

Avril 2023

Forum sur l'intelligence artificielle dans le secteur des services financiers:

# Une perspective canadienne sur l'intelligence artificielle responsable



# Table des matières

**03** Avant-propos

**04** Sommaire

**07** Introduction

**09** Explicabilité

- 10 Aspects de l'explicabilité
- 11 IA explicable
- 13 Communication d'informations
- 15 Explicabilité et confiance

**16** Données

- 18 Caractéristiques des données et défis connexes
- 20 Gouvernance des données
- 22 Données de tiers
- 23 Alignement de la stratégie des données et de la stratégie commerciale

**24** Gouvernance

- 25 Cadre de gouvernance de l'IA
- 27 Évolution des cadres de gouvernance existants pour l'IA
- 30 Compétences, culture et autres défis
- 31 Prochaines étapes

**33** Éthique

- 34 Répercussions juridiques, politiques et réglementaires
- 35 Considérations et défis
- 36 Objectifs commerciaux, données subjectives et injustice
- 37 Équité
- 39 Vie privée et droit de recours
- 41 Opérationnalisation de l'éthique et des structures organisationnelles de l'IA

**42** Réglementation

- 42 État de la réglementation et des politiques en matière d'IA des diverses administrations
- 46 Caractéristiques d'une réglementation efficace
- 47 Point de vue des organismes de réglementation

**49** Conclusions

**51** Annexe

- 51 Remerciements
- 51 Conférenciers visionnaires
- 52 Autres conférenciers
- 53 Les participants

**54** Références

## Avis de non-responsabilité

Le contenu de ce rapport reflète les points de vue et les idées des conférenciers et des participants du FIASSF. Le contenu du présent rapport ne doit pas être interprété comme des directives du Bureau du surintendant des institutions financières (BSIF) ou de tout autre organisme de réglementation, actuel ou futur.

# Avant-propos

En raison de l'évolution rapide, de la prolifération et de la nature transformatrice de la technologie de l'intelligence artificielle (IA), il est désormais indispensable de tenir compte des répercussions éthiques, juridiques, financières et sociales de son développement et de son déploiement. Il devient donc nécessaire de bien concevoir les pratiques de gestion du risque, notamment des cadres d'évaluation et de mises à l'essai solides, des processus décisionnels clairs et transparents, et des mécanismes de responsabilisation et de recours en cas de préjudice.

On met souvent l'accent sur la « moyenne » de distribution des résultats pour justifier l'utilisation et améliorer la validité des applications d'IA. Cependant, selon l'approche de réflexion basée sur le risque, il faut évaluer l'ensemble de ces « queues » de distribution avec ouverture d'esprit et créativité pour être en mesure d'atténuer les imprévus et les conséquences désastreuses possibles. C'est d'ailleurs dans une ambiance de créativité et d'exploration que s'est déroulé le Forum sur l'intelligence artificielle dans le secteur des services financiers (FIASSF). Cet événement rassemblait les meilleurs experts de l'application de l'IA du secteur des services financiers du Canada, ainsi que des représentants d'organismes gouvernementaux et du milieu universitaire. Les participants ont pu entendre d'éminents experts en IA et des représentants d'organismes de réglementation influents de plusieurs pays. Puis, ils ont eu l'occasion de discuter de ces apprentissages dans le but d'exploiter en toute sécurité le potentiel notable de l'IA dans le contexte de l'économie canadienne.

Le présent rapport résume ces discussions et offre un cadre d'examen des défis et des possibilités pour l'élaboration d'une réglementation efficace.

Malgré la multitude des points de vue exprimés au cours du forum, un message sans équivoque et unanime est ressorti : la nécessité et la volonté d'établir une collaboration multidisciplinaire, authentique et efficace. En effet, la création d'équipes multidisciplinaires facilite non seulement l'adoption plus rapide et harmonieuse de la technologie de l'IA par les entreprises individuelles, mais représente également une occasion pour les membres de cet écosystème élargi de s'attaquer aux risques et défis émergents communs.

Nous tenons à exprimer notre reconnaissance à l'endroit de tous les participants qui ont transmis leurs connaissances, donné de leur temps, fait preuve de candeur et apporté leurs précieuses contributions aux discussions. Nous espérons que ce rapport contribuera à l'adoption responsable et sécuritaire de l'IA dans les services financiers canadiens et qu'il guidera une discussion plus large sur l'IA.



Angie Radiskovic  
Directrice générale, Quantification des risques  
Bureau du Surintendant des Institutions  
Financières



Sonia Baxendale  
Présidente et directrice générale  
Global Risk Institute

# Sommaire

Grâce aux grandes améliorations qui ont été réalisées en matière de disponibilité et d'accessibilité des données ainsi qu'aux avancées dans les techniques de modélisation, des modèles sont maintenant appliqués à un nombre croissant d'activités dans les institutions financières réglementées au Canada. Les capacités de l'IA et l'utilisation qui en est faite ont par contre évolué plus rapidement que la réglementation.

Le BSIF s'est donc associé à l'Institut du risque mondial (Global Risk Institute) pour mettre sur pied une collectivité de leaders d'opinion en IA. Ils sont issus du milieu universitaire, des organismes de réglementation, des banques, des assureurs, des régimes de retraite, des

technologies financières et des centres de recherche. Ce groupe, appelé Forum sur l'intelligence artificielle dans le secteur des services financiers (FIASSF), a pour objectif de faire avancer la conversation sur les mesures de protection appropriées et la gestion des risques liés à l'utilisation de l'intelligence artificielle (IA) dans les institutions financières.

Au cours des ateliers du forum, les idées mises de l'avant pour soutenir le développement sécuritaire de l'IA ont été regroupées selon quatre principes E-D-G-E, soit l'explicitabilité, les données, la gouvernance et l'éthique.



**L'explicabilité** dépend à la fois du cas d'usage et du cadre de gouvernance; ce principe doit être pris en considération dès le début du processus de conception du modèle d'IA. Pour démontrer que les objectifs de modélisation ne se limitent pas au rendement, des exemples privilégiant la sélection d'un modèle explicable au lieu d'un modèle plus performant ont été fournis. Les participants se sont également penchés sur les motifs qui pourraient justifier, dans les cas d'usage ayant d'importantes répercussions, l'utilisation de modèles intrinsèquement explicables plutôt que le recours aux explications post-hoc.

Les institutions financières travaillent depuis longtemps avec **les données**, mais l'intégration de l'IA (ainsi que des sources de données et des algorithmes automatisés d'autorégulation qui en découlent) dans leurs opérations a présenté de nouveaux défis en matière de gestion et d'utilisation des données. Il est donc possible qu'elles éprouvent plus de difficulté à intégrer et à normaliser les contrôles permettant de gérer le risque lié aux données, surtout lorsqu'on tient compte de l'augmentation du volume de données, de l'intensification de la génération et de l'utilisation de ces données, et de l'émergence de nouvelles sources et de nouveaux types de données. Ces difficultés sont accentuées lorsque les données sont cloisonnées au sein d'une organisation ou lorsqu'elles proviennent de sources externes différentes. L'amélioration des données utilisées pour entraîner un algorithme aura une incidence directe sur la performance du modèle. Par conséquent, il est important que les institutions financières harmonisent leurs activités et leurs données pour ainsi recueillir, gérer et analyser les données qui cadrent avec leurs objectifs. D'ailleurs, pour que les systèmes d'intelligence artificielle fonctionnent efficacement, il est essentiel que les données

soient exactes, cohérentes et exhaustives, ce qui peut être obtenu grâce à une bonne gouvernance des données.

**La gouvernance** a beaucoup gagné en importance avec le temps. Au fil de son évolution, elle a acquis les propriétés suivantes : elle doit être intégrée et s'appliquer à tous les paliers de l'organisation, les rôles et responsabilités doivent être clairs, la tolérance au risque doit être bien définie et elle doit pouvoir s'adapter à mesure qu'une institution chemine dans son adoption de l'IA. En outre, la gouvernance du modèle d'IA exige une approche multidisciplinaire pour être efficace. Lorsque la gouvernance est exercée de manière machinale, la compréhension des risques est reléguée au second plan, et l'accomplissement des divers éléments du cadre prescrit prend le dessus, peu importe le risque associé à ces éléments ou leur pertinence.

Le concept **d'éthique**, quant à lui, est très nuancé et est naturellement subjectif, en plus de reposer sur des normes qui sont appelées à changer au fil du temps. D'ailleurs, la codification de ces normes dans les lois et règlements témoigne des défis et de la complexité de l'éthique de l'IA. Il n'existe, en effet, pas de définition universelle de l'équité. Qui plus est, la perception qu'on peut en avoir est influencée par le contexte et le point de vue. Dans le domaine de l'équité algorithmique, il existe différentes définitions mathématiques, dont plusieurs qui se contredisent. Si on fait abstraction de l'aspect juridique, le respect de la loi ne signifie pas toujours que les actions et les résultats sont équitables ou perçus comme tels. Il existe de nombreux cas d'usage pour lesquels le résultat souhaité est empreint de partialité, notamment les politiques de tarification et la stratification du risque. De plus, les données utilisées pour l'entraînement de l'IA peuvent être la source de biais et de résultats injustes. Les

institutions financières utilisent actuellement une approche « d'équité par l'ignorance » dans leurs modèles pour lutter contre la discrimination. Cette approche fait abstraction de certaines données personnelles dans la prise de décision et, en fait, ne les recueille même pas. Leurs modèles sont donc « aveugles » à ces données, mais ne sont pas nécessairement neutres sur le plan des résultats. De plus, l'absence de ces données complexifie la mise à l'essai des modèles. Les exigences de la société en matière de maintien de normes éthiques élevées ne cessent de croître. Les institutions financières courent alors des risques réputationnels bien réels et peuvent en subir les répercussions lorsque des préjudices imputables à un manquement éthique, réel ou perçu, sont causés aux clients. Il est donc important que les organisations fassent preuve de transparence, tant à l'interne qu'à l'externe, et communiquent de manière proactive ce qu'elles font pour garantir que leurs modèles d'IA respectent des normes éthiques élevées.

À l'échelle mondiale, les organismes de réglementation doivent trouver le juste équilibre entre la réglementation et l'innovation; c'est-à-dire l'équilibre entre l'établissement d'une réglementation solide et le maintien de la capacité d'adaptation et de la concurrentialité des institutions financières. Les différentes approches utilisées à travers le monde pour réglementer l'IA varient grandement. Certaines institutions comme la Banque d'Angleterre adoptent une approche fondée sur des principes, tandis que d'autres comme l'Autorité monétaire de Singapour y vont de directives normatives plus granulaires.

Le BSIF, quant à lui, se tourne vers l'avenir et propose une version améliorée de la ligne directrice E-23, *Gestion du risque de modélisation à l'échelle de l'entreprise*, dont une version à l'étude sera publiée aux fins de consultation publique en 2023. Les observations et les commentaires émis par les participants lors du FIASSF témoignent de la nécessité de collaborer et d'adopter une approche multidisciplinaire. On remarque également un fort appétit pour un dialogue continu sur le sujet dans le secteur canadien des services financiers.





# Introduction

Le recours à la numérisation et l'utilisation de l'IA se sont intensifiés rapidement dans le secteur des services financiers. À mesure que l'utilisation de ces technologies croît, il faut veiller à ce que les cadres actuels de gestion des risques liés à l'IA demeurent pertinents, tournés vers l'avenir et adaptatifs selon les besoins du secteur. Aux fins du présent rapport, les outils, les modèles, les applications et les systèmes d'IA seront utilisés de manière interchangeable.

En septembre 2020, le Bureau du surintendant des institutions financières (BSIF) a publié un document de travail qui énonçait les principes fondamentaux de la gestion des risques associés à l'utilisation de l'IA par les institutions financières. Ce document traitait de l'incidence de la robustesse, de l'explicabilité et de la responsabilité dans les modèles d'IA. Depuis la publication de ce document de travail sur les risques technologiques, le BSIF a mené des enquêtes sectorielles et des études approfondies auprès de certaines institutions financières et a tenu des échanges bilatéraux avec des centres de recherche et des forums sectoriels sur des sujets connexes afin de mieux comprendre les utilisations et les écueils de l'IA.

Le BSIF a reconnu qu'une approche sectorielle visant l'adoption sécuritaire de l'IA dans les services financiers était nécessaire. L'Institut du risque mondial (IRM) a donc dirigé un projet en partenariat avec le BSIF pour créer le Forum sur l'intelligence artificielle dans le secteur des services financiers (FIASSF). Ce forum, composé d'une série de quatre ateliers modérés, a attiré des participants des secteurs public et privé. Certains invités ont été appelés à présenter leurs points de vue, notamment sur l'évolution de l'IA, sur les thèmes émergents et sur les cas d'utilisation, ainsi que sur les défis liés à l'adoption de l'IA, à son développement, à son déploiement et à son utilisation dans les services financiers. Parmi les sujets abordés, on retrouvait entre autres : les éléments du secteur des services financiers qui peuvent être mis à profit en lien avec l'IA, les pratiques exemplaires à adopter pour relever les défis, les méthodes de gestion et d'atténuation des risques associés aux modèles et la gouvernance des modèles.

Le premier atelier de la série a permis aux participants de déterminer les thèmes les plus importants en lien avec les modèles d'IA, soit l'explicabilité, les données, la gouvernance et l'éthique, collectivement appelés les principes E-D-G-E. Ces thèmes ont ensuite constitué la base de discussion des ateliers suivants ainsi que la structure sur laquelle repose le présent rapport.

## Les principes E-D-G-E

---



**L'Explicabilité** permet aux institutions financières de renforcer la relation de confiance avec leurs clients. Lorsque les clients comprennent le raisonnement qui guide les décisions, ils sont plus à même de travailler vers l'atteinte de leurs objectifs financiers.



**Les Données utilisées** en IA permettent aux institutions financières de fournir des produits et services ciblés et personnalisés à leurs clients, d'améliorer la détection des fraudes, d'améliorer l'analyse et la gestion des risques, de stimuler l'efficacité opérationnelle et d'améliorer la prise de décision.



**La Gouvernance** appuie la réalisation du potentiel de l'IA en veillant à ce que l'institution financière dispose de la culture, des outils et des cadres appropriés pour soutenir le cycle de vie de l'IA.



**L'Éthique** encourage les institutions financières à prendre en compte les retombées sociétales plus larges de leurs systèmes d'IA et à définir soigneusement le rôle qu'elles aimeraient jouer dans le façonnement du monde qui les entoure.

Le présent rapport se penche donc sur les principales conclusions tirées lors de ce forum qui explorait l'intégration de l'IA dans le secteur des services financiers. Il s'appuie sur les discussions pour faire ressortir les points d'accord et de désaccord. En outre, il tient compte des points de vue des conférenciers principaux pour offrir une perspective plus large de chacun des thèmes du forum. Nous avons espoir que ce rapport contribuera à promouvoir le renforcement des pratiques de gestion des risques tout en aidant les praticiens à gérer de manière efficiente les défis que présente l'IA, à tirer profit de son potentiel et à prendre connaissance des approches à privilégier en ce qui a trait à l'adoption et à la gestion de la technologie.





# Explicabilité

Il est difficile d'expliquer le processus des modèles; il s'agit d'ailleurs de l'un des enjeux les plus importants et persistants associés à l'utilisation des techniques d'IA.

L'explicabilité permet d'examiner la théorie, les données, la méthodologie et d'autres aspects fondamentaux d'une approche afin de confirmer que le modèle est adapté à l'utilisation prévue avant de passer à l'étape de la vérification de sa performance.

L'explicabilité permet de faciliter l'évaluation de l'équité des modèles d'IA et, de manière générale, contribue à lutter contre les préjugés et la discrimination. Les organismes de réglementation commencent à se pencher officiellement sur la question de l'explicabilité, qui d'ailleurs fait l'objet de plus en plus d'études.

Les questions clés suivantes sur l'explicabilité ont été abordées pendant le forum :

Quels sont les degrés d'explicabilité possibles des systèmes d'IA?

Quels facteurs permettent de déterminer le degré approprié d'explicabilité d'une application particulière?

Quelles sont les approches à adopter pour parvenir à l'explicabilité et quels sont les risques associés?

Comment l'explicabilité est-elle liée au concept plus général de transparence?

Quel est le rôle de l'explicabilité dans le renforcement de la relation de confiance?

**“L'explicabilité permet notamment de s'assurer que la bonne décision est prise pour les bonnes raisons.”**

Alexander Wong, professeur et titulaire d'une chaire de recherche du Canada, Université de Waterloo



# Aspects de l'explicabilité

Le degré d'explicabilité requis pour un modèle doit être pris en considération dès le début du processus de sélection et de conception du modèle; il est établi en fonction du cas d'utilisation et du cadre de gouvernance. Par exemple, l'explicabilité peut être utile aux experts en science des données pour faciliter le débogage et permettre un meilleur repérage des moyens d'améliorer la performance et la robustesse des modèles d'IA. L'explicabilité peut aider un propriétaire d'entreprise à comprendre et à mieux gérer les risques découlant de l'utilisation d'outils d'IA en plus d'aider les organismes de réglementation à certifier la conformité. Les clients peuvent avoir besoin d'explications pour comprendre une décision et les modifications qu'ils peuvent apporter à leurs comportements pour influencer les décisions futures. De façon générale, l'explicabilité d'un modèle aide les institutions financières à évaluer si les décisions prises par l'IA correspondent à leurs valeurs d'entreprise, en plus de contribuer à une utilisation responsable de l'IA.

## Degrés adéquats d'explicabilité

Les différents degrés d'explicabilité reflètent notre compréhension du fonctionnement d'un modèle pour arriver à des résultats. Les modèles entièrement transparents ont un degré d'explicabilité élevé comparativement aux modèles moins transparents dont le degré d'explicabilité est faible. Il demeure toutefois possible d'expliquer localement et globalement les résultats des modèles dont le degré d'explicabilité est faible grâce à des techniques d'analyse post-hoc.

Le degré d'explicabilité adéquat d'un modèle selon le cas d'utilisation ou le groupe d'intervenants fait l'objet de discussions continues. Selon certains, l'explicabilité est importante dans les cas d'utilisation puisqu'elle permet d'assurer une utilisation éthique des modèles d'AI. Par exemple, au titre du Règlement général sur la protection des données (le RGPD) de l'Union européenne (UE), les utilisateurs ont un « droit d'explication » à l'égard de toutes décisions prises par des systèmes automatisés. L'UE est une des premières compétences à avoir adopté une telle réglementation, et on note un intérêt à l'échelle mondiale pour l'élaboration de directives de ce genre. Bon nombre de participants au forum ont fait valoir qu'ils seraient à l'aise d'utiliser des modèles en « boîte noire », moins explicables, pour prendre des décisions s'il y avait suffisamment de contrôles en place pour empêcher le modèle de fonctionner en dehors des limites prédéfinies.

Les participants ont d'ailleurs tous convenu que le degré adéquat d'explicabilité requis dépend de plusieurs facteurs, notamment :

- **Que faut-il expliquer?** Certaines approches offrent la possibilité d'expliquer la signification de chaque variable d'une prédiction donnée (p. ex. l'importance des caractéristiques comme dans les valeurs de Shapley). Il est toutefois possible de s'efforcer d'atteindre un degré d'explicabilité plus élevé pour comprendre le fonctionnement global du modèle (p. ex. en utilisant des modèles interprétables tels que les arbres de décisions). Un autre élément qu'il pourrait être nécessaire d'expliquer est le rôle



de l'IA dans la décision définitive. En effet, il pourrait être important de savoir si la décision a été prise directement par un système d'IA ou s'il s'agissait d'une prise de décision assistée par l'IA.

- **À qui faut-il fournir une explication?**

Le degré d'explication varie en fonction du destinataire (expert en science des données, propriétaire d'entreprise, organisme de réglementation, client, etc.). En effet, une explication qui convient pour un client pourrait être insuffisante pour un expert en science des données puisque les raisons qui motivent le besoin d'explicabilité ne sont pas les mêmes.

- **Est-ce qu'il s'agit d'un cas d'utilisation de grande importance?**

Le besoin d'explications est moins important pour les robots conversationnels ou les modèles d'IA utilisés en commercialisation que pour les modèles d'IA qui servent à la prise de décisions sur le crédit ou à la mesure des fonds propres. Il est communément admis que des degrés plus élevés d'explicabilité sont requis pour les applications dont les fonctions ont une importance relative plus élevée.

- **Quelle est la complexité du modèle?**

Certains modèles peuvent être complexes à un point tel que le modèle lui-même ainsi que les résultats générés sont très difficiles à expliquer. Ces modèles pourraient alors être considérés comme inappropriés dans certains cas d'utilisation.

## IA explicable

Certains modèles sont conçus pour être explicables, alors que d'autres sont considérés comme des modèles en boîte noire qui nécessitent des techniques supplémentaires pour permettre d'en comprendre les résultats. Ces techniques supplémentaires sont communément appelées techniques post-hoc. Les modèles qui sont conçus pour être explicables sont également appelés modèles intrinsèquement interprétables ou explicables; leurs mécanismes internes peuvent être intrinsèquement analysés et compris.

L'utilisation de techniques pour aider à comprendre ou à interpréter les résultats des modèles constitue le domaine appelé *l'IA explicable*. Ces techniques permettent de créer un modèle distinct qui vise à reproduire le comportement du modèle d'IA en boîte noire. En d'autres mots, l'« IA explicable » admet tacitement qu'il existe deux modèles : le modèle en boîte noire fournissant des décisions qui doivent être expliquées et le modèle intrinsèquement explicable conçu pour reproduire le comportement du modèle en boîte noire.





### Explications impartiales

Des chercheurs de l'Université de Californie (Irvine) et de l'Université Harvard ont démontré que les techniques d'explication LIME et SHAP pourraient ne pas être fiables. En effet, grâce à une évaluation approfondie de plusieurs ensembles de données du monde réel, ils ont démontré comment des classificateurs tendancieux conçus par leur cadre pouvaient facilement tromper les algorithmes LIME et SHAP et les amener à générer des explications anodines qui ne reflétaient pas les biais sous-jacents. <sup>[1]</sup>

### Approches de l'explicabilité

Lors des discussions sur l'IA explicable, il a également été question des différents types d'explications et de la distinction entre « comprendre une décision donnée » (explication locale) et « comprendre un modèle » (explication globale). Dans le cas des explications locales, un expert en science des données peut être en mesure de fournir une explication de la décision prise (par exemple, l'octroi d'un prêt), mais il ne connaîtra peut-être pas le fonctionnement interne plus large du modèle pour toutes les décisions.

Certaines techniques servent à fournir des explications locales; d'autres à fournir des explications globales. Par exemple, pour expliquer la contribution de chaque donnée utilisée dans la prise d'une décision particulière par un modèle de type boîte noire, on peut utiliser la méthode d'interprétation locale agnostique LIME (Local Interpretable Model-agnostic Explanations) ou la méthode des explications additives de Shapley (SHAP).

Ces techniques fournissent des informations sur les résultats des modèles, mais elles ne sont pas parfaites. En effet, selon l'opinion commune, l'explicabilité peut aider à détecter des biais. Cependant, il faut savoir que les algorithmes LIME et SHAP pourraient ne pas détecter certains biais

(c.-à-d. que les biais sont présents, mais ne se reflètent pas dans les explications).

De plus, il convient de noter que certaines techniques ou certains modèles utilisés pour expliquer un modèle de type « boîte noire » peuvent donner lieu à différentes explications pour la même prédiction. Pour résoudre le problème des techniques ou modèles d'explication post-hoc, il est recommandé de valider les modèles de type « boîte noire » utilisés ainsi que les modèles correspondants utilisés pour les expliquer.

### Compromis interprétabilité-performance

Les participants au forum ont exploré la possibilité d'un compromis entre l'interprétabilité et la performance du modèle, particulièrement dans le cas de certains modèles complexes et puissants, tels que les réseaux neuronaux dont la performance est élevée, mais qui présentent des degrés d'explicabilité faibles, contrairement aux modèles intrinsèquement interprétables.

Les participants au forum étaient divisés quant à un compromis entre l'interprétabilité et la performance. Certains estimaient que la meilleure approche, malgré sa dépendance au cas d'utilisation, consistait à utiliser des modèles intrinsèquement interprétables et à consacrer du temps et de l'argent à l'amélioration des données. Plusieurs participants au forum ont



convenu qu'avec des améliorations significatives (qualité, disponibilité) des données, la performance d'un modèle intrinsèquement interprétable pourrait être comparable à la performance d'un modèle en boîte noire. D'autres participants étaient à l'aise de s'appuyer sur les explications post-hoc fournies pour leurs modèles très performants et peu interprétables en soi.

### Explicabilité et performance

Il existe une croyance répandue selon laquelle les modèles plus complexes sont plus précis, ce qui signifie qu'une boîte noire compliquée est nécessaire pour optimiser les performances prédictives. Cependant, c'est souvent faux selon Cynthia Rudin, professeur à l'université Duke. Même pour des applications telles que la vision par ordinateur, où l'apprentissage profond entraîne des gains de performance importants et où l'interprétabilité est beaucoup plus difficile à définir, certaines formes d'interprétabilité peuvent être intégrées directement dans les modèles sans perte de précision.<sup>[2]</sup>

## Communication d'informations

La communication d'informations est importante pour les institutions financières parce qu'elle contribue à promouvoir la transparence et la responsabilisation.

Les institutions financières sont tenues de communiquer des renseignements adéquats et pertinents aux investisseurs, aux organismes de réglementation et au public pour les aider à comprendre la santé et le rendement financiers de l'institution. Ces renseignements permettent aux investisseurs de prendre des décisions éclairées quant aux occasions d'investissement au sein de l'institution et aide les organismes de réglementation à s'assurer que l'institution se conforme aux lois et règlements. De plus, cette communication contribue à renforcer la confiance des clients envers leur institution financière, car elle démontre que l'institution est ouverte et honnête au sujet de ses opérations.

La communication d'informations adéquates et pertinentes au sujet des modèles d'IA est également importante pour les clients des institutions financières. Selon les pratiques exemplaires, et si la loi l'exige, il est nécessaire de fournir une explication au sujet de la prise de la décision au client qui en fait l'objet. Le défi pour les concepteurs de modèles consiste à concevoir ces systèmes de manière à satisfaire aux objectifs d'exactitude et de communication. À l'instar de leur discussion précédente sur les degrés adéquats d'explicabilité, les participants ont convenu que les détails à inclure lors de la communication dépendent du cas d'utilisation (risque et importance relative).



## Explicabilité

Il est également généralement admis que les clients doivent en être informés lorsqu'ils traitent avec l'IA. Le Secrétariat du Conseil du Trésor du Canada, dans sa Directive sur la prise de décisions automatisée, a établi des exigences selon lesquelles les clients doivent être avisés que la décision rendue sera prise en tout ou en partie par un système décisionnel automatisé.

### Propriétés d'une bonne communication

Il est dit qu'une bonne communication doit être concise, simple, pertinente, intuitive et pratique. De plus, elle doit être rédigée dans un langage simple et compréhensible pour les utilisateurs sans expertise en IA, et accompagnée d'exemples et de recommandations, le cas échéant. Enfin, elle doit être présentée de manière logique, organisée et adaptée aux besoins précis et aux connaissances du public ciblé.

### Risques de communication

Comme nous l'avons souligné, il existe de nombreux avantages pour les institutions financières à divulguer des informations liées à l'IA. Cependant, elles doivent faire preuve de prudence lorsque vient le temps de décider des renseignements à communiquer. Les facteurs suivants doivent être pris en compte pour atténuer les conséquences imprévues d'une communication excessive :

- **Cybersécurité** : La communication de renseignements qui semblent inoffensifs peut nuire à la cybersécurité des institutions financières.
- **Processus d'intégrité de l'IA** : La communication ne doit pas compromettre le processus d'IA. Par exemple, les attaques par empoisonnement des IA contradictoires visent à influencer les données utilisées dans la formation ou la rééducation de l'algorithme.

Ces données contaminées sont alors introduites dans l'algorithme et provoquent un mauvais apprentissage machine. L'intégrité de l'IA est particulièrement importante lorsque les données utilisées par les algorithmes sont accessibles au public et susceptibles d'être compromises.

- **Avantage concurrentiel** : Lorsqu'une organisation partage trop d'informations sur ses produits, services, stratégies d'IA ou autres informations exclusives en matière d'IA, elle peut perdre son avantage concurrentiel.

En général, pour maintenir leur avantage concurrentiel, les entreprises doivent trouver un équilibre entre la communication suffisante d'informations pour satisfaire les clients et les investisseurs, et la préservation de la confidentialité des informations de nature délicate.

### Communication des tiers

Les institutions financières ne sont pas en mesure de communiquer adéquatement avec leurs clients lorsqu'elles n'ont pas un accès global aux modèles tiers. En effet, ces fournisseurs traitent leurs produits comme exclusifs afin de protéger leur propriété intellectuelle (PI).

Les participants au forum ont discuté de solutions possibles pour améliorer l'explicabilité et la communication liées aux modèles d'IA de tiers. Une des solutions proposées consistait à intégrer des exigences d'explicabilité dans le processus d'approvisionnement visant des outils de tiers. Certains participants au forum étaient d'avis qu'une telle solution serait difficile à mettre en œuvre, car les fournisseurs tiers ne seraient pas d'accord avec ces conditions, à moins qu'elles ne deviennent la norme et soient imposées par toutes les institutions financières. Une autre



solution proposée pour l'explicabilité des modèles d'IA tiers était la certification de ces modèles tiers par un organisme indépendant. Le succès de la solution proposée ne pouvait cependant pas être déterminé, car il n'existe pas de cadre établi pour la certification de modèles tiers. En outre, compte tenu de la nature « sur mesure » des modèles d'IA et de la différence parmi les cas d'utilisation, il est difficile de fournir des approbations générales à des tiers pour leurs produits. De la création d'un organisme indépendant pourrait découler le risque de monopole qui étoufferait l'innovation ou diluerait la qualité des certifications de tiers. Une troisième solution proposée consistait à communiquer les renseignements relatifs à une application tierce directement à l'organisme de réglementation, au cas par cas, afin de permettre à l'organisme de réglementation de vérifier l'intégrité de cette application.

Au lieu d'une certification obligatoire, les participants au forum ont convenu que des normes cohérentes à l'échelle du secteur pourraient être adoptées pour les modèles fournis par des tiers (une certification de type ISO). L'adoption volontaire de certifications et de normes ne garantirait cependant pas l'approbation par l'organisme de réglementation.

Quelques participants au forum ont déconseillé l'utilisation des modèles d'IA tiers dans la mesure du possible. Ils privilégient plutôt le développement des modèles à l'interne afin d'éviter les problèmes liés à l'explicabilité.

## Explicabilité et confiance

La relation entre l'explicabilité et la confiance (c.-à-d. une croyance ferme dans la fiabilité du modèle d'IA) n'est pas aussi évidente qu'elle pourrait paraître. Souvent, l'explicabilité est largement perçue comme un outil pour établir la confiance. Comprendre le mécanisme de décision d'un modèle n'est pas gage de confiance absolue en celui-ci. D'autres aspects peuvent affecter la confiance dans un modèle, notamment la précision et l'absence de biais.

Il existe même des cas où l'explicabilité peut nuire à la confiance, par exemple lorsque le client n'est pas d'accord avec l'explication d'une décision donnée. La confiance générale accordée au système d'IA peut être affectée par une explication difficile à comprendre.

L'explicabilité, accompagnée du bon niveau de communication (au bon public), est l'un des nombreux facteurs qui contribuent à établir la confiance entre une institution financière et ses clients. Inévitablement, une confiance accrue dans l'IA permet d'en étendre les limites d'utilisation et de continuer d'innover.



# Données

Les données sont une ressource essentielle au développement et à la mise en œuvre de la technologie de l'IA. À mesure que de nouvelles données deviennent disponibles, les applications d'IA peuvent s'améliorer et prendre de l'ampleur. Les institutions financières ont été en mesure d'utiliser les données à leur avantage, notamment pour adapter le service à la clientèle, pour améliorer la détection des fraudes et pour augmenter leur efficacité opérationnelle. Toutefois, l'intégration de l'IA dans leurs opérations a aussi mis en évidence les défis liés à la gestion et à l'utilisation des données.

L'un des principaux problèmes est de garantir la qualité des données, puisqu'il implique que les données doivent être exactes, fiables, complètes, représentatives, cohérentes et conformes aux réglementations pertinentes. Un autre problème rencontré porte sur la protection de la vie privée, car les institutions financières doivent prendre des mesures pour protéger les renseignements personnels et financiers de nature délicate. De plus, l'alignement de la stratégie de données avec la stratégie commerciale est devenu plus complexe. Il est de plus en plus important de parvenir à une saine gouvernance des données afin de régler ces problèmes.

**“Everyone talks about models, models, models. We need to focus on proper data analysis first... shift [the] mindset to [data] stewardship.”**

Ima Okonny, Chief Data Officer at Employment and Social Development Canada (ESDC)





Les questions clés suivantes en matière de données ont été abordées dans le cadre du forum :

Quelle incidence peut avoir un ensemble de données variées sur la qualité des données?

Quels sont les défis posés par l'IA en matière de gouvernance des données?

Comment peut-on gérer les risques liés à l'IA qui découlent de l'exposition à des tiers?

Quelles sont les difficultés que présente l'alignement des données et des stratégies commerciales?

L'IA peut être entraînée sur une variété de types de données, y compris :



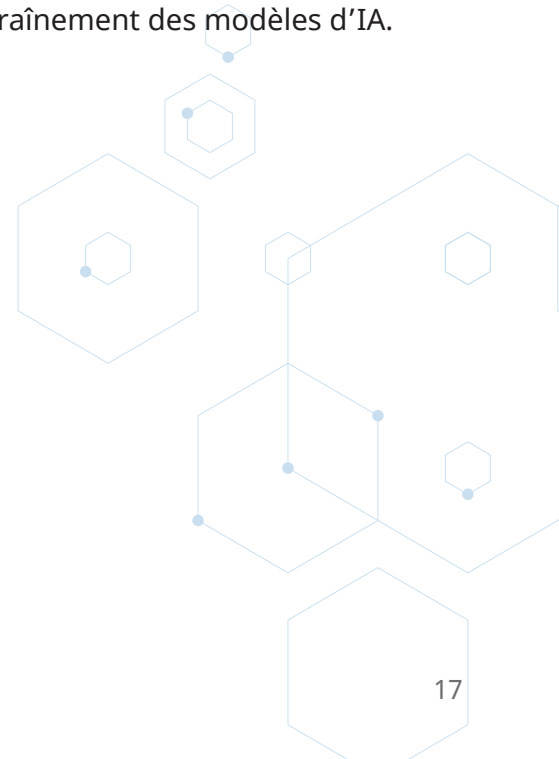
**Données structurées** : données ayant une structure bien précise. Parmi les données structurées, on retrouve notamment les transactions financières, les informations sur le client et les données d'inventaire.



**Données non structurées** : données qui ne sont pas organisées selon un modèle ou une structure prédéfinie, notamment du texte en format libre, des images et des vidéos. Parmi les données non structurées, on retrouve notamment les publications sur les réseaux sociaux, les courriels, les avis des clients et les rapports du conseil d'administration.



**Données synthétiques** : données générées artificiellement. Ce type de données peut être utilisé pour compléter ou remplacer les données empiriques utilisées pour l'entraînement des modèles d'IA.





## Caractéristiques des données et défis connexes

Les données utilisées pour l'entraînement et le développement de l'IA présentent un certain nombre de caractéristiques qui, lorsqu'elles sont exploitées par l'IA, offrent un large éventail de possibilités. Cependant, ces caractéristiques peuvent complexifier l'intégration et la normalisation des données ainsi que la gestion des risques connexes pour les institutions financières. Ces caractéristiques sont expliquées ci-dessous :

- **Volume et variété des données** : Les modèles d'IA peuvent traiter de grandes quantités de données et peuvent gérer des données structurées et non structurées, ce qui augmente les défis liés au maintien de la qualité et de l'intégrité de ces données.
- **Gestion des versions des données** : Les modèles d'IA peuvent être itératifs et entraînés à partir de plusieurs versions de données. Cette fonctionnalité peut complexifier le suivi de la version des données utilisées pour entraîner un modèle particulier ainsi que le suivi de l'évolution du modèle au fil du temps.
- **Agilité des données** : En plus de la vitesse à laquelle les données sont générées et exploitées par les modèles d'IA, certains systèmes en temps réel réagissent aux flux de données en direct. Une gouvernance solide est nécessaire dans ces cas d'utilisation et doit comporter une capacité d'intervention humaine de type « human-in-the-loop » et des modèles hors ligne qui peuvent être activés si le système en temps réel répond de manière inattendue aux flux de données en direct.

### Qualité des données

La qualité des données est essentielle dans les applications d'IA. La tâche de maintien d'une qualité élevée de données s'est complexifiée, comme nous l'avons mentionné précédemment. En outre, il existe d'autres problèmes qui accentuent la difficulté de déterminer la qualité des données, notamment :

- **Incohérence** : Les données peuvent être très incohérentes; elles peuvent avoir des formats, une structure et des niveaux de détail variables. Ces incohérences peuvent complexifier le repérage de tendances dans les données.
- **Bruit** : Les fautes de frappe, les erreurs grammaticales et les informations non pertinentes souvent trouvées dans les données peuvent rendre l'extraction d'informations utiles plus difficile.
- **Manque de contexte** : Sans contexte approprié, il peut devenir difficile de comprendre la signification de certaines données. Par exemple, lorsqu'un client inscrit dans un formulaire de commentaire « *Je ne suis pas satisfait du service* », il ne précise pas clairement le service duquel il se plaint.
- **Qualité des sources** : Les données peuvent provenir de diverses sources telles que les médias sociaux, les commentaires des clients et les articles de presse, dont la qualité et la fiabilité sont variables.
- **Double sens** : Certains types de données peuvent avoir une interprétation vague ou être



compris différemment selon l'utilisation qui en est faite. Par exemple, la valeur notionnelle de transactions financières ou de produits dérivés.

Pour surmonter ces problèmes, les experts et les ingénieurs en science des données peuvent faire l'exploration, le nettoyage, la validation et l'intégration des données. Cependant, il convient de noter que même avec ces approches, il peut encore être difficile d'assurer la qualité des données.

Les participants au forum ont également fait remarquer que même si les données synthétiques peuvent être utilisées pour résoudre certains problèmes de qualité des données (p. ex. partialité, équité, déséquilibre), la recherche dans ce domaine en est toujours à ses balbutiements. De plus, il est important de souligner que les données synthétiques peuvent avoir leurs propres limites, telles que l'impossibilité de reproduire pleinement la complexité et la variabilité du monde réel, ce qui pourrait conduire à des modèles d'IA qui ne sont pas suffisamment performants.

### Données synthétiques

Comme l'a souligné Stuart Davis, vice-président à la direction, Gestion du risque de crimes financiers, Banque Scotia, les données synthétiques constituent un domaine d'émergence important. En raison de la protection de la vie privée et d'autres restrictions, les institutions financières ne peuvent pas toujours fournir de données réelles à leurs partenaires. Les données synthétiques représentent une bonne solution de remplacement.<sup>[3]</sup>

### Agrégation de données

Lors de la collecte et de la combinaison de données de formats différents qui proviennent de sources multiples, il est important de reconnaître les incohérences et les erreurs qui peuvent survenir. D'ailleurs, une étape importante dans de nombreuses applications d'apprentissage supervisé est l'annotation/étiquetage des données. Cependant, si plusieurs personnes ou fournisseurs de données contribuent à l'étiquetage des données, il est possible qu'ils utilisent des normes ou définitions différentes, ce qui pourrait entraîner des incohérences dans les données et poser un défi au moment d'agréger de tels ensembles de données. La performance du système d'IA pourrait même être amoindrie, si des étiquettes incohérentes ou incorrectes nuisent à son apprentissage. Pour éviter ces problèmes, il est recommandé d'avoir des directives et des protocoles clairs en matière d'étiquetage, de traçabilité et de gestion des données.





## Gouvernance des données

Une bonne gouvernance des données peut contribuer à garantir l'exactitude, la cohérence, la sûreté et l'exhaustivité des données, toutes des caractéristiques essentielles au bon fonctionnement des systèmes d'IA. La gouvernance des données est également essentielle pour les institutions financières compte tenu de la nature délicate et confidentielle des données financières et des données clients. Les participants au forum ont exploré certains aspects associés à la gouvernance des données.

### Propriété des données

La propriété des données peut présenter certains défis pour les organisations puisque les ensembles de données nécessaires à la construction d'un modèle d'IA peuvent provenir de différents secteurs d'activité. Le fait de devoir obtenir l'autorisation d'utiliser les données des différents propriétaires, bien que nécessaire, pourrait ralentir le processus de développement de l'IA. Même si chaque secteur d'activité est responsable de ses propres données, c'est l'équipe qui construit le modèle qui doit s'assurer que les données utilisées pour construire le modèle sont correctes et complètes.

### Cloisonnement des données

Au cours de sa présentation, Andrew Moore, directeur de Google Cloud AI, a discuté du travail fait par Google en lien avec son outil de lutte contre le blanchiment d'argent, tâche qui s'est avérée difficile à accomplir en raison de la disponibilité des données. Il a mentionné que, dans le secteur des services financiers, les données sont généralement cloisonnées au sein d'une organisation et, lors de la construction de l'outil, Google a dû intégrer des milliers de bases de données. Ainsi, en raison des problèmes liés à l'obtention de données, le déploiement d'outils d'IA à même les applications technologiques pourrait être beaucoup plus difficile en finance.<sup>[4]</sup>

### Confidentialité et sécurité des données

Les organisations recueillent de grandes quantités de données pour que leurs modèles d'IA puissent exploiter et générer des informations, mais certaines données peuvent être personnelles ou de nature délicate. Lorsqu'il est question de données personnelles ou délicates, la confidentialité et la sécurité de ces données sont essentielles. Malheureusement, les risques de fuites de données ou d'accès non autorisés aux données continuent d'augmenter. Une saine gouvernance des données atténue ces risques.





### Limites régionales des données

L'extensibilité des solutions basées sur l'IA dans différentes zones géographiques revêt une importance particulière. L'un des défis rencontrés concerne l'utilisation de données protégées puisqu'il n'y a pas de consensus établi entre les compétences sur ce qui est considéré comme protégé. Les enjeux liés aux données protégées pourraient alourdir la gouvernance des données (par exemple, le suivi, la gestion, l'octroi d'autorisations) et ralentir le processus de développement de modèles.

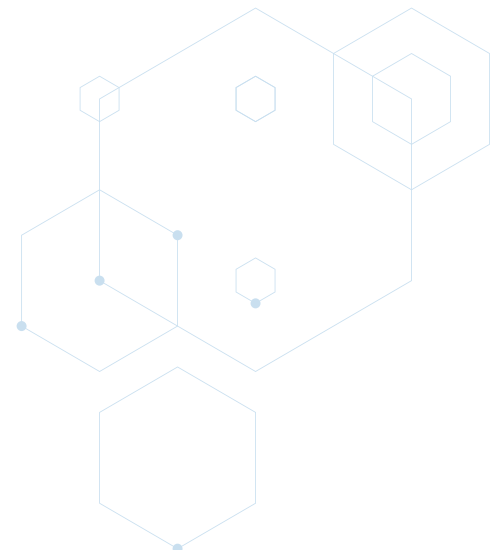
Autre défi de taille pour les institutions financières qui font des affaires dans différentes régions : l'interaction entre les systèmes ou processus dans ces différentes compétences, y compris l'infrastructure existante et même la fréquence de la collecte des données.

### Approche centrée sur les données

Pour améliorer les performances des applications d'IA et des modèles traditionnels, on peut bonifier continuellement les données utilisées pour entraîner ces modèles. Cette approche est connue sous le nom d'approche centrée sur les données. Au lieu de mettre l'accent sur l'itération et le recyclage des algorithmes pour rehausser les performances, l'intégration d'une approche centrée sur les données optimise le potentiel du modèle. Une saine gouvernance des données est nécessaire à l'adoption d'une approche de développement de modèles d'IA centrée sur les données.

### Littératie des données

L'instauration d'une solide culture de littératie des données a également été jugée essentielle pour les organisations qui déploient et utilisent activement l'IA. En effet, pour soutenir l'adoption généralisée de l'IA, il est nécessaire de sensibiliser l'ensemble de l'organisation aux divers risques qui découlent d'une utilisation inadéquate des données. Par conséquent, les organisations devraient envisager des activités de formation continue pour leur effectif sur un éventail d'aspects liés aux données.





## Données de tiers

La valeur de l'IA dépend de la qualité et de la quantité de données utilisées. Les organisations ont besoin de données au-delà de celles générées par leurs activités commerciales. Elles doivent donc compter sur des tiers pour les leur fournir. Les participants au forum ont exploré les aspects liés à l'utilisation de données de tiers.

### Collecte de données

Dans les situations où les données utilisées par l'institution financière sont recueillies par un tiers, la responsabilité que porte l'institution financière en cas d'atteinte à la protection des données du tiers pourrait être imprécise. Dans de tels cas, même si le tiers est responsable de la fuite de données, l'institution financière pourrait quand même s'exposer à un risque d'atteinte à la réputation.

Compte tenu de la grande variété de types de données qui peuvent être utilisés dans les modèles d'IA, il est important d'accorder une attention particulière au processus d'agrégation des données provenant de différents tiers afin d'éviter les incohérences et les erreurs dans les données agrégées.

Les organisations pourraient également éprouver de la difficulté à vérifier que les données ont été recueillies correctement par le tiers, de sorte qu'il pourrait être nécessaire de mettre en place des exigences pour assurer une collecte de données appropriée. Par exemple, il pourrait être nécessaire de confirmer que le fournisseur de données a obtenu le consentement des personnes concernées avant de fournir les données à d'autres tiers ou à l'institution financière.

Les participants au forum ont généralement convenu que les dispositions réglementaires doivent préciser les types de données de tiers qui peuvent être utilisés par les institutions financières ainsi que le niveau de diligence raisonnable requis dans chaque cas. Lorsque les données sont obtenues auprès de tiers, certaines protections sont essentielles pour assurer la gouvernance de la source, de la provenance et de la qualité des données. Certains participants ont même suggéré qu'il pourrait être envisageable d'adopter une approche de type « étiquette nutritionnelle » pour les données. Selon cette approche, des renseignements clés sont divulgués au sujet d'ensembles de données, comme le moment où elles ont été recueillies, la façon dont elles ont été recueillies, l'utilisation prévue au moment de la collecte, etc.<sup>[5]</sup>

### Échanges de données

L'IA ne peut être considérée indépendamment du mouvement de numérisation croissant de notre société. L'Internet des objets (IdO) et le système bancaire ouvert sont les deux avancées majeures qui ont fait l'objet de discussions approfondies lors du forum. Il a été noté, d'une part, que l'IdO et le système bancaire ouvert pourraient exacerber les problèmes de confidentialité des données, car un plus grand nombre d'acteurs du marché pourraient accéder aux données. À cet égard, il pourrait être plus difficile de faire porter la responsabilité dans un système bancaire ouvert où les données d'une partie pourraient être utilisées par une autre partie. D'autre part, l'échange de données par l'entremise du système bancaire ouvert pourrait contribuer à la détection des biais et à l'amélioration de la performance



globale des modèles d'IA puisqu'il y aurait de plus grands ensembles de données pour entraîner les modèles d'IA.

## Alignement de la stratégie des données et de la stratégie commerciale

L'alignement des stratégies de données avec les stratégies commerciales est essentiel pour permettre aux organisations d'exploiter efficacement leurs actifs de données de manière à améliorer les résultats de l'entreprise. La stratégie de données encadre l'approche de l'organisation en matière de gestion et d'exploitation des données, tandis que la stratégie commerciale encadre ses objectifs globaux. Lorsque les stratégies de données sont en accord avec les stratégies commerciales, les organisations peuvent veiller à recueillir, gérer et analyser les bonnes données pour soutenir leurs objectifs commerciaux.

En outre, la stratégie commerciale globale peut devoir être adaptée aux capacités de l'IA à mesure que l'IA développe de nouvelles capacités et de nouvelles occasions, notamment l'automatisation des processus, la découverte de nouvelles perspectives et la création de nouveaux produits et services.

Les participants ont admis que l'IA comportait de nombreux avantages. Ils ont également reconnu que les institutions financières pourraient avoir

du mal à harmoniser leur stratégie de données avec les exigences spécifiques de l'IA et avec la stratégie commerciale globale de l'organisation. Comme les participants l'ont indiqué, il n'est pas toujours évident de quantifier les avantages d'investir dans la stratégie des données, en particulier lorsqu'ils sont comparés à ceux d'autres projets commerciaux immédiatement rentables, ce qui pourrait entraîner la sous-priorisation des projets stratégiques liés aux données.





# Gouvernance

L'adoption d'une approche proactive en matière de gouvernance par les institutions financières démontre leur engagement envers l'utilisation responsable de l'IA et l'établissement d'un climat de confiance avec leurs clients et les intervenants. Un cadre de gouvernance solide sert de fondation à la culture de responsabilisation à l'égard de l'utilisation de l'IA de l'organisation. Ce cadre permet aux institutions financières de tirer parti de l'IA tout en évitant de causer du tort aux clients et à la population en général.

Les questions clés suivantes en matière de gouvernance ont été abordées lors du forum :

The forum addressed the following key questions on governance:

Qu'est-ce qui constitue une bonne gouvernance des modèles d'IA?

Les approches actuelles de gestion du risque de modélisation sont-elles suffisantes?

Quels sont les défis auxquels les institutions sont confrontées dans la mise en œuvre des cadres de gouvernance de l'IA?

Quels outils ou pratiques exemplaires peut-on utiliser pour atténuer les risques et les défis liés à la gouvernance?

**“L'utilisation de l'IA dans les institutions financières évolue rapidement. L'IA est un catalyseur commercial dont les avantages peuvent aller de l'accélération de la productivité à la stimulation de la croissance, et peuvent être liés à la performance de l'entreprise. Mais quel est le juste équilibre à atteindre entre le risque et l'innovation d'un point de vue commercial et réglementaire?”**

Donna Bales, directrice principale de la recherche, Bureau du directeur des systèmes d'information, Info-Tech Research Group, et fondatrice de l'Association canadienne de la technologie réglementaire (ACTR)





## Cadre de gouvernance de l'IA

La gouvernance englobe certes la surveillance, mais elle constitue un concept plus vaste. La gouvernance désigne les structures, systèmes et pratiques mis en place par une organisation pour la prise de décision, la reddition de comptes, le contrôle, le suivi et l'atténuation des risques, ainsi que les rapports de performance..<sup>[6]</sup>

Les institutions financières mettent déjà en œuvre des pratiques de gouvernance à différents degrés de sophistication. Les participants au forum ont soulevé la nécessité de modifier les cadres de gouvernance pour appuyer un contrôle adéquat des techniques d'IA au vu de l'utilisation croissante de celles-ci.

La gestion du risque de modélisation a été mise en évidence à la suite de la crise financière mondiale de 2008, avec l'introduction de lignes directrices réglementaires, notamment la directive SR11-7 de la Réserve fédérale américaine (2011)<sup>[7]</sup> et la directive E-23 du BSIF (2017)<sup>[8]</sup>. Ces lignes directrices rappellent l'importance de la gestion du risque de modélisation en tant qu'élément d'une bonne gouvernance et la placent au centre d'un cadre de gestion du risque d'entreprise plus vaste.

Les modèles d'IA présentent bon nombre des mêmes risques que les modèles traditionnels et existent souvent au sein d'un écosystème qui interagit avec d'autres fonctions de risque et de gouvernance établies.

Les participants ont souligné les caractéristiques d'une bonne gouvernance de l'IA dans les institutions financières :

- **Elle doit être intégrée et s'étendre à l'ensemble de l'organisation.** Il est important que tous les intervenants internes comprennent les principes fondamentaux de l'IA afin d'être en mesure de contribuer à la gestion des risques qui en découlent. La haute direction (ce qui inclut les membres du conseil d'administration) doit comprendre les avantages, les risques et les limites de l'utilisation des modèles d'IA pour assurer une prise de décision judicieuse. Des processus et des cadres qui s'étendent à l'ensemble de l'organisation sont essentiels à la bonne gouvernance.
- **Les rôles et les responsabilités doivent être clairs et bien articulés.** Il est nécessaire de définir la responsabilité à l'égard des données, des modèles d'IA et des résultats, ainsi que la structure des approbations relatives aux différents risques. Une formulation claire des mandats de ces différents groupes permet d'éviter les lacunes dans la gestion des risques.
- **La tolérance au risque doit être bien définie.** Les institutions financières doivent définir ou mettre à jour leur tolérance au risque en tenant compte des risques accrus qui découlent de l'utilisation de l'IA, y compris les cyberrisques et les risques liés aux tiers.
- **Elle doit cadrer avec le risque lié aux cas d'utilisation.** Une transition culturelle au sein des institutions financières est requise pour passer d'une approche fondée sur des règles à une approche fondée sur les risques. Cette transition permettra la prise de risques conformément à l'énoncé de tolérance au

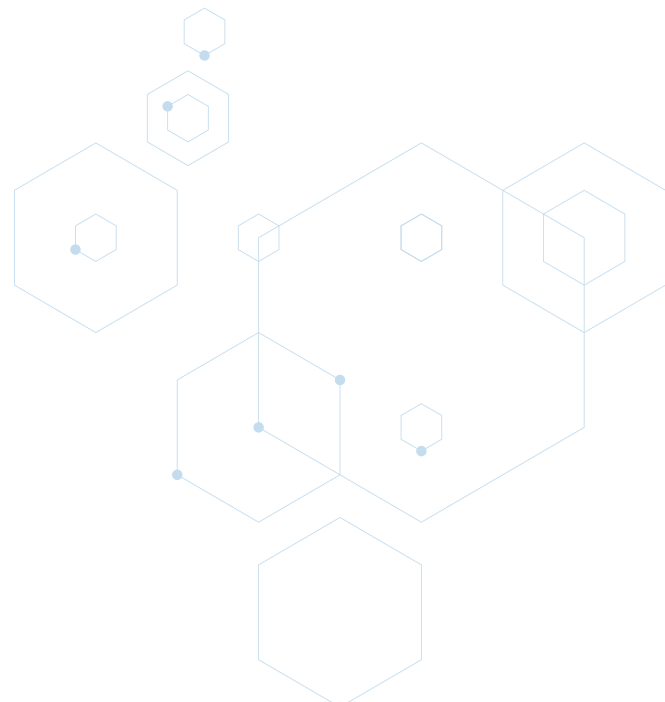


risque de l'entreprise. Le niveau de diligence raisonnable appliqué au cas d'utilisation de l'IA doit être proportionnel au risque.

- **Elle doit s'adapter à mesure qu'une institution financière progresse dans son adoption de l'IA.** Les institutions financières qui adaptent avec succès leurs cadres de gouvernance existants aux modèles d'IA bénéficieront d'une flexibilité et d'une agilité, et seront en mesure de se repositionner, au besoin, à mesure qu'apparaissent de nouvelles techniques et de nouveaux risques.
- **Elle doit s'adapter à mesure qu'une institution financière progresse dans son adoption de l'IA.** Financial institutions that successfully adapt existing governance frameworks for AI models will embrace flexibility and agility, pivoting where needed as new techniques and risks emerge.

### Outil d'évaluation de l'incidence algorithmique du Secrétariat du Conseil du Trésor (SCT)

Benoit Deshaies, directeur, Données et intelligence artificielle, SCT, présente un exemple d'outil de bonne gouvernance. L'outil d'évaluation de l'incidence algorithmique du SCT est un outil d'évaluation obligatoire qui doit être utilisé par tous les ministères du gouvernement du Canada pour évaluer les décisions automatisées sur un éventail de sujets. Il a été élaboré de façon collaborative, a une portée et une application bien définies, est fondé sur les risques et met en œuvre le principe de proportionnalité, car il établit un résultat en fonction de l'incidence sur le client. Les pratiques d'atténuation des risques sont ensuite traduites en exigences spécifiques de gouvernance en fonction du niveau d'incidence. Il s'agit d'un outil d'autoévaluation appuyé par un processus d'examen par les pairs pour les systèmes automatisés qui peuvent l'exiger.<sup>[9]</sup>





## Évolution des cadres de gouvernance existants pour l'IA

Le Forum public-privé sur l'IA qu'a tenu la Banque d'Angleterre a permis à l'institution de constater que les structures et cadres de gouvernance existants constituent un bon point de départ pour les modèles et les systèmes d'IA. La Banque d'Angleterre a d'ailleurs conclu que la gouvernance des modèles devait être adaptée en fonction du risque et de l'importance relative du cas d'utilisation, mais aussi accorder une attention particulière aux vulnérabilités de gouvernance exacerbées par les modèles d'IA.

Les participants au FIASSF étaient généralement d'accord avec la conclusion de la Banque d'Angleterre selon laquelle il était préférable d'étendre les cadres de gouvernance existants que d'élaborer une série de nouveaux processus et de nouvelles procédures propres à l'IA. Les institutions financières ne sont pas toutes rendues au même point en matière d'adoption de l'IA et de mise en œuvre des cadres de gouvernance. L'élaboration d'un cadre de gouvernance robuste qui tient compte de l'IA pourrait entraîner un changement de culture important, car de plus en plus de secteurs d'activités des institutions financières ont recours aux techniques d'IA.

Selon le rapport de synthèse sur la gouvernance de l'apprentissage automatique du Institute of International Finance (IIF) (2020), il existe des différences d'approches d'une zone géographique à l'autre. Les institutions financières au Canada mettent l'accent sur l'amélioration des cadres existants.<sup>[10]</sup>

Les participants au FIASSF se sont penchés sur les principaux défis liés à l'application et

à l'adaptation des cadres de gouvernance existants. Comme c'est le cas pour tout cadre de gouvernance, l'un des principaux risques est que la conformité devienne un exercice de « case à cocher », comme l'explique l'article « Model Risk Management Lessons Learned: Tracing Issues from the Pandemic to the Great Recession ».<sup>[11]</sup> Lorsque la gouvernance est exercée de manière machinale, la compréhension des risques est reléguée au second plan, pour faire place à l'accomplissement des divers éléments du cadre prescrit, peu importe le risque associé à ces éléments ou leur pertinence.

### Notation du risque de modélisation

Un cadre de gouvernance décrit normalement la notation du risque de modélisation pour tenir compte des nuances de risque associé à l'utilisation des modèles. Les institutions ont élaboré différentes approches pour noter le risque de modélisation. En général, ces approches tiennent compte de l'importance relative, de l'incidence financière, de la complexité de la méthodologie et de la complexité de l'infrastructure, entre autres. Au forum, on a également mentionné l'importance de l'incidence sur le client en raison du risque d'atteinte à la réputation. Il faut donc considérer non seulement l'importance financière d'un modèle d'IA, mais aussi l'effet possible sur les clients. Cette approche adopte une vision globale des risques associés aux modèles d'IA et en examine à la fois les aspects financiers et non financiers. Elle sera d'ailleurs pertinente pour respecter la *Loi sur l'intelligence artificielle et les données*, présentée



pour adoption dans le cadre du projet de loi C 27, en cours.

### Répertoire des modèles

Généralement, les institutions dotées de cadres de gouvernance éprouvés tiennent un répertoire de modèles qui peut être élargi pour intégrer tous les modèles d'IA. Une approche sophistiquée repère les interrelations entre les modèles d'IA, signale les résultats d'un modèle qui servent à en nourrir un autre et indique les modèles utilisés pour expliquer les résultats d'autres modèles. La notation du risque de modélisation et le répertoire des modèles sont des outils qui peuvent faciliter le suivi du travail de gouvernance requis pour chaque modèle selon une approche axée sur les risques.

### Approches de la gouvernance de l'IA

La composition exacte d'une structure de gouvernance de l'IA variera d'une institution financière à l'autre en fonction de facteurs tels que la taille de l'institution et la sophistication de son cadre de gouvernance d'entreprise. L'enquête de l'IIF mentionnée ci-dessus montre que 35 % des institutions financières mondiales ont mis en place une équipe centralisée de spécialistes, telle qu'un conseil de gouvernance de l'IA, qui travaille à harmoniser les objectifs organisationnels, à évaluer la valeur et la priorisation des cas d'utilisation de l'IA et à approuver les modèles d'IA avant le déploiement. Pour les petites institutions ayant des cas d'utilisation à faible risque, la responsabilité peut demeurer au sein de l'unité opérationnelle qui utilise l'outil d'IA. Il existe une troisième approche, hybride, entre une structure de gouvernance centralisée et fédérée, où l'équipe de coordination rassemble différents

intervenants pertinents pour développer et gouverner le cas d'utilisation de l'IA.

### Approche multidisciplinaire

La gouvernance des modèles d'IA peut être plus complexe que la gouvernance des modèles traditionnels, car elle nécessite une approche multidisciplinaire pour être efficace. Typiquement, la gouvernance des modèles est soumise à un comité d'examen des modèles composé de cadres supérieurs représentant les équipes de conception, de validation et d'utilisateurs du modèle. En raison de la portée, de l'ampleur et de la complexité accrues de certains cas d'utilisation de l'IA, la capacité de ce groupe d'intervenants a fait l'objet de discussions. En effet, certains participants étaient d'avis que la responsabilité à l'égard des questions éthiques devrait être explicitement définie dans le processus de gouvernance. L'utilisation de modèles d'IA peut également introduire des considérations juridiques qui devraient faire partie du processus de gouvernance. Les équipes juridiques et de conformité pourraient être des intervenants clés dans le développement de modèles d'IA.



La figure 1 présente divers groupes d'intervenants qui pourraient être intégrés dans le cadre d'un processus de gouvernance du modèle.

### Figure 1

Équipe multidisciplinaire de gouvernance de l'IA





### Responsabilité

En ce qui concerne la responsabilisation, les participants au forum ont discuté de la nécessité d'intégrer la gouvernance des données et des modèles. Les participants ont reconnu le caractère distinct de la responsabilité des données de celle des modèles, car les compétences, les outils et les processus nécessaires à la gouvernance des modèles diffèrent de celle des données. Cependant, les participants ont convenu qu'une gouvernance efficace des données est fondamentale et alimente directement la gouvernance du modèle, d'où la nécessité d'établir un lien étroit entre elles. Les propriétaires et les concepteurs de modèles

doivent s'assurer que les données utilisées pour le modèle sont adéquates et qu'ils disposent de contrôles appropriés pour prévenir les risques tels que les biais, les résultats injustes, le surajustement ou le sous-ajustement et le manque de représentativité. Pour ce faire, il faut une bonne gouvernance des données et une communication appropriée entre les équipes d'intervenants.

### Compétences, culture et autres défis

Les participants au forum ont discuté des défis liés à la mise en œuvre d'un cadre de gouvernance solide pour les modèles d'IA. Ces défis ainsi que leurs répercussions sont présentés ci-dessous. Dans l'ensemble, l'intégration efficace des outils d'IA dans le secteur des services financiers nécessite une combinaison d'expertise technique, de changement organisationnel et de changement culturel.

#### Concurrence pour les ressources

La compétitivité du marché du travail a eu une incidence sur la capacité d'attirer et de retenir l'expertise en IA au sein du secteur financier. Cette situation est davantage accentuée par la disponibilité limitée de personnes possédant à la fois l'expertise en matière d'IA et la connaissance du domaine.

#### Indépendance des fonctions de développement, de validation et d'audit

Lorsque les institutions financières attirent les talents appropriés, ces personnes sont souvent embauchées pour occuper des postes de première ligne (développement de modèles). Il y a donc un écart de compétences en IA entre les fonctions de développement et de validation ou





d'audit. Cette situation, quoique normale dans des cas d'émergence de nouvelles technologies, peut lamener l'efficacité du processus de gouvernance.

Un autre aspect de l'indépendance qui présente un défi est l'utilisation d'une approche « agile » de développement de modèles. Les participants au forum étaient divisés quant à l'inclusion d'une équipe de validation du modèle dès le début du processus d'élaboration. L'un des avantages cités était la détection précoce des problèmes, tandis que l'affaiblissement de la gouvernance représentait un défi.

### Réduction de la surveillance humaine

L'éventuelle prise de décision autonome par les systèmes d'IA pourrait entraîner une réduction de la surveillance humaine sur les aspects clés de la gouvernance. Il est nécessaire d'apporter des changements au cadre de contrôle, par exemple une surveillance plus fréquente, pour atténuer certains des risques liés à la prise de décision automatisée.

### Solutions de tiers

Les systèmes d'IA peuvent s'appuyer sur des modèles de fournisseurs, des outils ou des données de tiers pour lesquels le manque de transparence peut créer des problèmes de gouvernance.

### Source ouverte

Le risque découlant des données et des outils de source ouverte est différent du risque lié aux tiers, car ces données ne peuvent pas être associées à un accord contractuel. Certains participants estimaient que le code source ouvert offrait un avantage par rapport au code source vendu par un fournisseur tiers. En effet, les partisans du code source ouvert ont souligné sa transparence totale et son examen par de

nombreux utilisateurs. Cependant, les détracteurs ont soulevé des préoccupations concernant la responsabilisation, puisqu'avec ce type de code, il incombe à l'institution de l'examiner et d'en assurer la solidité plutôt que de remettre cette responsabilité à un tiers.

## Prochaines étapes

Les participants ont porté leur regard vers l'avenir de la gouvernance de l'IA dans le secteur des services financiers et ont proposé les domaines clés suivants :

### Gouvernance des outils et des technologies

Une gouvernance robuste de l'IA doit tenir compte de l'écosystème des outils et des technologies dans lequel le modèle est développé, déployé et surveillé. Elle comprend également les outils conçus explicitement pour soutenir la gouvernance des modèles d'IA, notamment les opérations d'apprentissage automatique (MLOps)<sup>[12]</sup>, qui peuvent contribuer à automatiser certains aspects du cycle de vie du modèle tels que le développement et la surveillance. Lorsqu'elles sont mises en œuvre efficacement, les opérations d'apprentissage automatique permettent une normalisation et une cohérence du processus puisqu'elles intègrent les données existantes et les cadres de gouvernance des modèles dans le processus. Elles devraient donc être considérées comme un outil de soutien à la gouvernance et non comme un substitut à celle-ci. Lorsque les outils d'IA sont utilisés en soutien direct à la gouvernance, il faut maintenir une interaction humaine dans le processus pour repérer les angles morts et les lacunes dans la gouvernance.



### Gouvernance par des tiers

Les participants au forum se sont entendus sur le fait que les cadres existants de gestion des risques liés aux tiers pourraient constituer un bon point de départ pour évaluer les risques propres à l'IA découlant de l'exposition à des tiers. Les attentes en matière de gouvernance pour les solutions développées à l'interne et les solutions fournies par les fournisseurs devraient être semblables. Pour résoudre le problème de la « propriété intellectuelle » des tiers, les participants au forum ont exploré un certain nombre d'options.

En raison du penchant du secteur pour les technologies ouvertes, les participants au forum ont discuté de la gouvernance nécessaire pour les logiciels et bibliothèques ouvertes utilisés pour développer et mettre en œuvre des modèles d'IA. Ils ont d'ailleurs convenu qu'une gouvernance du code source ouvert proportionnelle au risque du cas d'utilisation particulier était nécessaire. Dans les cas d'utilisation à faible risque et lorsque la tolérance au risque de l'entreprise le permet, les institutions financières peuvent utiliser du code source ouvert sans d'abord procéder à son examen rigoureux. Dans les cas d'utilisation à risque plus élevé, les institutions financières devraient procéder à un examen indépendant du code source ouvert.

### Aspects organisationnels

Les institutions financières peuvent avoir besoin de créer de nouveaux rôles ou de réorganiser les équipes existantes pour tirer efficacement parti des compétences et de l'expertise des experts en science des données et d'autres professionnels de la technologie. En outre, les organisations peuvent avoir besoin d'offrir des occasions de formation et de perfectionnement pour aider les employés

opérationnels et des technologies à s'adapter aux nouvelles technologies et aux nouveaux flux de travail.

### Compétences et éducation

Certaines institutions financières s'attaquent de façon proactive à la lacune en matière de compétences au moyen de formations internes. De façon générale, les programmes de formation par affectations successives entre les lignes de défense sont encouragés pour faciliter la circulation des talents et l'éducation dans tous les domaines de l'organisation. Certaines institutions financières mettent en œuvre des programmes de formation propres à l'IA qui incluent des notions de droit (notamment la protection de la vie privée, la gouvernance, les droits de la personne), de cybersécurité et du domaine. Les institutions qui mettent l'accent sur la formation sont mieux placées pour comprendre et gérer les risques qui découlent des applications basées sur l'IA.

### Collaboration

Les participants ont suggéré diverses formes de collaboration continue à l'échelle du secteur afin de créer une vaste collectivité de pratique et ainsi permettre aux institutions de parler de leurs pratiques exemplaires. Le dialogue et la collaboration continus entre les différents intervenants, comme le milieu universitaire, le secteur et les organismes de réglementation, pourraient contribuer à faire progresser l'innovation grâce à la mise en commun des connaissances.





# Éthique

L'éthique en affaires représente les principes moraux et les valeurs qui régissent la façon de prendre des décisions d'une organisation. Les principes éthiques comprennent des aspects tels que le droit à un recours, l'équité et la vie privée, et sont intégrés au code de conduite et aux valeurs des organisations. Comme le secteur se concentre de plus en plus sur des concepts comme l'investissement responsable, il devient d'autant plus nécessaire de démontrer un engagement à l'égard de la prise de décisions éthiques.

Les participants au forum ont eu l'occasion de discuter des multiples nuances du concept d'éthique. Premièrement, la relativité des principes et des valeurs éthiques a une incidence sur la façon dont les organisations de compétences différentes perçoivent l'éthique. De plus, la mise en œuvre des normes éthiques peut varier en fonction de ses applications précises au sein d'une organisation. Par exemple, l'utilisation de l'historique de voyage pour détecter la fraude pourrait être considérée comme appropriée en raison de son potentiel de détection des activités suspectes ou frauduleuses en lien avec ces données, mais son utilisation pour évaluer

la solvabilité peut être considérée comme contraire à l'éthique, car cette donnée n'est pas directement liée à la solvabilité d'une personne et peut potentiellement discriminer certaines personnes en fonction de leurs antécédents de voyage. Enfin, les normes éthiques pourraient changer au fil du temps. À titre d'exemple, l'accent est davantage mis sur les facteurs environnementaux, sociaux et de gouvernance (ESG) aujourd'hui qu'il y a quelques décennies.

## Intersection de l'IA et de l'éthique

David Leslie, directeur de la recherche sur l'éthique et l'innovation responsable à l'Institut Alan Turing, a défini l'éthique de l'IA comme un ensemble de valeurs, de principes et de techniques qui s'appuient sur des normes largement acceptées de bien et de mal pour guider la conduite morale dans le développement et l'utilisation des technologies de l'IA.<sup>[13]</sup>

Divers aspects de l'éthique en contexte d'IA dans le secteur des services financiers ont alimenté les discussions.

**“Nous ne voulons pas d'une technologie qui va à l'encontre de nos valeurs telles que le droit de contester une décision. Les problèmes surgissent lorsque nous sommes guidés par la technologie plutôt que par nos valeurs.”**

Carole Piovesan, associée directrice, INQ Law



Les questions clés suivantes en matière de données ont fait l'objet de discussions durant le forum :

Comment la relation entre l'éthique et la loi s'applique-t-elle à l'IA?

Quels sont les différents points de vue sur les directives réglementaires en matière d'éthique de l'IA?

Quels sont les défis liés à l'éthique de l'IA et comment ces défis peuvent-ils être surmontés?

Quelle est la définition universelle de l'équité?

## Répercussions juridiques, politiques et réglementaires

On ne peut explorer l'éthique sans tenir compte des normes juridiques. Les notions de droit et d'équité, quoique liées, sont différentes. Les normes juridiques sont établies par le gouvernement, tandis que les normes éthiques sont fondées sur des principes et des valeurs qui peuvent aller au-delà de ce qui est légalement requis. Par conséquent, il est possible pour une organisation de satisfaire à toutes les exigences légales tout en agissant de manière contraire à l'éthique. Les participants au forum ont convenu que les organisations doivent tenir compte des considérations juridiques et éthiques dans leurs processus décisionnels.

Même si l'éthique de l'IA n'est pas contraignante, les principes et valeurs associés ont été et continuent d'être codifiés dans des lois, qui, elles, le sont. À mesure que les normes éthiques sont intégrées dans les lois, les institutions doivent faire face à des situations où des décisions ont été prises avant l'adoption de lois connexes. Par exemple, le Règlement général sur la protection des données (RGPD) de l'Union européenne est entré en vigueur en mai 2018, mais les institutions européennes ont pris des décisions relatives aux données des clients avant 2018. Cette situation met en lumière la nécessité d'affiner et de mettre à jour continuellement les pratiques en fonction de l'évolution des normes et des lois.

La pyramide suivante illustre l'importance de tenir compte des aspects juridiques et éthiques tout en les équilibrant avec les valeurs de l'organisation (voir la figure 2). Au bas de la pyramide, on

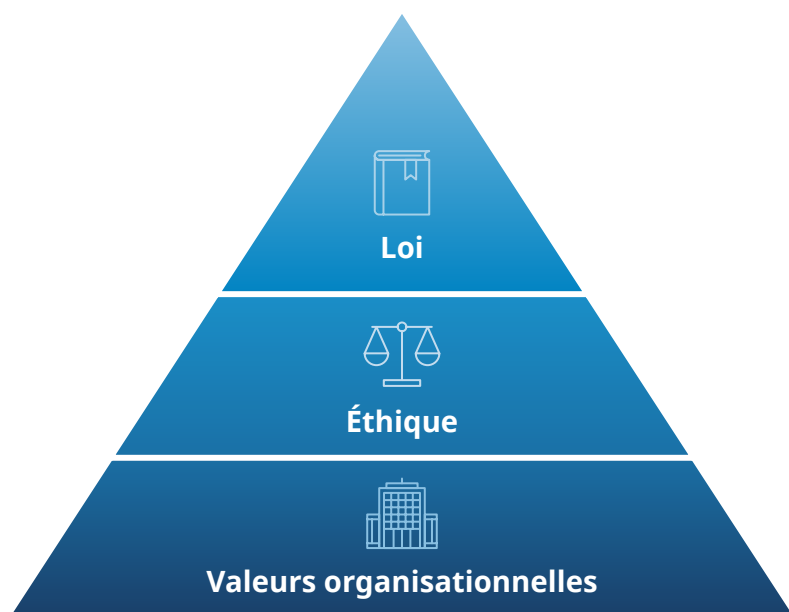




retrouve les valeurs organisationnelles. Les décisions prises doivent soutenir les valeurs importantes pour l'organisation. Au niveau suivant, on retrouve l'éthique; c'est-à-dire les aspects des valeurs canadiennes qui sont généralement acceptés, mais qui ne sont pas explicitement codifiés dans les lois. Le sommet de la pyramide représente la loi, car toutes les institutions qui cherchent à faire des affaires au Canada doivent respecter la législation canadienne.

### Figure 2

Différents niveaux de prise de décision éthique



## Considérations et défis

Les défis que présente l'éthique de l'IA ainsi que sa complexité sont mis en lumière par la subjectivité du concept d'éthique en soi, par le fait que les normes en matière d'éthique sont appelées à changer au fil du temps ainsi que par la codification même des normes éthiques dans les lois et règlements. Pour les institutions financières opérant dans plusieurs endroits, les défis et la complexité de l'éthique sont d'autant plus exacerbés.

Un certain nombre de considérations visant à atténuer ces défis ont fait l'objet de discussions lors du forum :

- **Perspectives multidisciplinaires.** Il est nécessaire, à toutes les étapes du développement et de l'utilisation des applications d'IA, de mobiliser des équipes multidisciplinaires et diversifiées (p. ex. informaticiens, avocats, experts en science des données financières, éthiciens) afin de tirer profit d'une vaste perspective du sujet.
- **Nouveaux rôles.** Les organisations pourraient envisager d'investir davantage dans l'éthique de l'IA en ajoutant de nouveaux rôles tels que celui de dirigeant principal de l'éthique ou de la confiance.



- **Normes.** Les normes sont distinctes des règlements ou des lois; cependant, elles peuvent être facultatives ou obligatoires. Les organismes de normalisation peuvent établir des lignes directrices en matière d'éthique en fonction des conventions établies dans le secteur des services financiers afin de soutenir la gestion des risques liés aux technologies d'IA. Les associations professionnelles peuvent également élaborer des normes éthiques. L'Institut canadien des actuaires (ICA) a d'ailleurs élaboré des normes minimales de déontologie auxquelles ses membres doivent se conformer.
- **Désignation de l'IA.** Pour s'assurer que les questions éthiques appropriées sont prises en compte au cours du processus de conception, d'élaboration et de mise en œuvre des systèmes d'IA, il a été suggéré de créer une désignation qui exige une formation sur les questions éthiques. Par exemple, Singapour a créé un titre d'ingénieur agréé en IA en réponse à ce problème.

## Objectifs commerciaux, données subjectives et injustice

Les participants au forum ont discuté de diverses approches pour faire face aux questions de partialité, en particulier en ce qui concerne les processus historiques et actuels qui désavantagent certains groupes. Les discussions ont également porté sur le rôle des institutions financières dans la création d'un monde plus équitable, plus juste, et dans l'atteinte d'une responsabilité sociale.

Les données utilisées pour l'entraînement et le développement de l'IA peuvent être la source de biais et de résultats injustes. Une approche qu'utilisent les institutions financières pour lutter contre les biais discriminatoires possibles est « l'équité par l'ignorance ». Cette approche fait abstraction de certains attributs personnels lors de la conception du modèle. Le modèle est donc « aveugle » à ces attributs, mais peut tout de même ne pas être neutre sur le plan des résultats.

De plus, dans certains cas, l'exclusion de données protégées ou de nature délicate du processus d'entraînement de l'IA ne mène pas nécessairement à des résultats sans conséquences. Certaines données retenues pourraient permettre de déterminer les données exclues et conduire à des résultats injustes. L'exclusion des données protégées et de certaines variables pourrait causer un dilemme si elle réduisait de façon significative la quantité de données disponibles et affaiblissait la performance des modèles d'IA. Ce problème



devra alors être résolu en fonction des valeurs éthiques de l'institution financière.

Une autre approche possible consiste à demander aux institutions financières d'examiner les résultats et de veiller à l'équité du modèle par la mise à l'essai de celui-ci en fonction de divers attributs personnels. Selon les résultats des essais, les institutions financières pourraient ensuite modifier les modèles ou créer des modèles distincts pour les sous-populations de clients afin d'assurer un traitement uniforme de tous ses clients. Cependant, cette approche nécessiterait la collecte et l'enregistrement de ces attributs qui les distinguent. L'exclusion des données protégées a été encouragée à des fins d'équité et de protection de la vie privée; la collecte de données protégées est interdite dans un certain nombre de pays. Toutefois, l'absence de ces données nous empêche d'évaluer l'équité de l'IA. Les participants au forum ont convenu qu'un changement de culture était nécessaire pour que la collecte et l'utilisation de l'information soient considérées comme utiles pour évaluer les décisions de modélisation, au lieu de se faire au détriment des groupes de consommateurs. La collecte et l'utilisation de données protégées pour l'IA sont des domaines qui demeurent à explorer.

Néanmoins, il pourrait être judicieux que les organisations examinent leurs stratégies de données avec les équipes juridiques, de conformité et de commercialisation afin de s'assurer que leurs objectifs répondent aux exigences réglementaires et légales en matière d'utilisation des données ainsi qu'aux attentes des clients. Les organisations pourraient ainsi mieux éviter les problèmes de conformité légale et réglementaire. De plus, en faisant preuve de transparence à l'égard de l'utilisation de leurs

données, ils pourraient établir une meilleure relation de confiance avec les clients.

Il a été reconnu que la société est de plus en plus exigeante envers les institutions financières pour ce qui est du maintien de normes éthiques élevées. En outre, les institutions financières s'exposent à un risque d'atteinte à la réputation réel et peuvent en subir les répercussions lorsque des préjudices, réels ou perçus, sont causés aux clients.

## Équité

Il n'existe pas de définition universelle de l'équité et, en plus, notre perception de l'équité est influencée par le contexte. Dans le domaine de l'équité algorithmique, il existe différentes définitions mathématiques, dont certaines qui se contredisent<sup>[14]</sup>. Lorsqu'il est question d'équité dans le contexte de l'IA, il s'agit plutôt d'éviter la discrimination à l'égard de personnes ou de groupes de personnes. D'un point de vue juridique, la *Charte canadienne des droits et libertés* établit des règles claires pour éviter la discrimination.

L'iniquité peut concerner une personne ou un groupe particulier, ce qui explique que des mesures d'équité « individuelles » et « collectives » ont été élaborées. L'équité individuelle signifierait la parité entre des personnes semblables, tandis que l'équité de groupe signifierait la parité entre les groupes, tels que les groupes démographiques. Bien que les deux soient importants, il s'agit de deux concepts qui se définissent différemment et qu'il



n'est généralement pas possible d'optimiser simultanément.

Malgré l'aspect juridique, le respect de la loi ne signifie pas toujours que les actions et les résultats sont équitables ou perçus comme tels. Par exemple, la commercialisation de produits uniquement auprès de certains groupes de personnes peut être considérée comme inéquitable par certaines personnes et équitable par d'autres. Une commercialisation ciblée de la sorte par les institutions financières découle de la segmentation générée par l'IA afin de transmettre le bon message ou le bon produit au bon client de manière rentable.

Il faudrait faire preuve de beaucoup de rigueur pour comprendre l'équité des décisions prises ou des mesures prises en fonction des résultats générés par l'IA. En outre, les institutions financières auraient besoin de définir ce qui est considéré comme « équitable » en fonction de l'utilisation particulière d'une application d'IA.

### Biais

Les termes biais et préjugés sont souvent utilisés pour parler d'iniquités; ces termes sont souvent utilisés dans le même contexte. Un préjugé, ou biais (traduction de l'anglais « bias »), est communément défini comme « une inclination ou un parti pris pour ou contre une personne ou un groupe, en particulier d'une manière considérée comme injuste ».

Les biais ou préjugés peuvent provenir de différentes sources et prendre diverses formes. Par exemple :

- Les préjugés liés à des données de nature

délicate ou protégées, telles que le sexe, la religion ou l'origine ethnique, peuvent être dus à des facteurs sociaux reflétés dans les données historiques, à un manque de représentation dans les données recueillies par les organisations ou à une collecte de données inadéquate.

- Les experts en science des données peuvent introduire un biais par le choix des données ou de l'algorithme sélectionné.
- Les humains qui utilisent les résultats des modèles d'IA pour prendre des décisions peuvent introduire un préjugé lorsqu'ils décident d'ignorer les résultats des modèles d'IA.
- Un biais peut apparaître dans un modèle mal entraîné, même lorsque les données d'entraînement sont entièrement impartiales.
- Le préjugé peut découler du choix de la mesure des résultats.

Cependant, il a également une signification plus technique qui cadre avec les objectifs du modèle. Cette définition du biais est le résultat souhaité d'un modèle. Par exemple, les modèles d'assurance comportent un biais contre les conducteurs ayant de mauvais antécédents de conduite et les obligent à payer des primes plus élevées.

### Évaluation et intervention

La mesure des résultats et de l'équité sont des moyens de détecter la discrimination, il a donc été recommandé que les institutions se munissent de cadres de gouvernance similaires à ceux utilisés pour surveiller la performance des modèles. Bien sûr, l'équité pourrait être prise en compte par l'institution financière au moyen de nouvelles



évaluations, révisions et conceptions préalables de modèles, mais un écart pourrait subsister pour les anciens modèles qui n'ont pas fait l'objet de ces révisions.

Il a également été reconnu que la mesure des résultats en matière d'équité dépend du contexte, car différentes mesures d'équité s'appliqueraient dans différents contextes. Par exemple, l'équité dans une décision d'octroi de crédit peut être définie différemment que dans un processus d'embauche. Par conséquent, différentes mesures d'équité seraient plus appropriées selon les contextes.

Il est vrai que des biais et de l'injustice peuvent découler des données, mais ces dernières contiennent également des solutions. Les institutions financières pourraient mettre l'accent sur la représentation des données, explorer l'utilisation de données synthétiques ou utiliser des techniques de traitement interne ou de post-traitement pour lutter contre la discrimination.<sup>[15]</sup> Comme on l'a souligné ci-dessus, des données protégées seraient normalement nécessaires pour l'évaluation et l'intervention.

#### Utilisation de l'apprentissage machine (ML) contradictoire dans la prise de décisions équitables

Edwards et Storkey (2016) ont utilisé l'apprentissage machine contradictoire pour régler le problème d'équité que soulevait l'apprentissage machine. Selon leur approche, une décision est équitable si elle ne dépend pas de données de nature délicate telles que le sexe, l'âge ou la race. Ensuite, ils ont construit des données synthétiques qui préservent les informations sur les données d'origine, à l'exception de la dépendance à la variable sensible.<sup>[16]</sup>

## Vie privée et droit de recours

Le secteur des services financiers fait l'objet d'un examen plus minutieux que les autres secteurs en ce qui concerne le consentement des clients et la protection de la vie privée. On s'attend à ce que les organisations consacrent des ressources adéquates à la protection de la vie privée et des données des clients.

Avec l'utilisation croissante des décisions fondées sur l'IA, on doit veiller à ce que les clients aient accès aux informations sur l'utilisation de leurs données, dans la transparence, en plus de disposer de recours efficaces et efficaces pour contester les décisions prises par l'IA.

### Consentement du client

Le caractère subjectif de l'éthique et des valeurs pourrait avoir une incidence sur le consentement. Pour les mêmes données, certaines personnes peuvent librement donner leur consentement tandis que d'autres y seraient réticents. Le consentement du client est requis pour recueillir et utiliser ses données. À cet égard, les participants au forum ont signalé des défis et formulé certaines recommandations. Ils ont forgé le terme « dérive de consentement » pour parler des cas où les clients donnent leur consentement pour que les données soient utilisées à une fin particulière, mais au fil du temps, les mêmes données sont utilisées pour autre chose. De tels cas nécessiteraient une gestion continue du consentement.

Pour éviter les effets négatifs sur les clients, les institutions financières devraient s'assurer que les clients sont bien informés des conséquences de leur consentement. Pour ce faire, les documents



de consentement (p. ex., les modalités) doivent être courts, faciles à lire et à comprendre. Il est essentiel que les institutions financières tiennent compte des défis auxquels les clients peuvent être confrontés lorsqu'ils donnent leur consentement. Par exemple, une certaine donnée peut sembler inoffensive en soi, mais elle peut avoir des conséquences inattendues si elle est combinée à d'autres données. Les clients peuvent avoir du mal à anticiper les conséquences de leur consentement. Le consentement est nécessaire et devrait être demandé. Cependant, s'il est impossible d'obtenir le consentement des clients ou des clients possibles, les institutions financières pourraient éprouver des difficultés dans l'adaptation de leurs produits et dans l'amélioration de l'inclusion financière.

La numérisation est l'un des facteurs qui peuvent promouvoir la sensibilisation des clients aux répercussions du consentement. Les clients peuvent trouver qu'il est difficile de poser des questions, de demander des éclaircissements ou d'avoir une discussion avec les employés de l'organisation s'ils doivent utiliser un formulaire numérique.

Certaines questions éthiques (et même juridiques) pourraient découler de l'information recueillie à grande échelle lorsqu'il est impossible d'obtenir le consentement. Ces situations pourraient notamment causer des atteintes à la vie privée, ce qui peut être particulièrement préoccupant lorsque des renseignements personnels sensibles, comme des données financières et médicales, sont recueillis.

### **Assurer la confidentialité des données**

Les techniques d'amélioration de la protection de la vie privée permettent la collaboration et la communication de renseignements sensibles

d'une manière qui préserve la vie privée. Parmi ces techniques, notons le chiffrement homomorphe, l'apprentissage fédéré, le calcul multipartite sécurisé, la confidentialité différentielle et la pseudonymisation.

L'utilisation de certaines techniques héritées, telles que la pseudonymisation, ne garantit cependant pas une confidentialité entière des données, et les organisations doivent en tenir compte. Par exemple, certains types d'informations privées peuvent être recrées lorsque des ensembles de données pseudonymisés sont combinés à d'autres ensembles de données.

Ces techniques ont démontré beaucoup de potentiel dans certains cas d'utilisation et sont très prometteuses pour l'avenir. Par contre, leur utilisation générale ne garantit pas la conformité aux lois sur la protection des renseignements personnels telles que la *Loi canadienne sur la protection des renseignements personnels et les documents électroniques* (LPRPDE).

Il pourrait également y avoir des moyens non techniques, tels que des mécanismes de gouvernance appropriés, qui seraient utiles pour améliorer la confidentialité des données et, dans une certaine mesure, assurer la protection des données et des modèles.





## Opérationnalisation de l'éthique et des structures organisationnelles de l'IA

L'opérationnalisation de l'éthique de l'IA est essentielle. Il est donc important que les organisations fassent preuve de transparence, tant à l'interne qu'à l'externe, et communiquent de manière proactive la méthode utilisée pour garantir que leurs modèles d'IA respectent des normes éthiques élevées. De plus, puisque les normes éthiques ont tendance à changer, il est nécessaire de documenter la prise de décisions. Cette même documentation est également nécessaire à des fins d'audit.

Les participants ont recommandé de faire appel à des tiers ou à des organismes indépendants pour effectuer des évaluations de l'incidence des risques, de la protection de la vie privée, des biais et de l'équité, mais ils sont conscients des importants défis en matière de responsabilité que présentent ces recommandations. Certains de ces défis sont résumés dans la section « Explicabilité » des principes E D G E, sous la rubrique « Communication par des tiers ». En outre, pour effectuer une telle évaluation, il faudrait accorder à l'organisme indépendant un accès aux variables protégées. Le recours à des tiers pourrait faire l'objet d'une étude plus approfondie.



# Réglementation

La réglementation est au service de la société; elle permet d'assurer la sécurité et la solidité du système financier et protège les consommateurs.

La question de l'équilibre entre la réglementation et l'innovation fait l'objet d'un débat permanent, c'est-à-dire qu'il s'agit d'établir une réglementation solide tout en veillant à ce que les institutions financières continuent d'innover et de rester concurrentielles.

L'IA a fourni et continuera d'offrir des avantages aux institutions financières et à leurs clients. Avec de tels avantages, on s'attend à ce que les institutions financières intègrent de plus en plus l'IA dans leurs produits, processus et prises de décision. Cependant, la concrétisation et l'intensification de nouveaux risques associés à une telle technologie obligent diverses compétences à formuler des règlements.

Les questions clés suivantes en matière de réglementation ont été abordées durant le forum :

Quel est l'état de la réglementation en matière d'IA à l'échelle mondiale?

Qu'attend-on des institutions financières?

Quelles sont les positions des organisations de réglementation?

**“L'IA est un outil. Son utilisation et sa réglementation sont propres au contexte.”**

Oliver Carew, chef d'équipe senior, Technologies financières (Fintech) chez EY, auparavant à la Banque d'Angleterre



## État de la réglementation et des politiques en matière d'IA des diverses administrations

La sensibilisation accrue au sujet des avantages et des risques de l'IA ainsi que l'intensification de discussions sur ces sujets ont poussé les organismes de réglementation et les décideurs à prendre des mesures pour prévenir ou atténuer les risques tout en veillant au maintien des avantages. Ces mesures contribuent de manière importante au maintien de la confiance du public dans l'utilisation de l'IA et dans les systèmes financiers.

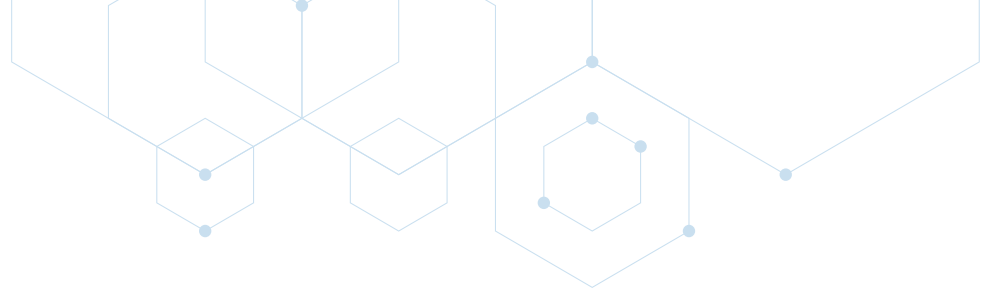
Au cours des dernières années, les décideurs et les organismes de réglementation de diverses administrations ont examiné les lois et règlements existants, sollicité les commentaires des intervenants concernant l'IA et rédigé de nouveaux règlements et de nouvelles politiques pour faire face aux risques liés à l'IA.

Les participants au forum ont exploré les activités récentes, et les publications connexes, menées dans diverses administrations, notamment au sein de l'Union européenne (UE), au Royaume-Uni (R.-U.), aux États-Unis (É.-U.) et ici au Canada.<sup>[17]</sup> La plupart des publications sont à l'état provisoire, mais elles donnent un aperçu des règlements et des politiques auxquels les institutions financières devront se conformer.

Les approches en matière de principes et de règlements pour l'IA des diverses administrations varient en termes de portée et d'incidence, mais elles s'harmonisent en ce qui concerne :

- La solidité de conceptualisation des modèles d'IA (p. ex., précision, fiabilité, robustesse, durabilité).
- L'importance de l'explicabilité des décisions à fort enjeu (y compris celles ayant une incidence sur le client).
- La mise en place d'une structure de gouvernance appropriée qui répond aux défis créés par l'IA (p. ex., transparence, responsabilité).
- L'IA ne doit pas causer de torts aux personnes ni à la société (p. ex., biais, discrimination, considérations éthiques, préoccupations en matière de protection de la vie privée).





## Caractéristiques d'une réglementation efficace

Les participants au forum ont exploré les éléments de qualité et de contenu des règlements qui peuvent en garantir le succès pour le secteur financier, c'est-à-dire favoriser l'innovation et gérer les risques connexes.

Avec le développement et l'adoption accrus de l'IA, une approche fondée sur des principes a été considérée comme le meilleur moyen de s'assurer que les orientations restent solides au fil du temps. De plus, il est important d'adopter des règlements clairs et équilibrés. Par exemple, les institutions veulent avoir la capacité d'innover; elles préféreraient avoir des règlements qui sont clairs sur ce qui est permis, mais pas normatifs au point d'étouffer l'innovation.

### Cohérence

Les institutions financières devront se conformer aux attentes des différents organismes de réglementation; certaines institutions exercent leurs activités sous différentes compétences. Le besoin d'uniformité a été soulevé afin d'éviter un fardeau réglementaire et des coûts liés à la conformité. Malgré la difficulté d'assurer l'uniformité entre les différents pays, il est important d'assurer l'uniformité au sein du Canada, c'est-à-dire entre les organismes de réglementation fédéraux et provinciaux. La nécessité d'harmoniser les règlements et les lois a également été soulignée en raison de la publication du projet de loi C-27, qui porte sur la protection de la vie privée, les données et l'IA. Les règlements devraient être adaptés au contexte canadien pour s'assurer qu'ils répondent

aux besoins des populations et qu'ils sont culturellement acceptables.

### Pratiques exemplaires

Les directives fondées sur des principes laissent souvent place à l'interprétation. Les participants au forum ont indiqué qu'il faudrait des orientations supplémentaires sur les pratiques exemplaires de mise en œuvre et de gouvernance efficaces de l'IA. Ces pratiques exemplaires représentent des recommandations sectorielles et ne sont pas des attentes réglementaires.

### Tiers

L'utilisation de données et de produits d'IA de tiers est inévitable, surtout pour les petites institutions. Des normes relatives à la gestion des risques liés aux tiers et/ou à l'examen indépendant des tiers ont été jugées nécessaires, et ont été envisagés afin d'éviter l'arbitrage réglementaire entre les solutions internes et les solutions fournies par les fournisseurs. Ces normes peuvent également porter sur les types de données de tiers que peuvent utiliser les institutions financières. Lorsque les données sont obtenues auprès de tiers, certaines protections sont essentielles pour assurer la gouvernance de la source, de la provenance et de la qualité des données.

### Commentaires

À l'étape de l'élaboration des lignes directrices, il est nécessaire de tenir compte des commentaires des intervenants. De plus, il devrait y avoir un



mécanisme pour recevoir les commentaires des intervenants afin de permettre l'amélioration de la réglementation.

### Bacs à sable et plateformes

La création de bacs à sable réglementaires pourrait également permettre aux institutions d'innover et de mettre à l'essai de nouvelles idées tout en sensibilisant et en suscitant la participation des organismes de réglementation. Parmi les suggestions : les organismes de réglementation devraient encourager le secteur à créer une plateforme où les institutions financières peuvent discuter de sujets clés liés à l'IA et explorer les pratiques exemplaires.

### Proportionnalité

Les règlements devraient tenir compte des différences de taille, d'importance relative et de capacités organisationnelles entre les institutions financières. Les petites institutions financières pourraient avoir besoin de s'appuyer sur des parties externes durant le processus d'adoption de l'IA.

En plus des exigences en matière de qualité et de contenu d'une réglementation efficace, les participants au forum ont recommandé que les organismes de réglementation fassent la promotion de la littératie en IA, y compris sur les données et le consentement. Ces organismes devraient également encourager les institutions financières à faire de même afin d'élargir l'inclusion financière au Canada.

## Point de vue des organismes de réglementation

La sollicitation des points de vue des participants au forum concernant l'élaboration de règlements solides et efficaces a bien sûr été bénéfique, mais il était également important d'obtenir directement l'opinion des organismes de réglementation. Malgré la longueur d'avance prise par le secteur financier sur de nombreux autres secteurs en ce qui concerne la compréhension des risques liés aux modèles, il est nécessaire que les organismes de réglementation se tiennent au courant des nouveaux risques découlant de l'adoption de l'IA.

### Harmonisation

Même si les organismes de réglementation ont des mandats différents et que leur portée n'est pas la même, il est nécessaire d'harmoniser la réglementation le mieux possible. Il faut également tenir compte des questions qui pourraient surgir à mesure que de nouvelles technologies ou solutions, comme un système bancaire ouvert, se mêlent aux technologies existantes. Donc, il est nécessaire d'harmoniser parmi les organismes canadiens de réglementation, et cette même uniformité est aussi requise à l'échelle internationale entre les diverses administrations.

### Innovation

While financial institutions need to innovate and Ce ne sont pas que les institutions financières qui doivent innover et gérer les risques qui découlent de cette innovation, les organismes



## Éthique

de réglementation doivent également innover, mener des essais, en tirer des leçons et tirer parti des observations d'autres secteurs sur l'utilisation de l'IA. D'ailleurs, certains organismes de réglementation se sont déjà munis d'un bureau de l'innovation. Il est important que les organismes de réglementation ne soient pas perçus comme un obstacle à l'innovation des institutions financières. Discuter et échanger avec le secteur privé pourrait être bénéfique.

### Collaboration

La collaboration entre les organismes de réglementation peut contribuer à l'harmonisation de la réglementation et soutenir l'innovation en matière de réglementation. Les organismes de réglementation provinciaux sont plus petits et ont parfois un double mandat, prudentiel et déontologique, et s'appuient parfois sur les organismes de réglementation fédéraux plus grands. La collaboration est d'autant plus importante.

### Orientation multidisciplinaire

Various risks arise or increase with AI, such as legal and compliance. AI needs a different approach for model risk governance. Consideration should be given to AI embedded

into software applications and systems. Diversity of thought is fundamental.

### Éducation

Étant donnée le rythme auquel se produit l'innovation en IA et les répercussions de celles-ci, les associations sectorielles pourraient contribuer à l'éducation de leurs membres sur les aspects clés de l'IA tels que les biais, l'éthique et la gouvernance des modèles, ainsi que sur les défis et les considérations.

### Petites institutions financières

La proportionnalité et l'externalisation doivent être prises en compte dans la réglementation des petites institutions financières. Ces institutions financières peuvent trouver difficile d'adopter l'IA, ce qui aura une incidence sur leur capacité à innover et à croître. En outre, ces institutions financières peuvent devoir s'appuyer davantage sur des tiers et/ou externaliser le développement de l'IA.





Éthique





# Conclusions

Lorsque l'IRM et le BSIF ont commencé à organiser le FIASSF, ils avaient l'intention de réunir les intervenants du secteur pour faire progresser les pratiques exemplaires de certains aspects de la gestion des risques liés à l'utilisation de l'IA dans le secteur canadien des services financiers.

Pendant le forum, il était gratifiant de voir les échanges et la contribution des participants, ce qui a démontré l'importance des occasions de discuter de sujets d'intérêt et de collaborer pour trouver des solutions aux défis. La collaboration entre les experts en IA de différentes institutions a été très enrichissante et a permis d'obtenir des informations importantes.

Il est évident que les différents sujets ont atteint différents stades de compréhension des défis et des solutions pratiques connexes. Parmi les sujets E D G E, les données et la gouvernance font partie des cadres des institutions depuis longtemps, tandis que l'explicabilité et l'éthique ont seulement récemment gagné en importance. Ces stades de maturité se sont reflétés dans la longueur de la discussion ou la profondeur de la conclusion tirée pour les différents sujets qui ont été abordés lors du forum.

Alors que l'utilisation de l'IA dans les institutions financières continue de progresser et d'évoluer, les conclusions tirées lors du forum étaient qu'une confiance plus forte en la technologie doit être établie pour aider à en accélérer l'adoption. Pour s'attirer la confiance des clients, il est nécessaire d'atteindre des niveaux appropriés d'explicabilité et de communication pour chaque cas d'utilisation, et d'adopter une approche centrée sur le client qui respecte les valeurs éthiques et protège la vie privée. Un cadre de gouvernance solide fondée sur le risque, qui comprend la gouvernance des données, avec une collaboration multidisciplinaire, favorise la confiance au sein des institutions financières. Un cadre de réglementation harmonisé qui respecte les valeurs culturelles canadiennes et qui est simple à respecter, tout en offrant des balises pour l'innovation, favorise la confiance dans l'ensemble du système financier et de la collectivité.

Les participants ont aussi parlé de l'importance de l'éducation générale au sujet de l'IA, de ses avantages et de ses limites. Celle-ci peut jouer un rôle de catalyseur dans l'adoption de l'IA. L'incidence de l'utilisation de l'IA doit être connue par l'ensemble du milieu des services financiers. Notons les aspects liés à l'utilisation des données, la gouvernance des données et la confidentialité des données, les considérations éthiques, l'incidence sur les clients, ainsi que les forces et les inconvénients des techniques d'IA particulières. L'intégration efficace des outils d'IA dans l'institution financière nécessite une combinaison d'expertise technique, de changements organisationnels et de changements culturels.

Les services tiers ont aussi fait l'objet de discussions, car ils sont présents dans presque tous les aspects du développement de solutions d'IA. Les outils de gestion des données, la gouvernance des données, le développement de modèles, la gouvernance de modèles, l'établissement de rapports, la gestion des risques liés aux modèles et le code source ouvert ne sont que quelques-uns des nombreux espaces où les services tiers peuvent être utiles dans

l'adoption de l'IA. Il a également été question de l'émergence des services tiers qui visent la conformité à la réglementation ou qui certifient les institutions en fonction des normes du secteur. Les avancées dans le domaine permettent de tirer parti du travail de différents acteurs, et entraînent également des défis qui n'étaient pas aussi répandus auparavant. Ce domaine nécessitera d'autres travaux et davantage d'exploration à mesure que l'utilisation de l'IA progressera dans le secteur des services financiers.

Il reste encore beaucoup à faire. Les idées et les discussions du forum représentent une première étape vers l'adoption de l'IA par les institutions financières canadiennes. Le forum a démontré la nécessité et les avantages d'une collaboration et d'une approche multidisciplinaire à différents niveaux.

# Remerciements

Nous tenons à exprimer notre sincère gratitude aux intervenants et aux participants aux ateliers pour leurs présentations significatives et leur engagement actif, qui se sont avérés déterminants pour les discussions et l'achèvement de ce rapport.

Nous remercions tout particulièrement Obim Okongwu du BSIF et Alexey Rubtsov de l'IRM pour leur gestion des différents aspects et intervenants du forum, du début à la fin. Nous sommes également très reconnaissants envers Conor Holler, Shelley Gronski, Paulina Prucnal et leur équipe pour l'excellent travail de conception et d'animation des ateliers. Nous remercions l'IRM et la Rotman School of Management d'avoir fourni des locaux pour les ateliers, ainsi que les équipes de l'approvisionnement, des affaires juridiques et des communications du BSIF pour leur soutien dans cette entreprise.

## Conférenciers visionnaires

### Avenir de l'IA

- Stuart Davis, vice-président à la direction, Gestion du risque de crimes financiers et chef de groupe, Lutte contre le blanchiment d'argent, Banque Scotia
- Foteini Agrafioti, scientifique en chef, RBC, et cheffe de Borealis AI
- Andrew Moore, vice-président et directeur général, IA infonuagique et solutions, Google

### Explicabilité

- Alex Wong, professeur et titulaire d'une chaire de recherche du Canada en IA et en imagerie médicale, Université de Waterloo
- David Heike, directeur général et chef du risque de modélisation, Services bancaires aux consommateurs et à la collectivité, JPMorgan Chase & Co.
- Agus Sudjianto, vice-président de la direction et responsable du risque de modélisation, Wells Fargo

### Données

- Ima Okonny, dirigeante principale des données, Emploi et Développement social Canada (EDSC)

### Gouvernance

- Donna Bales, co-fondatrice de l'Association canadienne de la technologie réglementaire
- David Palmer, analyste principal en surveillance financière, Réserve fédérale américaine

### Éthique

- Carole Piovesan, associée directrice, INQ Law

### Perspectives canadiennes et internationales

- Oliver Carew; expert en IA et chef d'équipe senior, Technologies financières chez EY, auparavant à la Banque d'Angleterre
- Qiang Zhang, directeur adjoint, Bureau de développement de l'IA, Autorité monétaire de Singapour (par vidéoconférence)

# Autres conférenciers

## Perspectives sur l'IA

- Bruce Choy; Directeur principal, recherche, Global Risk Institute
- Romana Mizdrak; Directrice principale, Quantification des risques, BSIF

## Perspectives du Canada

- Alexey Rubtsov; Professeur associé, Département des mathématiques, Toronto Metropolitan University
- Ana Garcia; Directrice, Quantification des risques, BSIF

## Présentations en vitrine

- Benoit Deshaies; Directeur – Données et intelligence artificielle, Conseil du trésor du Canada
- Bradley Fedosoff; VPS architecture, données et analytiques, CIBC
- Eugene Wen; VP groupe de l'analytique avancée, Manulife
- Greg Kirczenow; Directeur senior – Gestion des risques de modélisation IA, RBC
- Michaela Capra; VPA – Risques corporatifs – innovation numérique, Sunlife
- Shingai Manjengwa; Directeur – éducation technique, Vector Institute
- Stephanie Kelley; Professeur adjoint, Ivey Business School à University of Western Ontario




# Participants

- **Banque Amex du Canada**, Pat Smith
- **Antara Risk Management**, Sanjiv Talwar
- **Banque du Canada**, Maryam Haghighi
- **Autorité des services financiers de la Colombie-Britannique (BC Financial Services Authority)**, Steven Wright
- **Banque de Montréal**, Drew Galow, Letitia Golubitsky, Suyi Chen
- **Banque de développement du Canada**, Sherrilyn Lequin
- **Société d'assurance-dépôts du Canada**, Neville Arjani
- **Office d'investissement du Régime de pensions du Canada**, Brendon Freeman
- **Institut canadien des actuaires**, Joel Li
- **Institut canadien des actuaires**, Paul Childeross
- **Banque CIBC**, Brad Fedosoff, Ozge Yeloglu
- **Decca et McKinsey** (Anciennement auprès de) Matthew Killi
- **Autorité ontarienne de réglementation des services financiers**
- **Centre d'analyse des opérations et déclarations financières du Canada**, Ivy Ou, Nathalie Martineau
- **Institut du risque mondial**, Mike Stramaglia, Alexey Rubtsov, Bruce Choy, Mark Engel
- **Innovation, Science et Développement économique Canada**, Anastasiia Tryputen, Surdas Mohit
- **Intact Corporation financière**, Sebastien Bernard
- **Manuvie**, Eugene Wen, Henry Li
- **Northbridge Corporation financière**, Cheston Chiu
- **OMERS**, Richard Slessor, Sami Ahmed
- **Commission des valeurs mobilières de l'Ontario**, Levin Karg
- **BSIF**, Ana Garcia, Greg Caldwell, Karyn Leung, Mohamad Al-Bustami, Obim Okongwu, Patrick Cane, Regis Dahany, Romana Mizdrak, Sharon Chambers Creary, Stephen Manly
- **RiverRun Ventures GP**, Daniel Moore
- **Banque Royale du Canada**, Greg Kirczenow, Jun Yuan, Dominique Payette
- **Institut Schwartz Reisman pour la technologie et la société**, Monique Crichlow
- **Banque Scotia**, Carrie Chai, Gail Towne, Michel Valentik
- **Sunlife**, Mihaela Capra
- **Banque TD**, Paige Dickie, Baiju Devani
- **Secrétariat du Conseil du Trésor du Canada**, Benoit Deshaies
- **Université de Toronto**, Zissis Poulos, John Hull
- **Université de Toronto**, Andres Rojas, Shingai Manjengwa
- **Université Western**, Cristian Bravo

# Références

1. Slack, D., Hilgard, S., Jia, E., Singh, S., Lakkaraju, H. (2019) Fooling LIME and SHAP: Adversarial Attacks on Post hoc Explanation Methods. Repéré à <https://arxiv.org/pdf/1911.02508>
2. Cynthia Rudin (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead [Cessez d'expliquer les modèles d'apprentissage automatique en boîtes noires pour les décisions importantes et utilisez plutôt les modèles interprétables] Nature Machine Intelligence 1, 206-215
3. La valeur notionnelle peut représenter le montant nominal initial de la transaction ou le montant nominal ajusté en fonction du taux de conversion en devises étrangères, des intérêts courus, ou du multiplicateur de gains.
4. En effet, le traitement des données protégées est très variable : Singapour autorise la collecte et l'utilisation de données sur le sexe dans les modèles d'IA. L'Union européenne autorise la collecte de données sur le sexe, mais interdit l'utilisation du sexe dans les modèles d'entraînement et de filtrage utilisés dans les décisions de prêts individuels. Les États-Unis interdisent la collecte et l'utilisation de données liées au sexe. Le Canada n'interdit pas explicitement la collecte et l'utilisation de données protégées établies en vertu de la Charte des droits et libertés aux fins d'évaluation des préjugés. Au fédéral, les données protégées sont précisées dans la Charte canadienne des droits et libertés. Les provinces ont ajouté des données à cette liste (avec certaines variations selon la province). Même au sein d'une province, certains secteurs comme le secteur de l'assurance traitent les clients différemment sur la base de données protégées (p. ex., les jeunes hommes paient des taux d'assurance automobile plus élevés).
5. Le rapport final de la Banque d'Angleterre intitulé « Artificial Intelligence Public-Private Forum » (Forum public-privé sur l'intelligence artificielle) fait également mention d'une approche semblable de type « étiquetage des aliments ». Voir la page 20 du rapport qui peut être consulté sur le site suivant (en anglais seulement) : <https://www.bankofengland.co.uk/-/media/boe/files/fintech/ai-public-private-forum-final-report.pdf?la=en&hash=F432B83794DDF3F580AC5A454F7DFF433D091AA5>
6. Fondation canadienne pour l'audit et la responsabilisation, Guide pratique sur l'audit de la surveillance (<https://www.caaf-fcar.ca/fr/surveillance-concepts-et-contexte/qu-est-ce-que-la-surveillance-et-quel-est-son-lien-avec-la-gouvernance/qu-est-ce-que-la-gouvernance>)
7. Directive SR11-7 de la Réserve fédérale américaine (<https://www.federalreserve.gov/supervisionreg/srletters/sr1107.htm> (en anglais seulement))
8. Directive E-23 du BSIF : Gestion du risque de modélisation à l'échelle de l'entreprise dans les institutions de dépôts (<https://www.osfi-bsif.gc.ca/fr/fi-if/rg-ro/gdn-ort/gi-lid/Pages/e23.aspx>)
9. Secrétariat du Conseil du Trésor, gouvernement du Canada. Directive sur la prise de décisions automatisée (<https://www.tbs-sct.canada.ca/pol/doc-fra.aspx?id=32592>)
10. Rapport de synthèse sur la gouvernance de l'apprentissage automatique du Institute of International Finance ([https://www.iif.com/portals/0/Files/content/Innovation/12\\_4\\_2020\\_mlg\\_summaryreport.pdf](https://www.iif.com/portals/0/Files/content/Innovation/12_4_2020_mlg_summaryreport.pdf) (en anglais seulement))
11. Deniz Tudor; Model Risk Management Lessons Learned: Tracing Issues from the Pandemic to the Great Recession (<https://www.garp.org/risk-intelligence/operational/model-risk-management-lessons-220708> (en anglais seulement))
12. Opérations d'apprentissage automatique (ou MLOps) : l'expression fait référence aux pratiques, processus et outils que les organisations utilisent pour gérer la création et le déploiement de modèles. Il s'agit d'une extension du processus de développement et d'exploitation (DevOps), qui se concentre sur l'automatisation et la rationalisation du processus de développement logiciel, dans l'écosystème de l'apprentissage automatique. DevOps se concentre sur le développement et le déploiement de logiciels, tandis que MLOps touche également à l'ingénierie des données et au déploiement de modèles.
13. Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector [Comprendre l'éthique et la sécurité de l'intelligence artificielle : un guide pour la conception et la mise en œuvre responsable de systèmes d'IA dans le secteur public]. The Alan Turing Institute. <https://doi.org/10.5281/zenodo.3240529> (en anglais seulement)
14. Voir, notamment, le document « FEAT Principles Assessment Case Studies (Veritas Document 4) » (en anglais seulement), la Charte canadienne des droits et libertés (<https://www.justice.gc.ca/fra/sjc-csj/dlc-rfc/ccdl-ccrf/pdf/charte-poster.pdf>), ainsi que le document intitulé « Towards the Right Kind of Fairness in AI (GETD | AI Research & Thought Leadership, May 2021) » (en anglais seulement).
15. Les approches axées sur les données pour réduire la discrimination comprennent le sous-échantillonnage, le suréchantillonnage, le réglage des hyperparamètres en fonction du sexe et la modélisation probabiliste du sexe par substitution.
16. Edwards, H., Storkey, A. Censoring Representations with an Adversary (2016) <https://arxiv.org/abs/1511.05897> (en anglais seulement)
17. La législation sur l'intelligence artificielle de 2021 (IA Act) de la Commission européenne, le rapport du Forum privé-public sur l'IA de 2022 au Royaume-Uni (AI Public-Private Forum: Final report), le document de travail sur l'intelligence et l'apprentissage automatique de 2022 au Royaume-Uni (DPS/22 - Artificial Intelligence and Machine Learning), les principes de gestion des risques pour les banques de 2022 au Royaume-Uni (CP6/22 - Model risk management principles for banks), le rapport du consortium Veritas de 2022 à Singapour, la demande d'information sur l'IA de 2021 aux États-Unis (Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning), la loi sur la responsabilité algorithmique de 2022 aux États-Unis (Algorithmic Accountability Act of 2022), la loi américaine sur la confidentialité et la protection des données de 2022 (American Data Privacy and Protection Act), le document de travail du BSIF intitulé Renforcer la résilience du secteur financier dans un monde numérique (2020); la lettre aux institutions financières au sujet de la ligne directrice E-23 (2022), le projet de loi C-27 (2022).

Image hexagonale créée par kjpargeter / Freepik1



No de cat. : IN4-35/2023F-PDF

Bureau du surintendant des institutions financières

255 rue Albert – 12<sup>ième</sup> étage

Ottawa, ON K1A 0H2

Téléphone : 1-800-385-8647

Courriel : [information@osfi-bsif.gc.ca](mailto:information@osfi-bsif.gc.ca)

© Sa Majesté le Roi du Chef du Canada, 2023

Also available in English



| BSIF |  
| OSFI |

