- proposer des consignes mettant l'accent sur l'acquisition des réflexes prudents.

Ce guide est un document qui sera mis à jour de manière itérative. Le guide propose aux OP et à leur personnel des éléments à prendre en considération dans l'utilisation et dans l'encadrement de l'utilisation des outils d'IA générative, plutôt que d'imposer des interdictions ou des limitations quant à l'utilisation des outils d'IA générative publics. Étant donné que ces outils n'ont pas fait l'objet d'un processus de validation comme celui qui est exigé pour les outils officiellement déployés et approuvés par un organisme, des précautions spécifiques sont à appliquer.

Il s'agit de donner aux OP et au personnel les moyens de faire des choix responsables et de développer de bonnes habitudes applicables dans différents contextes, ce qui est essentiel pour utiliser correctement cette technologie en constante évolution. Au même titre que l'utilisation de tout autre outil numérique, comme les réseaux sociaux, l'utilisation des outils d'IA générative doit être responsable.

CADRE LÉGAL ET RÉGLEMENTAIRE

Les bonnes pratiques de ce guide doivent toujours être prises en compte parallèlement aux lois, règlements, directives, politiques et orientations concernant l'utilisation de la technologie et des services numériques, entre autres, et sans être exhaustif :

[Loi sur la gouvernance et la gestion des ressources informationnelles des organismes publics et des entreprises du gouvernement \(RLRQ, c. G-1.03\);](#)

[Loi concernant le cadre juridique des technologies de l'information \(RLRQ, c. C-1.1\);](#)

[Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels \(RLRQ, c. A-2.1\);](#)

[Politique gouvernementale de cybersécurité;](#)

[Directive gouvernementale sur la sécurité de l'information \(décret numéro 1514-2021 du 8 décembre 2021\);](#)

[Exigences en matière de ressources informationnelles au regard de l'utilisation de l'intelligence artificielle par les organismes publics \(arrêté numéro 2024-01 du ministre de la Cybersécurité et du Numérique en date du 28 février 2024\);](#)

[Modèle de classification de sécurité des données numériques gouvernementales \(Arrêté numéro 2024-05 du ministre de la Cybersécurité et du Numérique en date du 12 décembre 2024\);](#)

[Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics \(arrêté numéro 2025-02 du ministre de la Cybersécurité et du Numérique en date du 3 décembre 2025\);](#)

[Indication d'application contenant les mesures applicables lors de l'utilisation de l'intelligence artificielle générative \(IAG\) par les organismes publics \(indication IA-RI-2025-003-OP entré en vigueur le 5 décembre 2025\).](#)

STRATÉGIE D'INTÉGRATION DE L'IA DANS L'ADMINISTRATION PUBLIQUE

Afin de profiter pleinement du potentiel de l'IA comme outil d'optimisation et de modernisation de ses processus, le gouvernement du Québec a lancé officiellement en juin 2021 la [Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026](#) (ci-après « la Stratégie »).

La Stratégie vise à positionner l'administration publique comme acteur exemplaire de l'IA en prenant notamment appui sur le leadership du Québec dans ce domaine. Le présent guide s'appuie particulièrement sur les trois axes de la Stratégie :

- Des services publics renouvelés et optimisés par l'IA;
- Une administration publique outillée et proactive à l'égard des changements engendrés par l'IA;
- Une action gouvernementale en IA fondée sur des pratiques responsables.

De plus, ce guide concourt à la réalisation de deux objectifs de cette Stratégie :

- **L'objectif 2** : Soutenir les organismes publics afin qu'ils tirent profit de l'IA dans le déploiement des services publics.
- **L'objectif 8** : Encadrer la conception et l'utilisation de l'IA par des balises éthiques et des pratiques de sécurité adaptées à l'IA.

Enfin, le guide appuie les autres mesures prévues à la Stratégie et est complémentaire à ces dernières. Parmi celles-ci, notons celles de baliser l'utilisation de l'IA générative dans l'administration publique. Des références vers les documents qui y sont liés sont indiquées au début de ce guide.

CONTEXTE

À la fin de l'année 2022 et au cours de l'année subséquente, la possibilité d'utiliser des outils d'IA générative dans un contexte professionnel au sein de l'administration publique québécoise s'est étendue avec une offre d'outils plus performants. L'IA générative est donc apparue comme un moteur de changement majeur. À une époque où les services publics visent une plus grande efficacité et une adaptation aux besoins individuels, l'IA générative peut redéfinir la manière dont le secteur public est capable d'accomplir ses missions et de communiquer avec les citoyens. Dans ce contexte, l'usage d'outils d'IA générative publics peut être pertinent lorsque permis par les OP. Tout en gardant à l'esprit

les aspects positifs qui peuvent être générés par l'utilisation de ces outils, les utilisateurs doivent également prendre garde aux risques que cela peut soulever.

QU'EST-CE QUE L'IA GÉNÉRATIVE?

L'IA générative (IAG) peut être définie comme un « *système d'intelligence artificielle utilisé pour produire du contenu au moyen d'algorithmes et de mégadonnées, généralement sous forme de texte, de son, de vidéo ou d'image* »¹. Ces algorithmes sont issus de diverses branches, méthodes et techniques de l'IA telles que l'apprentissage automatique, l'apprentissage profond, le traitement automatique du langage naturel ou la vision par ordinateur pour produire du contenu sous diverses formes. L'IA générative se distingue de l'IA prédictive², laquelle est utilisée pour prédire des résultats, des événements, des comportements ou des tendances à partir des données.

Cette capacité « générative » existe depuis des dizaines d'années dans les algorithmes d'IA, notamment les modèles graphiques probabilistes (réseaux bayésiens, modèles de Markov cachés, champs aléatoires de Markov) et les réseaux de neurones artificiels (réseaux antagonistes génératifs).

L'IA générative a récemment gagné en visibilité grâce à la convergence de plusieurs avancées majeures, comprenant l'augmentation spectaculaire de la puissance de calcul, l'accès à de vastes ensembles de données et les progrès en apprentissage automatique, en particulier dans les domaines de l'apprentissage profond et de l'apprentissage par renforcement. Les approches génératives contemporaines les plus marquantes incluent notamment les modèles transformeurs³, notamment les transformeurs génératifs préentraînés ou GPT (*generative pre-trained transformer*), initialement conçus pour le traitement automatique des langues, notamment les grands modèles de langage et les outils de traductions), mais par la suite adaptés pour différents types de données (code informatique, image, vidéo, audio, molécule, etc.).

Actuellement, ces systèmes et applications comptent parmi les plus populaires et suscitent à la fois de l'engouement, mais aussi des débats sur son utilisation. Parmi les exemples bien connus du grand public, on peut citer les assistants virtuels ou les agents conversationnels intelligents appuyés par l'IAG (exemples : ChatGPT ou Microsoft Copilot), les assistants de programmation (exemples : GitHub Copilot) ou les hypertrucages.

¹ [Intelligence artificielle générative \(OQLF, consulté le 17 mars 2026\)](#).

² [Intelligence artificielle prédictive \(OQLF, consulté le 17 mars 2026\)](#).

³ [Modèle transformeur \(OQLF, consulté le 17 mars 2026\)](#).

PUIS-JE SAISIR DES RENSEIGNEMENTS CONFIDENTIELS DANS UN OUTIL D'IA GÉNÉRATIVE PUBLIC?

Il est interdit d'entrer des **renseignements confidentiels**, y compris des renseignements personnels, dans les outils d'IA générative publics. Ces plateformes ne sont pas exploitées dans les environnements sécurisés du gouvernement et leurs modalités d'utilisation peuvent permettre l'analyse ou la conservation des données, ce qui expose l'organisation à des risques juridiques, de sécurité de l'information ou relatifs à la réputation.

Les renseignements confidentiels incluent, sans s'y limiter, les renseignements personnels, lesquels sont définis comme tout renseignement concernant une personne physique et permettant de l'identifier directement ou indirectement, comme une adresse, un numéro de téléphone, une photo, un dossier médical, un revenu, etc. Certains renseignements personnels peuvent être sensibles et nécessitent un niveau de protection renforcé (exemples : données de santé et données biométriques).

Ces renseignements sont en principe confidentiels, sauf exception prévus par la *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels* (RLRQ, c. A-2.1) (p. ex. certains renseignements à caractère public, comme le nom et la fonction d'un haut dirigeant d'un organisme public).

Un renseignement confidentiel est défini comme tout renseignement dont l'accès est limité par la *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, notamment ceux dont la divulgation est susceptible de nuire à divers intérêts publics :

- les relations intergouvernementales;
- les négociations entre organismes publics, l'économie;
- l'administration de la justice et la sécurité publique;
- les décisions administratives ou politiques;
- la vérification.

Les renseignements confidentiels incluent, à titre indicatif : documents décisionnels et analyses internes non publiées, informations financières ou stratégiques non publiques, contenus relevant de la sécurité de l'État, de la cybersécurité ou des infrastructures critiques, ainsi que toute information couverte par une entente de confidentialité ou une classification interne des données numériques gouvernementales (exemples : Protégé A, B et C, Diffusion restreinte, Confidentiel, Secret, Très Secret, etc.).

Pour utiliser les renseignements confidentiels pour les cas d'utilisation approuvés par votre organisation, n'employer que des outils approuvés et hébergés dans les environnements sécuritaires conformes aux politiques gouvernementales de sécurité et de gestion de l'information. À titre d'exemple, voir la matrice de souveraineté pour l'hébergement et le traitement des données numériques gouvernementales dans le [Guide d'application des mesures applicables lors de l'utilisation de l'intelligence artificielle générative](#).

Définitions utiles :

Renseignement confidentiel	• Renseignement dont la divulgation serait susceptible d'avoir des incidences néfastes, notamment sur les relations intergouvernementales, les négociations entre organismes publics, l'économie, les tiers, l'administration de la justice et la sécurité publique, les décisions administratives ou politiques et la vérification. Sont également confidentiels les renseignements couverts par le secret professionnel et les renseignements personnels, sauf dans les cas prévus par la Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels.
Renseignement personnel	• Un renseignement personnel est un renseignement qui concerne une personne physique et qui permet, directement ou indirectement, de l'identifier. Il permet de reconnaître ou d'apprendre quelque chose au sujet d'une personne physique identifiable, se rapportant, par exemple, à la nature de l'individu, ses caractéristiques propres, sa situation personnelle, etc.
Renseignement personnel sensible	• Un renseignement personnel est considéré comme sensible lorsque, par sa nature notamment médicale, biométrique ou autrement intime ou en raison du contexte de son utilisation ou de sa communication, il suscite un haut degré d'attente raisonnable en matière de respect de la vie privée. Cette sensibilité doit donc être évaluée en fonction du degré d'attente raisonnable du renseignement personnel en matière de vie privée. Lorsque le degré est élevé, cette information doit être considérée comme un renseignement personnel sensible. • En d'autres termes, il s'agit de tout renseignement qui peut affecter davantage la personne concernée, compte tenu de sa source, du degré d'intimité auquel il est rattaché, et des répercussions potentiellement liées à sa divulgation ou à son utilisation. • Il peut s'agir, par exemple, de renseignements médicaux, biométriques, génétiques ou financiers, ou encore de renseignements sur la vie ou l'orientation sexuelle, les opinions politiques, les convictions religieuses ou bien l'origine ethnique.

CAS D'UTILISATION

L'IA générative offre des applications variées dans la prestation et la gestion des services publics. Sans remplacer la réflexion et le jugement humains, l'IA générative peut apporter du soutien dans divers domaines d'activités.

Voici quelques exemples de cas d'utilisation de l'IA générative :

Idéation/remue-méninge	L'IA générative peut servir de catalyseur pour stimuler la créativité et générer de nouvelles idées.	Génération de scénarios d'intervention : esquisser plusieurs scénarios de programme d'intervention avec avantages et inconvénients.
Création de contenu	L'IA générative peut être utilisée pour créer de nouveaux contenus.	Rédaction de brouillons pour des documents : aider à la rédaction de brouillons de présentations, de rapports et de notes. Rédaction de courriels : aider à la rédaction de correspondance courante tout en maintenant une communication claire et professionnelle. Création d'actualités, contenus pour la publicité et le marketing : générer des contenus visant à informer le public des initiatives gouvernementales, des changements de politique et des services disponibles, en personnalisant le contenu pour différents groupes démographiques et régions par exemple
Création de persona	L'IA générative peut aider à développer des personas pour les campagnes de communication publique et proposer des messages ciblés pour les différents segments de la population.	Messages adaptés : créer des personas afin de proposer des variantes de messages selon la littératie numérique et barrières d'accès.
Création de contenus multimédias	L'IA générative peut créer de nouveaux sons ou de nouvelles images, voix, vidéos ou musique.	Accessibilité : créer des descriptions audios pour images et documents.
Programmation de code informatique	L'IA générative peut appuyer les travaux de développement de systèmes informatiques, notamment pour le débogage, l'explication de code et la conversion à un autre langage de programmation.	Amélioration : suggérer des améliorations de lisibilité et de sécurité dans le code informatique.
Production de listes de contrôle	L'IA générative peut générer des listes de contrôle à partir de courriels ou de comptes-rendus de réunions pour assurer le suivi des actions et des décisions.	Synthèse des décisions d'une réunion : repérer les décisions, les actions à entreprendre et les dépendances entre équipes à partir de notes de réunion et générer une liste de contrôle concise pour le suivi opérationnel et la traçabilité des engagements.
Formation	L'IA générative peut être utilisée de différentes manières dans le contexte de formation du personnel.	Parcours de formation : proposer des parcours personnalisés pour le personnel en ciblant les compétences à bonifier et en recommandant des ressources éducatives pertinentes. Génération de contenu de formation : contribuer à créer du contenu de formation, tel que des manuels, des guides ou des supports de cours Conception de tests de connaissances : aider à créer des évaluations et des tests pour mesurer les compétences et les connaissances
Analyse de données	L'IA générative peut traiter et analyser des données, notamment pour aider à la prise de décision, en cernant les tendances, les modèles et les prévisions pour la planification et l'allocation des ressources.	Analyse automatisée de données opérationnelles : analyse rapidement des données opérationnelles et en dégager les tendances clés, les écarts et les facteurs de risques, afin d'aider les gestionnaires à réajuster leur planification et l'allocation des ressources.
Traitement de contenu	L'IA générative peut être utilisée pour faciliter le traitement de contenu déjà existant.	Vulgarisation de concepts complexes : transformer des informations techniques ou des données complexes en explications simples et en visuels clairs pour faciliter la compréhension. Correction de texte : effectuer la relecture et la correction de documents officiels, en s'assurant que le texte est exempt de fautes d'orthographe, de grammaire et de ponctuation.



- **Normalisation de texte** : aider à uniformiser le ton et le style des communications et s'assurer qu'elles répondent aux normes de l'administration publique et qu'elles sont cohérentes sur tous les canaux lors de communications officielles
- **Conversion de contenu** : faciliter la réutilisation de contenu d'une plateforme à une autre, par exemple en adaptant des documents textuels pour une présentation en ligne ou en transformant des données en infographie
- **Traduction de contenu** : traduire rapidement du contenu
- **Synthèse de textes** : résumer de longs documents en points clés

Dans tous les cas d'utilisation, il est proscrit de :

- saisir des renseignements confidentiels;
- demander d'utiliser ou de s'inspirer de matériel protégé par des droits d'auteur;
- demander à l'IA d'enfreindre ou d'aider à enfreindre une loi, une directive ou un règlement.

L'organisation et le membre du personnel demeurent imputables de l'utilisation des résultats générés par les outils d'IA.

RISQUES

Les outils qui utilisent l'IA ont le potentiel de libérer du temps pour se concentrer sur des tâches à plus forte valeur ajoutée. Cependant, il est important que le personnel de l'administration publique soit pleinement sensibilisé aux risques et adopte de bonnes pratiques d'utilisation responsable de ces technologies. Bien que cette section ne puisse remplacer une analyse de risques en bonne et due forme et qu'il puisse exister d'autres risques, voici trois grandes catégories de risques à garder en tête lors de l'utilisation de ces outils :

- **Catégorie 1** : risques liés à l'utilisation d'outils d'IA générative publics par le personnel de l'administration publique.
- **Catégorie 2** : risques liés à l'utilisation d'outils d'IA générative publics par les cybercriminels.
- **Catégorie 3** : risques découlant des vulnérabilités des modèles d'IA générative publics exploitées par les cybercriminels.

La catégorie 1 regroupe les risques découlant de l'utilisation d'outils d'IA générative publics par le personnel de l'administration publique, notamment la confidentialité des

informations utilisées et les résultats potentiellement biaisés, tandis que les catégories 2 et 3 concernent les risques associés à la cybercriminalité, soit par l'usage malveillant des outils d'IA générative par des acteurs hostiles (catégorie 2) ou par l'exploitation des vulnérabilités propres aux modèles d'IA générative publics (catégorie 3).

Catégorie 1 : Risques liés à l'utilisation d'outils d'IA générative publics par le personnel de l'administration publique

Confidentialité des renseignements et protection des renseignements personnels

Les systèmes d'IA générative sont souvent munis de composantes de collecte des données servant à des fins d'analyse et même de réutilisation pour améliorer les performances des modèles d'IA générative, notamment à l'aide de l'apprentissage automatisé. Par exemple, une personne peut, involontairement, fournir de l'information avec des renseignements confidentiels ou des renseignements personnels pour qu'elles soient traitées par l'outil comme données d'entrée. Les données saisies pourraient être conservées dans le système, pour des fins d'entraînement, et ensuite divulguées en réponse à la demande d'un autre utilisateur qui ne devrait pas normalement y avoir accès.

Biais

Les modèles de langage sont entraînés sur des données existantes pouvant contenir des biais implicites ou explicites, ce qui peut amener une production de texte ou d'autres actions biaisées. Ces biais peuvent provenir de différentes sources, telles que les préjugés sociétaux, les données déséquilibrées ou les erreurs humaines lors de l'étiquetage des données. Le biais dans l'IA se réfère à des tendances ou à des préjugés non intentionnels qui peuvent être présents dans les résultats produits par l'IA et créer un préjudice à un citoyen ou au personnel de l'administration publique, par exemple.

Fiabilité

L'IA générative peut parfois produire des erreurs ou des « hallucinations », c'est-à-dire inventer des faits dans sa réponse étant donné qu'elle est basée sur des modèles statistiques et ne comprend pas le sens du texte. En effet, l'IA générative peut créer de fausses informations qui semblent plausibles à première vue, d'où l'importance de toujours vérifier les faits, même pour des questions simples. Des décisions, prises à partir d'informations erronées, peuvent entraîner des conséquences indésirables pour l'organisme ou créer un préjudice à un citoyen ou au personnel de l'administration publique, par exemple.

Dépendance technologique

L'adoption de l'IA dans l'administration publique, au sein des entreprises ainsi que dans notre quotidien peut entraîner une dépendance excessive à ces systèmes. Au sein du monde professionnel, les personnes ayant recours à ces outils peuvent devenir moins enclines à effectuer des tâches analytiques ou décisionnelles sans l'assistance de l'IA, ce qui peut diminuer leur capacité à fonctionner indépendamment de la technologie.

Respect de la législation

L'IA générative propose des informations et des contenus qui pourraient soutenir une prise de décisions ayant un impact réel sur les clientèles des OP (individus, communautés ou entreprises). Si les orientations, les actions ou les décisions adoptées ne respectent pas les lois et la réglementation en vigueur, elles pourraient générer des préjudices et même possiblement des sanctions. Aussi, l'IA générative n'est pas nécessairement alimentée par les textes de loi en vigueur au Québec et au Canada et elle peut proposer des actions ou du contenu qui contreviendraient à des lois ou règlements en vigueur. Enfin, nourrir l'outil avec des renseignements personnels pourrait être considéré comme une infraction à la législation entourant la protection des renseignements personnels.

Respect de la propriété intellectuelle

La propriété intellectuelle (PI) est : « *le résultat d'un travail de création de l'esprit qui fait l'objet d'un droit* »⁴. On reconnaît notamment : le droit d'auteur, le brevet, le modèle industriel, la marque de commerce, l'appellation d'origine et le secret industriel.

Le risque lié à la PI concerne les incertitudes et les complications juridiques quant à la propriété et aux droits sur les contenus créés par des outils d'IA générative. En effet, ces systèmes peuvent imiter, reproduire ou s'inspirer de travaux protégés par différents types de droits de PI lors de la génération de contenu, soulevant des questions de violation potentielle de ces droits.

Génération de code non sécurisé

La génération de code non sécurisé désigne le risque que les outils basés sur l'IA produisent du code informatique qui ne respecte pas le cadre de développement sécuritaire et qui peut comporter des vulnérabilités ou du code malveillant. Ces vulnérabilités peuvent permettre l'accès non autorisé à des données, des intrusions dans des systèmes, des fuites d'informations sensibles, etc.

⁴ [Propriété intellectuelle \(OQLF, consulté le 17 mars 2026\)](#).

Catégorie 2 : Risques liés à l'utilisation d'outils d'IA générative publics par les cybercriminels

L'IA est devenue un moteur d'innovation incontournable pour les organisations, y compris les gouvernements. Cependant, le développement des modèles d'IA présente des défis uniques en matière de sécurité.

L'IA peut être utilisée pour amplifier les campagnes de mésinformation et de désinformation

Les modèles de langage peuvent générer automatiquement du contenu trompeur, tel que de fausses nouvelles, des théories du complot ou des hypertrucages, qui peuvent influencer l'opinion publique et saper la confiance dans les institutions. La diffusion massive de ces contenus générés par l'IA peut rendre plus difficile la distinction entre les informations véridiques et les manipulations.

L'IA peut également être exploitée pour des attaques d'hameçonnage plus sophistiquées

Les modèles de langage peuvent être utilisés pour créer des courriels personnalisés et convaincants, ciblant des individus spécifiques avec des informations obtenues à partir de leurs données personnelles. Ces attaques de piratage psychologique améliorées par l'IA augmentent les risques de compromission des identifiants et des informations sensibles.

L'IA peut jouer un rôle dans la propagation de rançongiciels

Les attaquants peuvent utiliser des modèles d'IA pour identifier les cibles potentielles, optimiser leurs méthodes d'infection et automatiser certaines parties de leurs attaques. De plus, l'IA peut être utilisée pour générer des variantes de rançongiciels plus difficiles à détecter par les solutions de sécurité traditionnelles, augmentant ainsi la probabilité de succès des attaques.

L'IA générative renforce les menaces de piratage psychologique

Les attaques de piratage psychologique⁵ (ou ingénierie sociale) se produisent lorsque des auteurs de menace se servent d'un lien social et de manipulation pour pousser ou inciter des utilisateurs à divulguer quelque chose qui va à l'encontre des intérêts d'une organisation, comme le fait de fournir des détails sensibles, des mots de passe ou de l'information financière. L'IA générative amplifie ces risques en permettant la création

⁵[À propos du piratage psychologique \(Québec.ca, consulté le 17 mars 2026\).](#)

d'hypertrucages audio et vidéo très réalistes, ainsi que de faux profils sur les réseaux sociaux, rendant ces tentatives de manipulation plus crédibles et difficiles à détecter.

Catégorie 3 : Risques découlant des vulnérabilités des modèles d'IA générative publics exploitées par les cybercriminels

Les modèles d'IA générative, bien que puissants et innovants, ne sont pas à l'abri des cyberattaques. Les cybercriminels peuvent exploiter activement les failles de sécurité de ces systèmes, compromettant ainsi leur intégrité, leur confidentialité et leur disponibilité. Parmi les principales menaces pesant sur les modèles d'IA générative, trois types d'attaques se distinguent particulièrement : les attaques par empoisonnement, l'extraction de données et les violations par abus.

Les attaques par empoisonnement

Les attaques par empoisonnement consistent à introduire des données trompeuses ou malveillantes dans les ensembles de données utilisés pour entraîner les systèmes d'IA. Cette manipulation peut amener l'IA à produire des résultats erronés, biaisés ou même dangereux, compromettant ainsi son intégrité et sa fiabilité.

L'extraction de données

Des individus malintentionnés peuvent essayer d'accéder à des informations avec renseignements confidentiels stockées dans les systèmes d'IA. Des données personnelles, des informations classifiées ou même des secrets d'État pourraient tomber dans les mains de cybercriminels.

Les violations par abus

Des personnes mal intentionnées détournent les capacités des systèmes d'IA pour générer du contenu trompeur, offensant, biaisé ou préjudiciable. Notamment, cela peut inclure la création de fausses nouvelles, de propagande ou de contenus discriminatoires, portant atteinte à la confiance du public et à la réputation des institutions.

Perte de confiance envers les institutions publiques et le gouvernement du Québec

Si certains des risques énumérés précédemment et associés à cette technologie devaient se concrétiser, cela pourrait affecter négativement la perception publique du gouvernement. En particulier, cela risque de diminuer la confiance des citoyens envers les actions et les décisions gouvernementales, notamment si ces risques touchent la sécurité, la confidentialité des données ou l'éthique.

RECOMMANDATIONS À L'ÉCHELLE ORGANISATIONNELLE

Pour une utilisation responsable et adéquate de l'intelligence artificielle, les OP de l'administration publique québécoise doivent suivre l'[Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics \(arrêté numéro 2025-02 du ministre de la Cybersécurité et du Numérique en date du 3 décembre 2025\)](#).

Plus précisément pour les outils d'IA générative publics, les OP devraient adopter une posture de prudence et préconiser l'application des principes spécifiques suivants :

- **Protection** : veiller à la sécurité des informations avec renseignements confidentiels, notamment des renseignements stratégiques, et à la protection des renseignements personnels. L'organisation demeure imputable de la protection des informations qu'elle détient.
- **Responsabilité** : assumer la responsabilité du contenu généré par l'IA, en veillant à ce qu'il soit factuel, légal et éthique. Il faut s'assurer que ces résultats sont cohérents, valides et pertinents dans le cadre du mandat concerné. L'organisation demeure imputable des résultats qu'elle produit et de l'information qu'elle diffuse.
- **Utilité** : s'assurer que les outils d'IA générative sont utilisés de manière efficace et pertinente pour répondre aux besoins de l'organisation ou des citoyens.
- **Diligence** : être diligent et proactive dans la gestion des risques et des incidents liés à l'IA.
- **Neutralité** : veiller à ce que les résultats générés par les outils d'IA générative soient validés et reformulés au besoin afin d'être exempts de biais et qu'ils soient utilisés de manière équitable.
- **Traçabilité** : évaluer la pertinence de suivre et de documenter l'utilisation de l'IA générative, offrant ainsi une forme de transparence dans le suivi et la justification des actions et décisions.
- **Sensibilisation et gestion du changement** : informer et sensibiliser sur le fonctionnement, les forces, les limites et l'utilisation responsable des outils d'IA générative.
- **Santé et bien-être** : s'assurer que le personnel a accès à des ressources en cas d'impact négatif par rapport à leur santé ou à leur bien-être, selon les ressources disponibles au sein de l'organisation.

Le contexte de chaque OP est différent. Les principes énumérés ci-haut peuvent être bonifiés pour s'adapter à leur contexte particulier. Les organisations ont généralement différents mécanismes entourant la sécurité et l'accès à l'information et la protection des renseignements personnels et ceux-ci pourraient être interpellés en lien avec l'utilisation de tels outils ou d'incidents liés à leur utilisation. Il pourrait être pertinent de désigner une personne responsable pour répondre aux questions ou gérer les incidents liés à ces outils.

BONNES PRATIQUES INDIVIDUELLES

Ces pratiques prudentes sont des mesures préventives et des comportements à intégrer dans la routine quotidienne afin de mitiger les risques associés à l'utilisation d'outils d'IA générative publics. Voici des réflexes prudents à adopter.

1. Ne jamais divulguer d'informations avec des renseignements confidentiels

La protection des données est primordiale. Il ne faut jamais soumettre des informations contenant des renseignements confidentiels ou qui font l'objet de propriété intellectuelle de l'organisme public à des systèmes d'IA générative publics. Le personnel demeure imputable de l'utilisation qu'il fait des données de son organisation. Pour des informations supplémentaires, les unités responsables de la protection des renseignements personnels et de la sécurité informatique de votre organisation peuvent être interpellées.

2. Reconnaître que l'outil, bien que puissant, peut contenir des erreurs ou des imprécisions

Bien que l'IA générative soit un outil performant, elle n'est pas exempte d'erreurs. Les réponses générées ne doivent pas être considérées comme des informations définitives, mais plutôt comme une base à la réflexion et aux vérifications supplémentaires. Il incombe donc à la personne utilisatrice de revoir, de réviser et de vérifier les informations générées. Pour y parvenir, vous pouvez entre autres :

- vérifier la source (si elle est disponible);
- vérifier l'auteur;
- croiser et diversifier les sources pour corroborer les réponses de l'IA générative;
- utiliser seulement les outils d'IA approuvés par votre organisation;
- solliciter l'expertise humaine;

- utiliser des outils de vérification des faits⁶, notamment :
 - [Vérificateur de faits sur HabiloMédias](#);
 - [Vérification des faits AFP Factuel](#).

Le recours à un outil d'IA générative devrait prendre en considération le contexte dans lequel il est utilisé. Par exemple, les conséquences de son utilisation lorsqu'on fait un remue-méninge ne sont différentes que lors d'une prise de décision critique, ce qui mérite des précautions supplémentaires.

3. Respecter les conditions d'utilisation d'un outil en ligne

Il faut s'assurer de lire et de respecter les conditions d'utilisation. Cela peut inclure des restrictions sur le type de contenu généré, des limitations de temps d'utilisation ou d'autres règles spécifiques.

4. S'assurer de la conformité au cadre législatif et aux normes éthiques en vigueur

Le matériel généré par les outils d'IA générative publics doit respecter les lois, règles et directives existantes. Le personnel de l'administration publique demeure imputable de respecter le cadre légal, les principes pour une utilisation responsable de l'IA et les normes éthiques gouvernementales et de son organisation, s'il y a lieu.

5. S'assurer que le canal de communication réseau avec l'IA est chiffré et vérifier toujours l'URL dans la barre d'adresse

S'assurer, par exemple, d'utiliser le protocole de communication « https » entre un client Web et un serveur Web qui recourt au chiffrement afin de sécuriser les échanges et d'éviter qu'un tiers malveillant puisse y accéder.

Même si des sites Web ont des certificats valides, cela ne garantit pas toujours qu'ils soient dignes de confiance. Il faut donc toujours vérifier l'URL dans la barre d'adresse pour s'assurer que l'on se trouve sur le site prévu avant de saisir des informations.

⁶ [Comment repérer la désinformation \(Québec.ca, consulté le 17 mars 2026\)](#).

6. Signaler toute information qui pourrait laisser croire à une fuite d'informations ou à la publication d'informations trompeuses, erronées ou diffamatoires au sujet de l'organisation

Certaines applications d'IA générative publiques conservent les données qui leur sont soumises pour les réutiliser (entraîner le modèle). Si en utilisant ces applications, les réponses obtenues laissent croire à une fuite d'informations ou à la publication d'informations trompeuses, erronées ou diffamatoires au sujet de l'organisation, il faut le signaler rapidement conformément à la procédure en vigueur au sein de cette dernière.

7. Être prudent dans le choix des outils et se tenir régulièrement informé des mises à jour, des améliorations et des éventuelles limitations de l'outil

Avec la multiplication des offres d'IA sur le marché, il est essentiel de faire preuve de discernement dans le choix des outils que vous utilisez. Il existe une multitude d'IA dans tous les domaines et champs d'applications, dont la sécurité et la fiabilité peuvent varier considérablement. Il est crucial d'être vigilant face aux IA malveillantes ou développées sans les mesures de sécurité adéquates, car elles peuvent exposer votre organisation à des risques importants, tels que la fuite de données sensibles, la manipulation des résultats ou encore la compromission des systèmes.

Pour garantir une utilisation sécurisée et responsable de l'IA, il est essentiel de privilégier les outils approuvés par votre organisation. Ces outils ont été évalués et validés pour répondre aux exigences de sécurité et de conformité. En cas de doute sur la fiabilité ou la sécurité d'une IA, il est impératif de consulter les autorités compétentes au sein de votre organisation, telles que le service informatique ou le responsable de la sécurité des systèmes d'information. Ils pourront vous guider dans le choix des outils les plus adaptés à vos besoins et vous aider à mettre en place les mesures de sécurité nécessaires pour protéger vos données et vos systèmes.

Enfin, étant donné la rapidité avec laquelle évolue la technologie, il est conseillé de se tenir régulièrement informé et de consulter les directives d'utilisation les plus récentes en visitant la [Bibliothèque du Centre gouvernemental de cyberdéfense](#). Ainsi, au fil du temps, il est possible que l'usage de certains outils précis se retrouve restreint, voire interdit, en raison de divers enjeux de sécurité. À titre d'exemple, l'[Indication sur l'interdiction d'utilisation au regard des assistants virtuels DeepSeek \(IA-SI-2025-001-OP\)](#).

QUI CONTACTER?

Pour des questions concernant le guide sur les bonnes pratiques d'utilisation des outils d'IA générative offerts par des services externes, n'hésitez pas à contacter la Direction du centre d'expertise en intelligence artificielle du MCN, à l'adresse encadrement.ia@mcn.gouv.qc.ca.

ANNEXE – RÉSUMÉ DES BONNES PRATIQUES INDIVIDUELLES

Résumé des bonnes pratiques individuelles

1. **Ne jamais divulguer d'informations contenant des renseignements confidentiels.**
2. **Reconnaître que l'outil, bien que puissant, peut contenir des erreurs ou des imprécisions.**
3. **Respecter les conditions d'utilisation d'un outil en ligne.**
4. **S'assurer de la conformité au cadre législatif et aux normes éthiques en vigueur.**
5. **S'assurer que le canal de communication réseau avec l'IA est chiffré et vérifier toujours l'URL dans la barre d'adresse.**
6. **Signaler toute information qui pourrait laisser croire à une fuite d'informations ou à la publication d'informations trompeuses, erronées ou diffamatoires au sujet de l'organisation.**
7. **Être prudent dans le choix des outils et se tenir régulièrement informé des mises à jour, des améliorations et des éventuelles limitations de l'outil.**

Ces bonnes pratiques devraient être appliquées pour toute solution informatique. Par exemple, si le correcteur automatique corrige une phrase et en change le sens, le membre du personnel demeure responsable du texte. Il en va de même pour les solutions d'IA générative.

