

**Recueil du Symposium de 2021 de Statistique Canada  
Adopter la science des données en statistique officielle pour répondre aux  
besoins émergents de la société**

**Utilisation responsable de  
l'apprentissage automatique à  
Statistique Canada**

par Keven Bosa et Yanick Beaucage

Date de diffusion : le 15 octobre 2021



Statistique  
Canada

Statistics  
Canada

Canada

## Utilisation responsable de l'apprentissage automatique à Statistique Canada

Keven Bosa et Yanick Beaucauge<sup>1</sup>

### Résumé

Un cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable a été élaboré à Statistique Canada. Le cadre comprend des lignes directrices pour l'usage responsable de l'apprentissage automatique et une liste de contrôle connexe, qui sont organisées en quatre thèmes : le respect des personnes; le respect des données; des méthodes éprouvées; une application rigoureuse. Les quatre thèmes mis en commun assurent l'utilisation éthique des algorithmes et des résultats de l'apprentissage automatique. Le cadre est ancré dans une vision qui cherche à créer un milieu de travail moderne et à fournir une orientation et un soutien à ceux qui utilisent les techniques d'apprentissage automatique. Il s'applique à tous les programmes et projets statistiques menés par Statistique Canada qui utilisent des algorithmes d'apprentissage automatique. Cela comprend les algorithmes d'apprentissage supervisés et non supervisés. Le cadre et les lignes directrices le supportant seront présentés dans un premier temps. Le processus de revue des projets utilisant l'apprentissage automatique, soit la façon dont le cadre est appliqué aux projets de Statistique Canada, sera ensuite expliqué. Finalement, des travaux futurs pour améliorer le cadre seront décrits.

Mots Clés : Apprentissage automatique responsable, explicabilité, éthique

### 1. Introduction

De plus en plus de données sont générées au quotidien. On n'a qu'à penser aux données de téléphonie cellulaire, d'images satellites, de navigation sur internet ou de lecteur optique. La profusion de données ouvre la voie à la création de nouvelles statistiques, plus détaillées et plus actuelles permettant de mieux servir la population. Comme plusieurs autres organismes nationaux de statistique, Statistique Canada a adhéré à cette nouvelle réalité et prend de plus en plus avantage des sources de données alternatives afin d'améliorer et moderniser ses différents programmes statistiques. Étant donné leur volume et leur vélocité, des méthodes d'apprentissage automatique sont souvent nécessaires pour extraire et transformer l'information précieuse contenue dans ces nouvelles sources de données.

Statistique Canada a mené plusieurs projets faisant appel à des méthodes d'apprentissage automatique au cours des trois dernières années. Par exemple, les scientifiques de données se sont servis du traitement de langage naturel pour classer des commentaires provenant de répondants au recensement ainsi qu'à d'autres enquêtes et ainsi accélérer leur résolution en les distribuant automatiquement aux spécialistes appropriés, sans avoir besoin de les lire. Des méthodes d'apprentissage non supervisé ont permis de partitionner la base canadienne de données des coroners et des médecins légistes en groupes homogènes afin d'améliorer la compréhension de certains phénomènes. Un algorithme d'apprentissage supervisé a été développé pour prédire le rendement des cultures. Des projets utilisant des réseaux neuronaux sur des images satellites sont actuellement en cours pour optimiser le programme de l'agriculture en détectant, par exemple, la présence de serres ou encore en identifiant les différents types de grandes cultures dans le but de réduire le fardeau de réponse des agriculteurs. Un algorithme a aussi été élaboré pour extraire de l'information financière provenant de documents. Les exemples précédents donnent une idée de la diversité des problèmes pour lesquels l'apprentissage automatique peut faciliter le travail des agences statistiques.

---

<sup>1</sup>Keven Bosa, Statistique Canada, 100 promenade Tunney, immeuble R.-H. Coats, Ottawa, Ontario, Canada, K1A 0T6, [keven.bosa@statcan.gc.ca](mailto:keven.bosa@statcan.gc.ca) ; Yanick Beaucauge, Statistique Canada, 100 promenade Tunney, immeuble R.-H. Coats, Ottawa, Ontario, Canada, K1A 0T6, [yanick.beaucauge@statcan.gc.ca](mailto:yanick.beaucauge@statcan.gc.ca)

L'utilisation de l'apprentissage automatique comprend son lot d'avantages : traitement des données volumineuses et non structurées, automatisation des processus, amélioration de la couverture et de la précision et bien d'autres. Toutefois, elle soulève aussi plusieurs questions. Par exemple :

- Est-ce que le processus protège l'intégrité et la confidentialité des données ?
- Est-ce que la qualité des données d'apprentissage est adéquate pour le but poursuivi ?
- Une fois l'algorithme mis en place, qui est responsable des résultats et des effets qui en découlent ?

Suite à ces questions et l'utilisation grandissante de méthodes d'apprentissage automatique à Statistique Canada, la Direction des méthodes statistiques modernes et de la science des données a reconnu le besoin d'encadrer et guider l'élaboration des processus d'apprentissage automatique et d'en faire des processus responsables.

Le Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada sera présenté à la section 2. La section 3 expliquera brièvement le processus de revue mis en place pour appliquer ce Cadre. Finalement, la section 4 conclura cet article en proposant quelques réflexions et en mentionnant quelques travaux futurs.

Cet article est une version plus détaillée d'un article présenté dans l'édition du mois de juillet 2021 du bulletin du Réseau de la science des données pour la fonction publique fédérale (2021a).

## 2. Présentation du Cadre

Avant de présenter le cadre de travail dont s'est doté Statistique Canada, nous ferons un bref survol de la Directive sur la prise de décisions automatisée établie par le Secrétariat du conseil du trésor tel que spécifié dans le document Gouvernement du Canada (2019). Celle-ci a d'ailleurs fait l'objet d'un article présenté dans l'édition du mois de juin 2021 du bulletin du Réseau de la science des données pour la fonction publique fédérale (2021b). Il y est mentionné que : « *La Directive vise également à s'assurer que les systèmes décisionnels automatisés sont déployés d'une manière qui permet de réduire les risques pour les Canadiens et les institutions fédérales, et qui donne lieu à une prise de décisions plus efficace, exacte et conforme, qui peut être interprétée* ». Il est aussi mentionné que la Directive « *...s'applique aux systèmes qui formulent ou qui aident à formuler des recommandations ou des décisions* ». Une évaluation de l'incidence algorithmique doit être complétée pour les projets assujettis à cette directive. Cette évaluation prend en fait la forme d'un questionnaire auto-administré où des points sont attribués en fonction des réponses fournies à chaque question. Plus le pointage final est élevé, plus les décisions prises par le système pourraient avoir un impact élevé sur les droits, la santé, le bien-être ou les intérêts économiques des personnes ou des collectivités et plus il y aura d'exigences à satisfaire.

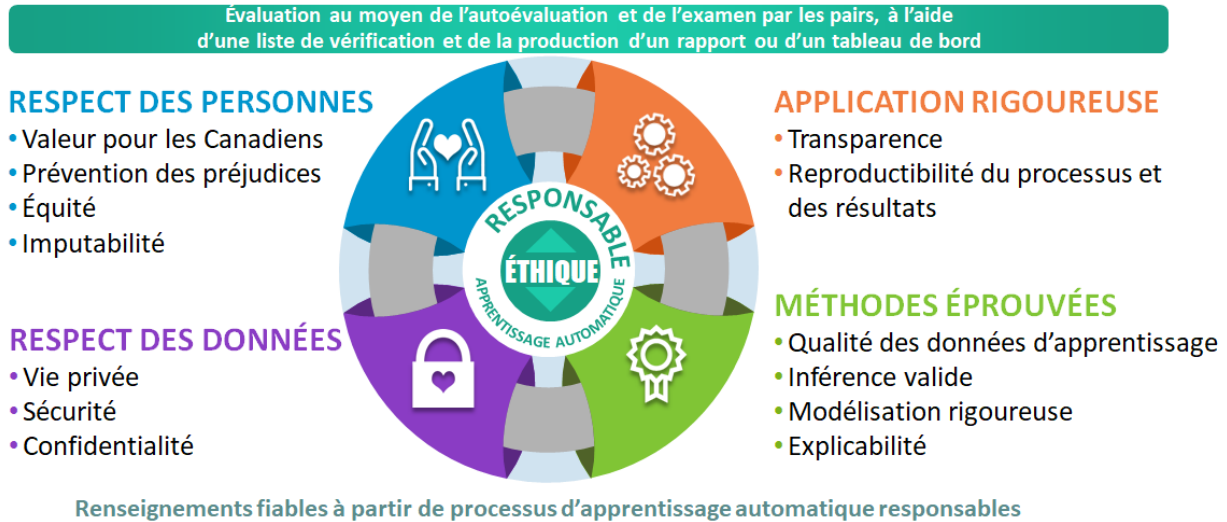
À Statistique Canada, tous les projets utilisant l'apprentissage automatique ou, de façon plus générale la modélisation, font partie de programmes statistiques dont le but n'est pas de prendre des décisions administratives sur un individu ou une entreprise, mais bien d'établir des statistiques pour décrire l'environnement dans lequel les Canadiens évoluent. Statistique Canada n'a donc pas encore eu à se conformer à cette Directive ou à utiliser l'Outil d'évaluation de l'incidence algorithmique. Toutefois, comme mentionné à la fin de la section précédente, Statistique Canada a été proactif en adoptant un cadre afin d'assurer une utilisation responsable de l'apprentissage automatique au sein de l'agence.

La figure 2-1 donne un bon aperçu du Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada. Ce dernier est organisé en quatre thèmes : respect des personnes; respect des données; application rigoureuse; méthodes éprouvées. Les quatre thèmes mis en commun assurent l'utilisation éthique des algorithmes et des résultats de l'apprentissage automatique. Chaque thème est défini par quelques attributs, listés à personne qui fournit sous chacun d'eux à la figure 2-1. Chaque attribut est par la suite défini à l'aide d'une ou plusieurs lignes directrices. Ces lignes directrices s'appliquent à tous les programmes et projets statistiques menés par Statistique Canada qui utilisent des algorithmes d'apprentissage automatique, particulièrement ceux mis en production. Cela comprend les algorithmes d'apprentissage supervisé et non supervisé. Le lecteur est invité à consulter

le Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada (2021) pour obtenir plus de détails ainsi que pour connaître les lignes directrices associées à chaque attribut.

Figure 2-1

### Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada



À Statistique Canada, nous souhaitons faire une utilisation efficace des ressources du gouvernement tout en produisant des renseignements qui aident les Canadiens à mieux comprendre leur pays. Le thème **respect des personnes** inclut quatre attributs : la valeur pour les Canadiens; la prévention des préjudices; l'équité; l'imputabilité.

1. Le concept de **valeur pour les Canadiens** dans un contexte d'apprentissage automatique implique que son utilisation doit avoir une valeur ajoutée, que ce soit dans les produits eux-mêmes ou par une plus grande efficacité dans le processus de production.
2. La **prévention des préjudices** nécessite d'être au courant des dangers potentiels et d'avoir un dialogue constructif avec les intervenants et les porte-parole du milieu avant la mise en œuvre d'un projet d'apprentissage automatique.
3. **L'équité** implique que le principe de la proportionnalité entre les moyens et les objectifs soit respecté, et qu'un équilibre soit maintenu entre des intérêts et des objectifs différents. L'équité veille à ce que les personnes et les groupes ne soient pas victimes de préjugés injustes, de discrimination ou de stigmatisation.
4. **L'imputabilité** est l'obligation juridique et éthique d'une personne ou d'une organisation d'être responsable de son travail et de communiquer les résultats du travail de façon transparente. Les algorithmes ne sont pas responsables; quelqu'un est responsable des algorithmes.

À Statistique Canada, nous prenons les données au sérieux. Le thème **respect des données** a trois attributs: la protection de la vie privée des personnes auxquelles les données appartiennent; la sécurité des renseignements tout au long du cycle de vie des données; la confidentialité de renseignements identifiables.

1. La **vie privée** est le droit de se retirer et de ne pas être sujet à une quelconque forme de surveillance ou d'intrusion. Lors de l'acquisition de renseignements de nature délicate, les gouvernements ont des obligations relativement à la collecte, à l'utilisation, à la divulgation et à la conservation des renseignements personnels. Le terme vie privée réfère généralement à des renseignements concernant des particuliers (définition tirée de la Politique sur la protection des renseignements personnels et la confidentialité de Statistique Canada (2020)).
2. La **sécurité** représente les dispositions fondées sur l'évaluation de la menace et des risques qu'utilisent les organisations pour empêcher l'obtention ou la divulgation inadéquate de renseignements confidentiels. Les

mesures de sécurité protègent aussi l'intégrité, la disponibilité et la valeur des fonds de renseignements. Cela englobe les protections matérielles, comme l'accès restreint aux zones où les renseignements sont entreposés et utilisés ou les autorisations de sécurité des employés, ainsi que les protections technologiques utilisées pour empêcher l'accès électronique non autorisé (définition tirée de la Politique sur la protection des renseignements personnels et la confidentialité de Statistique Canada (2020)).

3. La **confidentialité** fait référence à la protection contre la divulgation de renseignements personnels identifiables concernant une personne, une entreprise ou une organisation. La confidentialité suppose une relation de « confiance » entre celui qui fournit les renseignements et l'organisation qui les recueille; cette relation s'appuie sur l'assurance que ces renseignements ne seront pas divulgués sans l'autorisation de la personne ou sans l'autorité législative appropriée (définition tirée de la Politique sur la protection des renseignements personnels et la confidentialité de Statistique Canada (2020)).

Une **application rigoureuse** signifie de mettre en place, de maintenir et de documenter les processus d'apprentissage automatique de façon à ce que les résultats soient toujours fiables et que l'ensemble du processus puisse être compris et recréé. Ce thème a deux attributs : la transparence et la reproductibilité du processus et des résultats.

1. La **transparence** fait référence au fait d'avoir une justification claire de la raison pour laquelle cet algorithme et les données d'apprentissage sont les plus appropriés pour l'étude en cours. Pour être transparents, les développeurs devraient produire une documentation complète, y compris rendre accessible le code informatique à d'autres personnes, et ce, sans compromettre la confidentialité ou la protection des renseignements personnels.
2. La **reproductibilité du processus** signifie qu'il y a suffisamment de documentation et que le code informatique a été suffisamment partagé pour faire en sorte que le processus soit reproduit, à partir de zéro. La **reproductibilité des résultats** signifie que les mêmes résultats peuvent être reproduits de façon fiable lorsque toutes les conditions sont contrôlées. Il n'y a pas d'étapes qui modifient les résultats à la suite d'une intervention ponctuelle ou humaine.

Les **méthodes éprouvées** sont celles qui peuvent être invoquées de manière efficace et efficiente afin de produire les résultats espérés. Nous suivons habituellement des protocoles reconnus qui comportent une consultation avec des pairs et des experts, de la documentation et des tests lors de l'élaboration de méthodes éprouvées. Ce thème a quatre attributs : la qualité des données d'apprentissage; l'inférence valide; la modélisation rigoureuse; l'explicabilité.

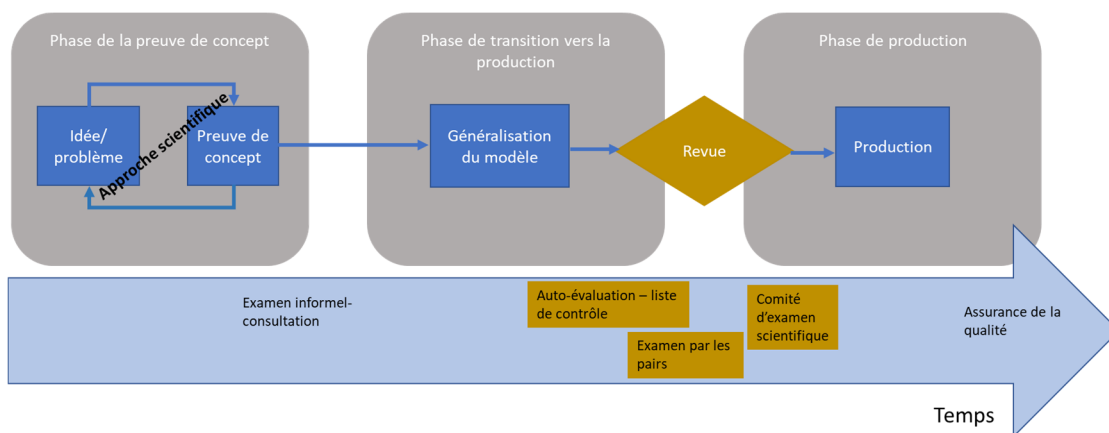
1. Dans un contexte d'apprentissage automatique, la **qualité des données d'apprentissage** est mesurée par la cohérence et l'exactitude des données étiquetées. La couverture, ce qui signifie que les données étiquetées et les descriptions couvrent tous les cas auxquels l'algorithme peut faire face lors de la production, est également importante pour réduire le risque de partialité ou de discrimination (équité). La couverture est également importante pour assurer la représentativité des variables, ce qui est important lorsqu'on veut obtenir des mesures de rendement réalistes.
2. Une **inférence valide** désigne la capacité d'obtenir, à partir d'un échantillon, des conclusions plausibles et une mesure de précision connue de la population cible. Dans un contexte d'apprentissage automatique, une conclusion valable signifie que, à partir de données tests (jamais utilisées pour la modélisation), les prédictions doivent être, dans une grande proportion, raisonnablement près de leurs vraies valeurs ou, dans le cas de données catégoriques, les prédictions sont exactes dans une grande proportion.
3. Une **modélisation rigoureuse** en apprentissage automatique consiste à s'assurer que les algorithmes sont vérifiés et validés. Cela permettra aux utilisateurs et aux décideurs de faire confiance à l'algorithme à juste titre du point de vue de l'adaptation des données à leur utilisation, de la fiabilité et de la robustesse.
4. Un modèle qui est **explicable** est un modèle qui est suffisamment documenté. Les documents doivent expliquer clairement le lien entre les données en entrée et les résultats. Ils doivent également permettre de déterminer quelles conclusions on peut tirer ou encore ce qui devrait être exploré plus en profondeur. En d'autres mots, un modèle explicable n'est pas une boîte noire.

### 3. Processus de revue

Le processus de revue constitue la mise en œuvre du Cadre faite à Statistique Canada. L'accent est mis sur les projets où l'utilisation de méthodes d'apprentissage automatique dans une ou plusieurs étapes mène à la production de statistiques officielles. La figure 3-1 permet de bien visualiser à quel moment le processus de revue d'un algorithme d'apprentissage automatique devrait être lancé.

Le premier encadré de la figure 3-1 représente la phase pendant laquelle la preuve de concept (ou prototype) est élaborée. Habituellement, cette phase débute par un problème à résoudre. Dans ce contexte précis, la preuve de concept considère des méthodes d'apprentissage automatique pour solutionner un problème. La démarche scientifique devrait être utilisée pour élaborer le prototype. Cette démarche débute par la postulation d'hypothèses en fonction des données disponibles. Ces hypothèses sont ensuite soumises à une expérience (ou simulation) permettant d'obtenir des résultats. Ces derniers sont interprétés, ce qui mène à une conclusion à propos de la véracité des hypothèses émises. En fonction de la conclusion, de nouvelles hypothèses peuvent être envisagées, ce qui correspond au cycle de l'approche scientifique représenté dans le carré gris à gauche de la figure 3-1. Lorsque le modèle développé durant la première phase est satisfaisant, une revue informelle du projet devrait être faite avant de passer à la phase suivante. Cette revue peut prendre la forme d'une présentation ou de consultations avec des collègues experts en science des données. S'il n'y a pas d'enjeux majeurs suivant la revue informelle, le modèle où la méthode sera généralisée pendant la phase de transition vers la production qui est représentée par le deuxième encadré de la figure. Par exemple, si seulement une partie des données a été utilisée durant la première phase, le modèle sera modifié pour traiter l'ensemble des données. La méthode pourrait aussi être appliquée à un plus grand nombre de variables lors de la généralisation. C'est vers la fin de la phase de transition que devrait débiter la revue formelle de la méthode. Le processus de revue comprend trois étapes : l'auto-évaluation à l'aide de la liste de contrôle ; l'évaluation par des pairs ; et une présentation du projet au comité d'examen scientifique de la Direction des méthodes statistiques modernes et de la science des données. Plus de détails à propos de ces étapes seront donnés aux sous-sections 3.1 à 3.3. Ultimement, le processus recommandera (ou non) le passage de l'application à l'étape de production. Un plan d'assurance de la qualité devrait faire partie de la phase de production pour réviser, en continu, les performances du modèle utilisé. Des ajustements comme le ré-entraînement du modèle ou la modification des hyperparamètres seront nécessaires si on détecte une détérioration de la qualité du modèle.

**Figure 3-1**  
**Visualisation du processus de revue pendant le cycle de vie d'un projet**



### 3.1 Auto-évaluation à l'aide de la liste de contrôle

**Figure 3.1-1**  
**Liste de contrôle**

**Méthodes éprouvées - Inférence valide**

English  Français

Ligne directrice no	Questions de la liste de contrôle	Oui	Non	Réponse écrite	Autoévaluation remplie par	Évaluation par les pairs menée par	Commentaires formulés pendant l'évaluation par les pairs
27	Décrivez le protocole de validation ou de diagnostic choisi et les mesures connexes et ce qui les rend appropriés compte tenu des objectifs et des mesures d'évaluation du programme statistique et du rôle de l'AA dans ce programme.						
28	Indiquez les mesures d'évaluation et les cibles et décrivez comment les cibles ont été établies.						
29	A-t-on mis de côté des données d'essai adéquates qui n'ont pas été utilisées dans le processus d'élaboration?	<input checked="" type="radio"/>	<input type="radio"/>				
30	Si les données sur l'apprentissage proviennent d'un échantillon, a-t-on utilisé des poids de sondage? Pourquoi ou pourquoi pas?	<input type="radio"/>	<input checked="" type="radio"/>				

L'équipe ayant développé le modèle ou l'application utilisant des méthodes d'apprentissage automatique doit faire une auto-évaluation concernant l'utilisation de ces techniques. Pour ce faire, l'équipe devra prendre connaissance du Cadre (voir Statistique Canada (2021)) et répondre aux questions de la liste de contrôle. La liste de contrôle prend la forme d'un questionnaire où, de façon générale, chaque ligne directrice du Cadre est reformulée sous forme d'une ou plusieurs questions. La figure 3.1-1 représente une partie du questionnaire concernant l'attribut *inférence valide* du thème *méthodes éprouvées*. Lorsque complété, ce questionnaire et la documentation du projet et des méthodes utilisées sont envoyés à l'équipe responsable de mener la revue.

### 3.2 Évaluation par les pairs

L'évaluation par les pairs consiste en une révision du projet faite par des experts de Statistique Canada qui n'ont pas été impliqués dans l'élaboration de la méthodologie proposée. Des réviseurs provenant de deux équipes différentes seront impliqués. D'une part, les questions et la documentation concernant les deux premiers thèmes du Cadre, « respect des personnes » et « respect des données », seront évaluées par l'équipe du Secrétariat de l'éthique des données. Le comité d'éthique interne ou externe pourrait également être impliqué en fonction de la sensibilité et des risques potentiels associés au projet faisant l'objet de l'évaluation. D'autre part, des experts en science des données seront responsables de l'évaluation de la méthode par rapport aux thèmes « application rigoureuse » et « méthodes éprouvées ». À la fin de cette évaluation, un rapport contenant des recommandations de toutes les parties impliquées dans la revue sera envoyé au gestionnaire du projet. La méthodologie proposée pourrait être réévaluée par les pairs en fonction des recommandations contenues dans le rapport et des réponses fournies par l'équipe travaillant sur le projet en question.

### 3.3 Présentation du projet au comité d'examen scientifique de la Direction des méthodes statistiques modernes et de la science des données

La dernière étape du processus de revue est la présentation du projet au comité d'examen scientifique de la Direction des méthodes statistiques modernes et de la science des données. Cette présentation expose la méthodologie utilisée lors du processus d'apprentissage automatique devant un comité d'experts. L'objectif de cette étape est de présenter le projet, les méthodes utilisées ainsi que la revue à un groupe d'experts afin d'obtenir leur approbation quant aux conclusions de la revue ainsi que leurs commentaires et recommandations afin d'assurer une utilisation responsable de l'apprentissage automatique. L'expertise de ce comité se situe au niveau des thèmes « application rigoureuse » et « méthodes éprouvées », présents dans le Cadre. Le rôle de ce comité est donc de revoir la méthodologie utilisée et de la remettre en question si nécessaire, notamment en identifiant certaines lacunes ou problèmes potentiels et en proposant des améliorations ou des corrections. Ultiment, ce comité recommandera (ou non) la mise en œuvre de la méthodologie proposée dans le contexte de production de statistiques officielles.

## 4. Conclusion

Cet article décrit les procédures actuelles qui ont été mises en place concernant l'utilisation responsable de l'apprentissage automatique à Statistique Canada. De nouvelles sources de données et méthodes d'apprentissage automatique émergent pratiquement chaque jour. Afin de demeurer pertinent, le Cadre présenté dans cet article devra être fréquemment adapté et révisé pour tenir compte des nouveaux enjeux d'éthique et de qualité. Des travaux de recherche sont actuellement en cours avec l'objectif d'améliorer le Cadre et de le rendre encore plus pertinent.

Un de ces projets de recherche est relié à l'interprétabilité et l'explicabilité des modèles d'apprentissage automatique utilisés. Ce domaine a suscité beaucoup d'intérêt dernièrement et fait l'objet de plusieurs articles. L'idée est de faire une revue de littérature dans ce domaine pour bien définir ces deux concepts et être au courant des méthodes utilisées pour aider à expliquer et interpréter des modèles complexes. Certaines méthodes d'interprétation seront ensuite utilisées sur des algorithmes développés à Statistique Canada pour déterminer la meilleure approche à prendre. Il serait aussi important de savoir si des travaux de recherche plus théoriques sont souhaitables et si de nouvelles méthodes d'interprétation doivent être développées dans l'éventualité où l'explicabilité/interprétabilité de certains modèles est jugée insatisfaisante même après l'utilisation de méthodes aidant à l'interprétation proposées dans la littérature. Un autre projet de recherche concerne l'étude des solutions d'apprentissage automatique prêtes à utiliser qu'offrent certains fournisseurs. Ces outils permettent l'utilisation de l'apprentissage automatique par des personnes ayant peu ou pas de connaissances dans ce domaine. Ils peuvent aussi aider des scientifiques des données à développer un algorithme d'apprentissage automatique plus rapidement. Un des objectifs de ce projet est de comparer les résultats de certaines de ces solutions génériques des fournisseurs avec des algorithmes spécifiques, développés à l'interne. Ainsi, on pourra évaluer et identifier les bénéfices et les risques associés à ces solutions afin de prévenir les utilisations non souhaitables de ces outils.

En conclusion, l'apprentissage automatique est de plus en plus considéré et utilisé afin d'améliorer et moderniser les différents programmes statistiques de Statistique Canada. Il est donc nécessaire de s'assurer que l'on fait une utilisation responsable de ces méthodes afin de préserver la confiance des Canadiens et de continuer à produire des statistiques plus détaillées, plus rapidement et de bonne qualité. C'est pour cette raison que le Cadre a été développé et qu'un processus de revue a été instauré. Une formation visant à informer les employés sur ce qu'est l'utilisation responsable de l'apprentissage automatique à Statistique Canada est actuellement en développement. Cette dernière aidera à faire la promotion du Cadre et à en favoriser l'utilisation.

## Bibliographie

Gouvernement du Canada (2019), « Directive sur la prise de décisions automatisée », URL : <https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592>

Réseau de la science des données pour la fonction publique fédérale (2021a), « Utilisation responsable de l'apprentissage automatique à Statistique Canada », URL : <https://www.statcan.gc.ca/fr/science-donnees/reseau/apprentissage-automatique>

Réseau de la science des données pour la fonction publique fédérale (2021b), « Utilisation responsable des systèmes décisionnels automatisés du gouvernement fédéral », URL : <https://www.statcan.gc.ca/fr/science-donnees/reseau/systemes-automatisees>

Statistique Canada (2020), « Politique sur la protection des renseignements personnels et la confidentialité », URL : [https://icn-rci.statcan.ca/31/31a/31a\\_002-fra.html](https://icn-rci.statcan.ca/31/31a/31a_002-fra.html)

Statistique Canada (2021), « Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada », disponible sur le réseau interne de Statistique Canada à l'adresse suivante : <https://www150.statcan.gc.ca/n1/pub/89-20-0006/892000062021001-fra.htm>