



Canada School
of Public Service

École de la fonction
publique du Canada



Introduction à l'intelligence artificielle (IA)

Aperçu

- 1 Qu'est-ce que l'IA ?
- 2 Domaines et techniques de l'IA
- 3 L'IA au fil du temps : les dernières percées
- 4 Compétences, données et infrastructure au service de l'IA
- 5 Utilisations et possibilités de l'IA au sein du gouvernement du Canada
- 6 Risques et considérations éthiques
- 7 Environnement juridique et politique
- 8 Annexes : orientation, apprentissage disponible et assistants IA

IA, données et algorithmes

Intelligence artificielle (IA)

Un **système d'IA** est un système basé sur une machine qui déduit comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions à partir des entrées qu'il reçoit.

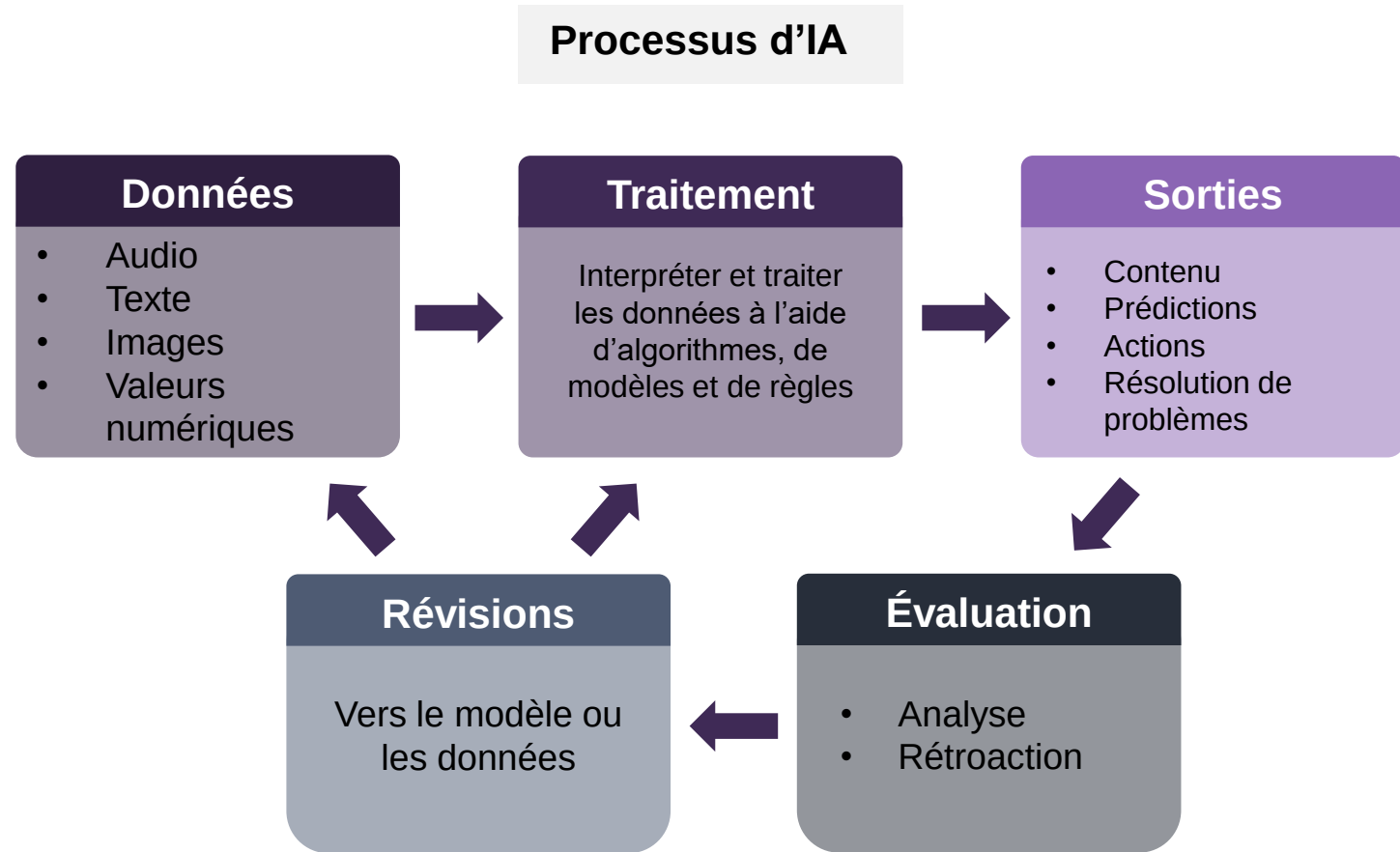
L'IA est aussi **une catégorie de technologies**, dont la définition courante est « la technologie qui effectue des tâches qui nécessiteraient normalement l'intelligence humaine pour être accomplies, telles que donner un sens au langage parlé, apprendre des comportements ou résoudre des problèmes. »¹

Données

Les **données** font référence à des valeurs structurées et non structurées telles que des nombres, du texte, des images et des vidéos. Les systèmes d'IA tirent leur valeur du traitement de quantités massives de données – et sont généralement nécessaires pour traiter ces données en premier lieu.

Algorithme

Un **algorithme** est un ensemble de règles ou d'instructions qu'une machine (et en particulier un ordinateur) suit pour atteindre un objectif particulier.



Qu'est-ce que l'IA ? Domaines et techniques de l'IA

L'IA doit être considérée comme un ensemble de domaines et de sous-domaines interconnectés. Plutôt qu'une seule technologie d'IA, une gamme de techniques et d'approches différentes sont utilisées pour résoudre différents problèmes.

Apprentissage automatique

Techniques de renforcement qui permettent aux ordinateurs d'améliorer les résultats au fil du temps en testant plusieurs approches de traitement dans le modèle et en évaluant les résultats par rapport aux critères de réussite, puis en effectuant des ajustements.

Vision par ordinateur

Méthodes d'acquisition et de compréhension d'images numériques, généralement divisées en reconnaissance d'activités, reconnaissance d'images et vision industrielle.

Réseaux de neurones

Une classe d'algorithmes vaguement calqués sur la structure neuronale du cerveau qui améliore ses performances sans être explicitement instruit sur la façon de le faire.

Traitement automatique du langage naturel

Outils qui interprètent le texte (ou la transcription de la parole) à des fins d'analyse ou pour permettre une interaction conversationnelle avec un logiciel (par exemple, les agents conversationnels, l'IA générative). Peut utiliser l'apprentissage automatique, mais ce n'est pas obligatoire.

De plus en plus répandue

L'IA générative, les grands modèles de langage, les assistants d'IA et les robots ont rapidement émergé et deviennent de plus en plus répandus.

IA générative

Une catégorie d'IA qui accepte le langage naturel et d'autres invites multimédias pour générer un nouveau contenu (texte, images, audio ou autres formes de données) qui est statistiquement probable en réponse à une invite.

Grands modèles de langage (GML)

Les GML alimentent l'IA générative. Ils utilisent des algorithmes d'apprentissage automatique pour traiter de grandes quantités de données et générer des réponses textuelles de type humain basées sur ces données.

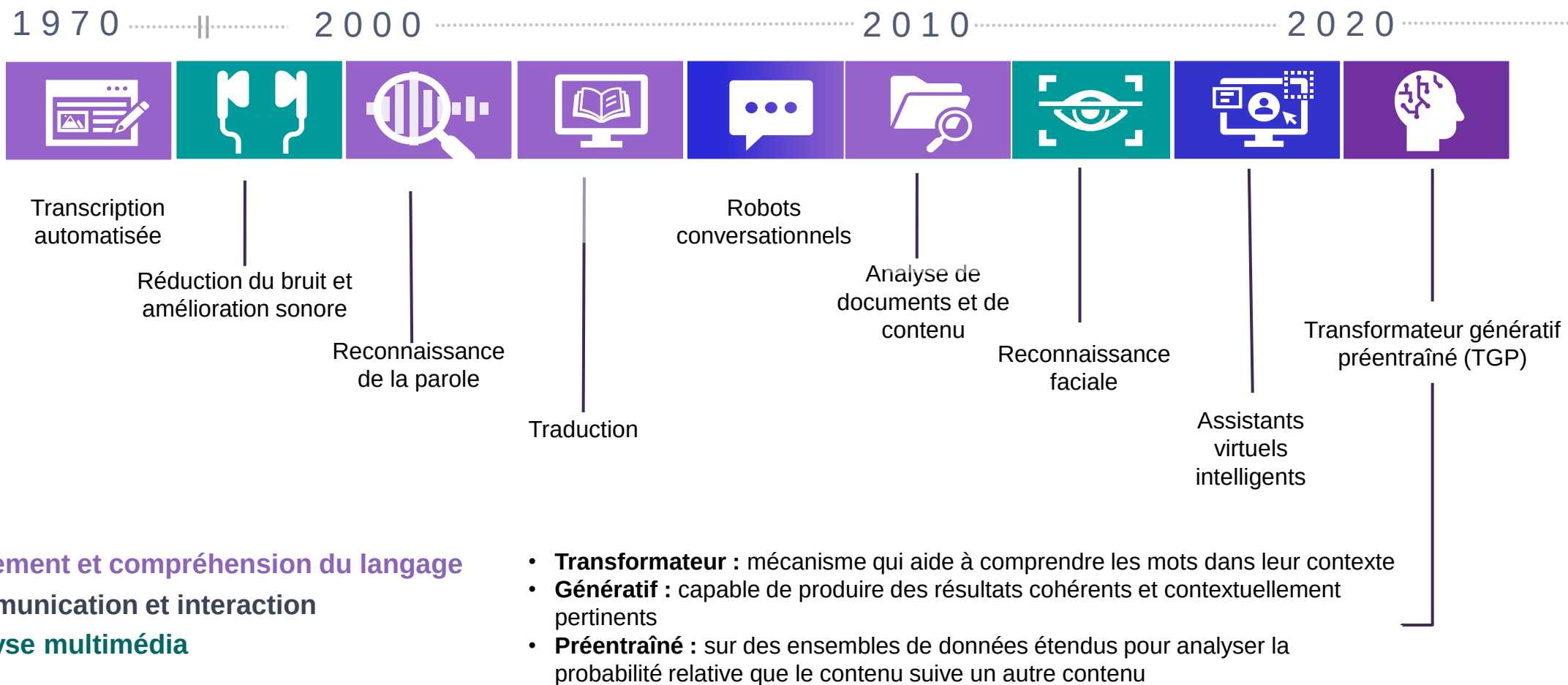
Assistant d'IA

Un logiciel qui utilise l'IA pour augmenter la productivité afin de rationaliser et d'automatiser les flux de travail, de générer du contenu, de connecter des logiciels, de gérer les calendriers, d'aider à la prise de décision, etc.

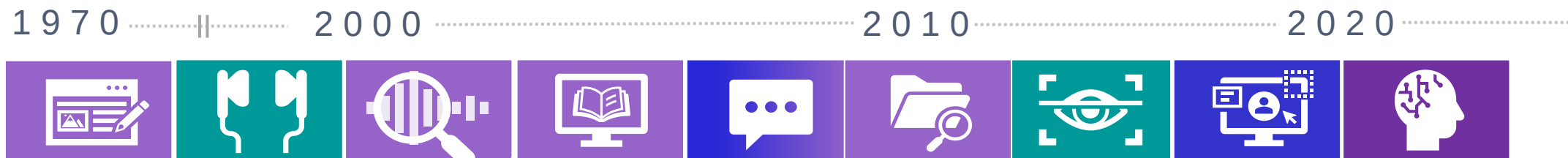
Robot

Une application logicielle qui effectue des tâches automatisées sur Internet et dans des systèmes basés sur des instructions humaines fournies par programmation.

L'IA au fil du temps



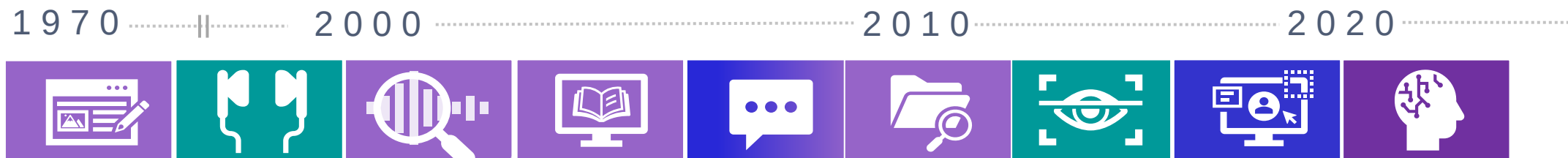
L'IA au fil du temps : les dernières percées



Causes des avancées fulgurantes durant la dernière décennie :

- Augmentation des **données disponibles**
- **Accès à des logiciels**, y compris à code source libre
- Augmentation de la **puissance de calcul** et diminution des coûts
- **Modèles, algorithmes et techniques** partagés ou propriétaires
- **Marchandisation et intégration des outils d'IA**

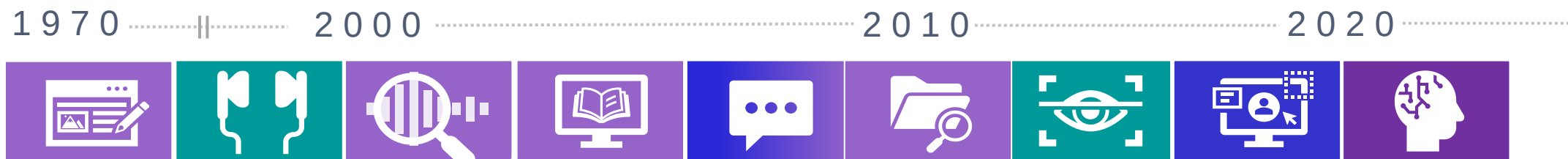
L'IA au fil du temps : les dernières percées



Les dernières percées :

- L'IA générative qui inclut des références à des sources (appelée **génération augmentée de récupération**) pour augmenter la fiabilité des systèmes d'IA
- L'IA générative incorpore des **données propres au GC** et aux programmes derrière les pare-feu du GC
- Génération **multimodale** – texte, audio, images et vidéos

L'IA au fil du temps : la suite



Et ensuite :

- **L'investissement et l'expérimentation** continus dans ce domaine généreront de nouveaux outils et de nouvelles utilisations, bien qu'il y ait des prédictions du cycle d'engouement. Les utilisations les plus fructueuses et les plus durables seront révélées au fil du temps.
- Les systèmes d'IA matures représentent toujours une **IA faible** (les systèmes d'IA sont bons pour des tâches précises), **et non une intelligence artificielle générale** (des systèmes d'IA qui peuvent s'autoaméliorer dans un large éventail de tâches.)

Compétences, données et infrastructure soutenant l'IA

Pour chaque utilisation de l'IA, il y a beaucoup de travail et d'outils sous la surface. En fin de compte, chaque projet d'IA dépend des données.

Application

« Utiliser l'IA pour faire [X] »



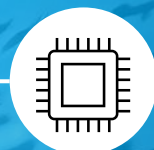
Données

Des données idéalement nettoyées, normalisées, impartiales, exemptes de données personnelles ou de droits d'auteur, et souvent massives et coûteuses à analyser. Le nettoyage et la préparation des données pour l'analyse et le traitement peuvent représenter plus de 50 % du travail et des coûts.



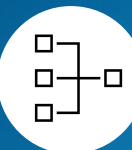
Infrastructure et outils

Une puissance de calcul substantielle et précise ou un accès à des outils infonuagiques, à des modèles libres ou propriétaires, et à des logiciels spécialisés.



Modèles

Données d'entraînement, choix des approches d'IA, méthodes statistiques, algorithmes



Compétences

Science des données, mathématiques, statistiques, recherche, collecte de données et programmation

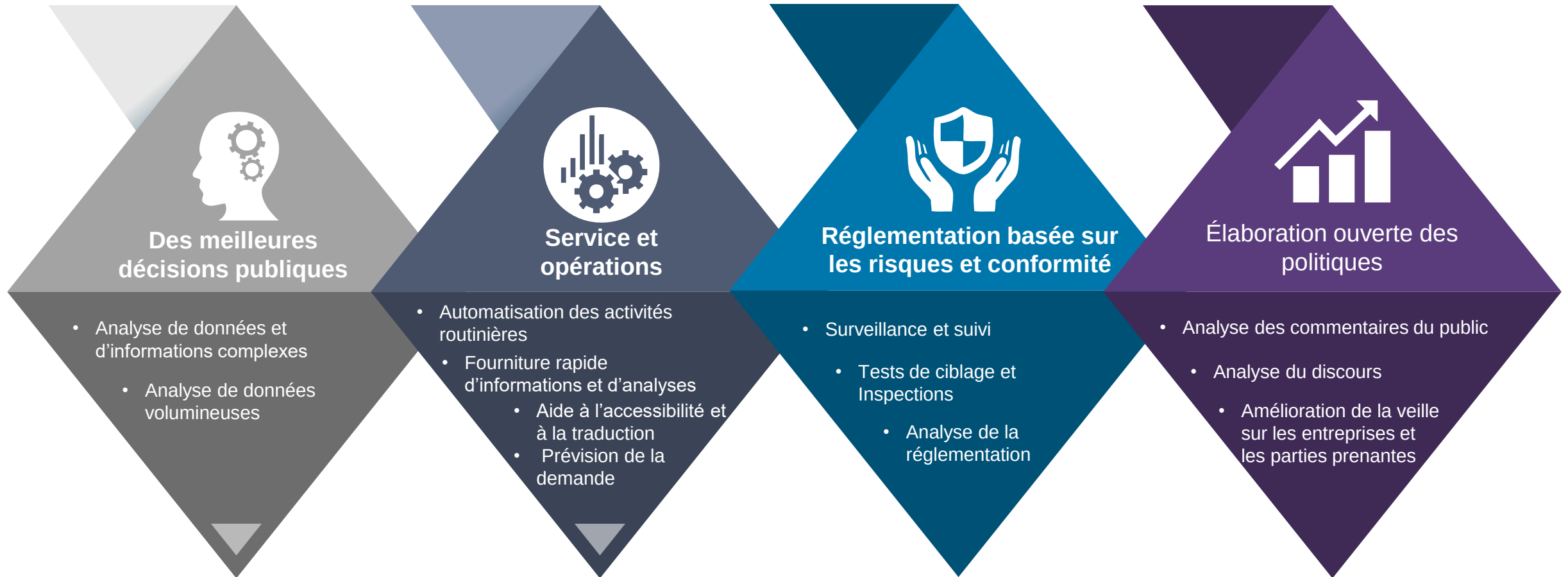


QUESTIONS D'AFFAIRES

En fin de compte, le fondement du travail sur l'IA est un problème commercial bien défini à résoudre



Possibilités et utilisations courant pour l'IA au GC



Risques de l'IA au GC

- 1 **S'adapter continuellement à l'évolution des technologies** dans les opérations.
- 2 **Accuser des retards d'efficacité** pendant que le gouvernement du Canada élabore, achète ou adopte des solutions et des approches conformes aux politiques.
- 3 **Conserver des fonds de données** de plus en plus importants, précieux et liés.
- 4 **Tenir compte des politiques publiques** concernant l'utilisation de l'IA dans tous les secteurs.
- 5 **Protéger la confiance du public** lors de l'utilisation de l'IA pour les activités gouvernementales.
- 6 **Subir l'avantage informationnel** des parties externes qui tirent parti de l'IA dans leurs relations avec le gouvernement.

Considérations éthiques

Au fur et à mesure que l'IA progresse et que son utilisation se généralise, il existe un risque accru qu'elle soit - même involontairement - utilisée à mauvais escient, perpétue les inégalités ou exacerbe les problèmes sociétaux existants. Cette liste ne comprend que quelques-unes des nombreuses considérations éthiques. D'autres peuvent inclure des répercussions sur le marché du travail, les conséquences environnementales et des questions sur la relation des humains avec la technologie.

Partialité et équité

Les biais dans l'IA se traduisent en décisions injustes ou en résultats erronés. Le gouvernement du Canada a la responsabilité de s'assurer que les outils d'IA traitent tout le monde équitablement et sans discrimination.

Le biais peut être un produit de l'**algorithme/modèle** ou des **données d'apprentissage ou d'entrée**.

En bref, les données historiques dans un contexte de racisme et de discrimination systémiques sont susceptibles d'entraîner une sortie biaisée de l'IA.

Transparence et responsabilité

La transparence est nécessaire quant au fonctionnement des systèmes d'IA et, s'ils soutiennent la prise de décision, à la manière dont les données ont été analysées pour produire des résultats. Cela inclut l'**ouverture, la clarté, la traçabilité et l'explicabilité** du système d'IA.

Les acteurs – individus ou organisations – qui exploitent les systèmes d'IA peuvent **ne pas se sentir responsables des actions, des extraits ou des décisions** prises par le système, ni ne les assumer.

Confidentialité, sécurité et gouvernance

Les systèmes d'IA traitent des quantités massives de données, et les outils d'IA sont souvent basés sur le nuage ou sur des ressources, du code ou des modèles créés de l'extérieur.

Tout **traitement de données à caractère personnel ou sensibles** doit être régi et protégé.

Provenance des données et droits d'auteur

La provenance des données à l'origine fait référence à la propriété, à la collecte et à la fiabilité des données sources. Les organisations qui utilisent des données peuvent avoir besoin de suivre et de documenter **les sources, les transformations et l'utilisation des données tout au long du cycle de vie des données**.

De nombreux ensembles de données alimentant l'IA générative en particulier ont des sources de données massives et opaques qui incluent probablement **des informations personnelles ou des œuvres directes ou dérivées protégées par le droit d'auteur**.

Manipulation et tromperie

Cette catégorie comprend les considérations éthiques liées à l'utilisation de l'IA dans la production et la diffusion d'informations trompeuses ou fausses.

Des entités externes peuvent **utiliser ou propager de la désinformation, de la mésinformation, ou des vidéos hypertruquées à l'aide de l'IA**.

L'IA générative peut créer du contenu qui contient de fausses informations.

IA et représentation

Des spécialistes du domaine ont décelé un certain nombre de risques et de méfaits potentiels et prouvés de l'IA. Celles-ci sont susceptibles d'avoir un impact disproportionné sur les communautés marginalisées.

Partialité et équité

Les biais dans l'IA signifient des décisions injustes ou des résultats faussés. Le gouvernement du Canada a la responsabilité de s'assurer que les outils d'IA traitent tout le monde équitablement et sans discrimination.

Le biais peut être un produit de l'algorithme/modèle ou des données d'apprentissage ou d'entrée.

En bref, les données historiques dans un contexte de racisme et de discrimination systémiques sont susceptibles d'entraîner une sortie biaisée de l'IA.

1

Absence ou sous-représentation dans les données d'entraînement : Par exemple, dans les simulations, on a constaté que les voitures autonomes ne s'arrêtaient pas pour les personnes en fauteuil roulant, ce qui était absent des données d'entraînement.*¹

L'IA générative sous-représente les communautés minoritaires dans les images générées de nombreuses professions; Les tentatives de corriger cela par le biais d'algorithmes de traitement ont jusqu'à présent plutôt créé des images stéréotypées codées racialement.²

Les médias sociaux ou les applications de diffusion de contenu sont susceptibles de générer des recommandations basées sur « ce que les gens aiment », ce qui cloisonne les communautés ayant des expériences de vie différentes.

2

Surreprésentation dans les données d'entraînement : Par exemple, des spécialistes ont démontré que les systèmes de police prédictive surveillaient systématiquement les communautés marginalisées, surreprésentées dans les données de formation.³

*Jutta Treviranus, directrice du Laboratoire de conception inclusive d'OCAD, fait remarquer que l'IA peut être une « un couteau à double tranchant » pour les personnes handicapées : il y a un risque que des décisions ou des systèmes soient fondés sur des données qui les excluent, ou que des systèmes produisent des résultats qui servent la majorité parce qu'ils sont conçus pour l'efficacité. D'autre part, les technologies d'automatisation (par exemple, les voitures autonomes) pourraient également créer des options et des soutiens pour les personnes.

L'environnement juridique et politique

Pour l'usage interne du gouvernement du Canada :

- [Directive sur la prise de décisions automatisée](#)
- [Guide sur l'utilisation de l'intelligence artificielle générative](#)
- Le GC vient de lancer une stratégie d'IA pour la fonction publique fédérale, qui devrait être terminée à l'automne 2025.

Pour l'industrie et la société :

- [Loi sur l'intelligence artificielle et les données](#) (en cours de développement)
- [Code de conduite volontaire visant un développement et une gestion responsables des systèmes d'IA générative avancés](#)

La Directive exige une évaluation de l'impact algorithmique dans le cadre de laquelle l'**utilisation de l'IA par le gouvernement du Canada appuie ou rend des décisions administratives** concernant des personnes.

La *Loi sur l'intelligence artificielle et les données* a été déposée en 2022 dans le cadre du projet de loi C-27. En date de janvier 2024, elle fait l'objet d'une étude de comité.

On s'attend à ce que les institutions fédérales **suivent les principes** d'une utilisation pertinente, responsable, équitable, transparente, éclairée, sécurisée < PRETES > de l'IA générative

Il faut consulter la **direction de l'information et de la sécurité** et utiliser des systèmes qui ont fait l'objet d'un contrôle de confidentialité et de sécurité ou figurent sur la liste d'approvisionnement de l'IA.

Annexe A : Principes pour une utilisation responsable de l'IA générative (PRETES)

Pertinente

Veiller à ce que l'utilisation d'outils d'IA générative réponde aux besoins des utilisateurs et de l'organisation et contribue à améliorer les résultats pour la population canadienne; déterminer les outils appropriés pour la tâche; les outils d'IA ne sont pas le meilleur choix dans toutes les situations.

Responsable

Assumer la responsabilité du contenu généré par ces outils, notamment veiller à ce que les informations soient factuelles, légales, éthiques, et conformes aux conditions d'utilisation.

Équitable

Veiller à ce que le contenu de ces outils ne comporte pas ou n'augmente pas les biais et qu'il respecte les droits de la personne, l'accessibilité et les obligations en matière d'équité de procédurale et de fond.

Transparente

Identifier le contenu produit à l'aide de l'IA générative; informer les utilisateurs qu'ils interagissent avec un outil d'IA; documenter les décisions et être en mesure de fournir des explications si des outils sont utilisés afin d'appuyer la prise de décisions.

Éclairée

Apprendre à connaître les points forts, les limites et l'utilisation responsable des outils; apprendre à créer des messages-guides efficaces et à relever les faiblesses potentielles des résultats.

Sécurisée

Garantir que l'infrastructure et les outils sont adaptés à la classification de sécurité des renseignements et que la vie privée et les renseignements personnels sont protégés.

[Guide sur l'utilisation de l'intelligence artificielle générative](#)

Annexe B : Utilisations des assistants d'IA

Inoffensif	Transcription automatisée	Traduction	Réduction du bruit et amélioration de la sonorité	Fonctionnalités d'accessibilité
De plus en plus attendu	Assistants virtuels intelligents	Établissement intelligent du calendrier	Résumés de réunion et mesures de suivi	Analyse des documents et du contenu
Contentieuse	Suivi des interventions	Reconnaissance des gestes	Analyse des sentiments	Reconnaissance faciale
	+ d'usages naissants à mesure que le marché des assistants d'IA connaît une expansion et une expérimentation rapides			

Annexe B : Les avantages opérationnels de l'information en temps réel

Les assistants d'IA sont déjà intégrés dans les principales applications logicielles d'entreprise (par exemple, Word, Excel, Teams, PowerPoint)

Dans un contexte de réunion, cela signifie une **intégration avec les systèmes de gestion de la relation client (CRM)** (par exemple, Dynamics, Salesforce)

Les parties rencontrant des fonctionnaires auront de plus en plus un **accès riche et en temps réel** à :

Des faits et des chiffres

L'analyse des sentiments

Les positions de principe

**Les invitations de conversation
suggérées**

**Les notes tirées de toutes les réunions antérieures avec une personne ou
une organisation**

Annexe B : Vie privée, éthique, gestion de l'information et sécurité

Si une réunion est **transcrite et résumée**, elle est enregistrée.

Les robots et les logiciels intégrés **ne déclenchent pas toujours un avis d'enregistrement**.

Les transcriptions et les enregistrements sont alors détenus par :

- d'autres organisations
- Souvent, par des fournisseurs d'assistants d'IA en nuage

Certains logiciels d'assistant d'IA pourraient avoir une **faible protection des données** – ou être **explicitement conçus pour recueillir des données**.

Pour les enregistrements et les transcriptions enregistrés par les fonctionnaires du GC, les **documents relèveront** des cadres de lois et de politiques sur **la gestion de l'information ainsi que l'accès à l'information et la protection des renseignements personnels**.

Annexe B : gestion responsable de la vie privée

Veiller à ce que **toutes les personnes participant** à une réunion **soient identifiées et connues**.

Se demander si les assistants – **qu'ils soient humains ou une IA** – devraient participer à la discussion de la réunion.

En général, nous **ne pouvons pas supposer qu'il n'y a pas d'IA et qu'il n'y a pas d'enregistrement** – un participant pourrait avoir un appareil autonome qui enregistre, ou encore un assistant à domicile comme Alexa.

En fin de compte, tout dépend de la confiance envers les personnes participant à la réunion et au type de renseignements abordés.

Annexe C : ressources d'apprentissage de l'EFPC

Cours

[Découvrez l'intelligence artificielle](#)

[Utiliser l'IA générative au gouvernement du Canada](#)

[Considérations éthiques liées à l'intelligence artificielle](#)

Microapprentissage

[Mieux comprendre les assistants d'IA servant aux vidéoconférences](#)

[Utilisation des grands modèles de langage \(comme ChatGPT\) dans la fonction publique fédérale](#)

[Démystifier l'intelligence artificielle](#)

[Aperçu de ChatGPT d'OpenAI](#)

Événements

[Série sur l'intelligence artificielle](#)