

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

**Donner le pouvoir aux analystes de tenir
compte de l'éthique de leur travail : une
étude de cas du cadre d'éthique des
données de la United Kingdom Statistics
Authority du Royaume-Uni**

par Natalie Shlomo

Date de diffusion : le 22 octobre 2021



Statistique
Canada

Statistics
Canada

Canada

Donner le pouvoir aux analystes de tenir compte de l'éthique de leur travail : une étude de cas du cadre d'éthique des données de la United Kingdom Statistics Authority du Royaume-Uni

Simon Whitworth¹

Résumé

Notre société devient de plus en plus numérique et offre de multiples possibilités de maximiser notre utilisation des données dans l'intérêt du public, par un éventail de sources, de types de données et de technologies pour nous permettre de mieux informer le public sur les questions sociales et économiques et de contribuer à l'élaboration et à l'évaluation efficaces des politiques gouvernementales. Un facteur important pour concrétiser le potentiel d'utilisation des données aux fins de recherche et de statistiques d'intérêt public est la garantie de l'utilisation des données de manière éthiquement appropriée. Au début de l'année, la United Kingdom Statistics Authority a lancé le Centre for Applied Data Ethics afin de fournir des services, des conseils, des formations et des orientations concernant l'éthique des données à la communauté des analystes du Royaume-Uni. Le Centre a développé un cadre et un portefeuille de services pour donner les moyens aux analystes de tenir compte de l'éthique de leur recherche rapidement et facilement, dès la phase de conception de la recherche, promouvant ainsi une culture de l'éthique dès la conception. Le présent article donne un aperçu de ce cadre, des services de soutien aux utilisateurs qui l'accompagnent et de l'incidence de ce travail.

Mots clés : éthique des données; données; recherche et statistiques

1. Introduction

La UK Statistics Authority (UKSA) est un organisme indépendant du gouvernement. L'UKSA a pour objectif légal de promouvoir et de protéger la production et la publication de statistiques officielles qui « servent le bien public », ce qui comprend l'information du public sur les questions sociales et économiques, la contribution à l'élaboration et à l'évaluation des politiques publiques, la réglementation de la qualité et la contestation publique du mauvais usage des statistiques. Comme il est énoncé dans la dernière stratégie quinquennale de l'UKSA, [Statistics for the Public Good](#), la mission de l'Autorité est de produire « des données et des analyses de grande qualité pour informer le Royaume-Uni, améliorer la vie de la population et bâtir l'avenir » (UK Statistics Authority, 2020). Les préoccupations au sujet de l'éthique des données sont souvent citées comme des obstacles à l'exploitation du pouvoir des données à des fins d'analyse. En tant que producteur et utilisateur responsable de données, l'UKSA relève ce défi en s'attaquant de front aux problèmes éthiques et en lançant le [UK Statistics Authority's Centre for Applied Data Ethics](#) (le Centre) cette année pour contribuer à donner à la communauté des analystes les moyens de cerner et régler les problèmes éthiques de la collecte et de l'utilisation de données à des fins d'analyse. Il est essentiel de donner aux analystes le pouvoir de cerner et de traiter efficacement et robustement les questions d'éthique dans le cadre de leur travail, pour leur permettre de collecter et d'utiliser les données d'une façon conforme à l'éthique, et de répondre ainsi aux préoccupations relatives à l'éthique des données.

Tous les travaux du Centre sont orientés par un comité consultatif indépendant composé d'experts du gouvernement et du milieu universitaire, notamment des représentants de l'Ada Lovelace Institute, du Central Digital and Data Office, du Centre for Data Ethics and Innovation, de l'Office for National Statistics, d'Administrative Data Research UK et de la section de l'éthique et de la gouvernance des données de la Royal Statistical Society. Le Comité fournit des conseils indépendants très précieux concernant l'orientation stratégique, les résultats et l'incidence des travaux du Centre.

¹ Simon Whitworth, UK Statistics Authority, 1 Drummond Gate, Londres, SW1V 2QQ, Royaume-Uni.

2. Les principes de l'éthique des données

Ce travail s'appuie sur les efforts déployés par le [National Statistician's Data Ethics Advisory Committee \(NSDEC\)](#) (Comité consultatif sur l'éthique des données du statisticien national), qui a été établi en 2016 afin de fournir des garanties et des avis indépendants en matière d'éthique au statisticien national. Le NSDEC, composé de spécialistes de la recherche, de l'éthique des données, du droit des données et d'un membre profane pouvant donner un point de vue de non-spécialiste, examine de façon transparente les projets de recherche de toute la communauté des analystes. Le comité a étudié un éventail d'utilisations novatrices des données pour éclairer des politiques gouvernementales importantes.

À ses débuts, le NSDEC a contribué à la rédaction des six principes éthiques de l'UKSA, qui sont énumérés ci-dessous. Tous nos services d'éthique évaluent les projets de recherche en fonction de ces principes. Ces derniers sont examinés régulièrement de façon à ce qu'ils restent conformes aux normes et aux développements les plus récents du paysage des données en constante évolution. Les voici.

- L'utilisation des données apporte des avantages évidents aux utilisateurs et sert le bien public.
- L'identité du sujet des données (qu'il s'agisse d'une personne ou d'une organisation) est protégée, la confidentialité des renseignements est préservée et les renseignements protégés, et la question du consentement est examinée adéquatement.
- Les risques et les limites des nouvelles technologies sont pris en compte et la surveillance humaine suffit à ce que les méthodes employées soient conformes aux normes reconnues d'intégrité et de qualité.
- Les données utilisées et les méthodes employées sont conformes aux exigences juridiques du Royaume-Uni, telles que la *Data Protection Legislation* (Loi sur la protection des données), la *Human Rights Act 1998* (Loi de 1998 sur les droits de la personne), la *Statistics and Registration Service Act 2007* (Loi de 2007 sur le Service de statistique et d'enregistrement), et le devoir de confidentialité de la common law.
- Les opinions du public sont prises en compte à la lumière des données utilisées et des avantages perçus de la recherche.
- L'accès, l'utilisation et le partage de données sont transparents et sont communiqués de manière claire et accessible au public.

3. Aller au-delà des principes éthiques

L'élaboration de principes éthiques de haut niveau est une bonne chose, mais comment nous assurer qu'ils sont employés par l'ensemble de la communauté analytique et qu'ils ont une incidence sur le terrain? En effet, à quoi bon définir des principes éthiques s'ils ne sont pas utilisés? Nous avons relevé ce défi et sommes passés de la parole aux actes en élaborant un outil d'autoévaluation éthique, le [UK Statistics Authority's ethics self-assessment tool](#). L'outil donne aux analystes les moyens de cerner et d'atténuer les risques éthiques de leurs projets. La cote de risque éthique générée par l'outil permet également de déterminer les projets qui sont considérés comme plus litigieux sur le plan éthique afin qu'ils puissent ensuite être examinés par les experts indépendants du NSDEC.

Les analystes ne sont pas livrés à eux-mêmes face à l'outil. Une équipe d'experts propose aux utilisateurs orientation, formation, soutien et surveillance. Les rétroactions des utilisateurs indiquent que cela facilite l'utilisation de l'outil, ce qui a contribué à sa popularité dans la communauté des analystes. Au moment de la rédaction, il avait été utilisé plus de 470 fois par des analystes du gouvernement, du milieu universitaire, du secteur commercial et d'organismes de bienfaisance. Pour que la recherche se fasse en temps opportun, il est important que ce processus soit efficace. Or, actuellement, le délai moyen pour obtenir d'un membre de l'équipe des conseils et du soutien à propos d'un outil d'autoévaluation rempli est de deux jours.

Parce que nous avons travaillé avec des analystes sur plus de 470 autoévaluations de l'éthique, nous sommes très bien placés pour déterminer les domaines précis sur lesquels des conseils éthiques transversaux seraient utiles à

la communauté des analystes. C'est pourquoi nous avons lancé un programme de recherche visant à publier des conseils pratiques à l'intention des utilisateurs sur une série de sujets relatifs à l'éthique appliquée des données. Jusqu'à maintenant, le programme a publié des documents d'orientation sur :

- [les considérations éthiques dans l'utilisation des données géospatiales aux fins de recherche et de statistiques;](#)
- [la prise en compte du bien public dans les recherches et les statistiques : directives en matière d'éthique;](#)
- [la prise en compte des points de vue et de la participation du public concernant l'utilisation des données à des fins de recherche et de statistiques;](#)
- [les considérations éthiques dans l'utilisation de l'apprentissage automatique aux fins de recherche et de statistiques;](#)
- [une liste de vérification de l'éthique de haut niveau aux fins d'utilisation de données de tiers.](#)

Pour ceux qui manquent de temps, ces documents d'orientation comprennent une liste de vérification de l'éthique qui résume les principaux points abordés dans le document, que les analystes peuvent parcourir rapidement. Bien que ces listes de vérification puissent être utilisées comme des résumés autonomes, elles sont également conçues pour fournir un soutien supplémentaire aux chercheurs qui remplissent un outil d'autoévaluation de l'éthique pour un projet de recherche portant sur ces sujets en particulier. Ces documents d'orientation sont publiés sous forme de premières ébauches ouvertes, ce qui permet de recevoir les commentaires et les rétroactions de la communauté des utilisateurs, qui éclaireront les itérations futures et l'élaboration continue des documents.

4. Présence internationale

Un objectif important du Centre est d'être « des chefs de file reconnus en matière d'application pratique de l'éthique des données dans les statistiques et la recherche » (Centre for Applied Data Ethics Strategy, 2021). Bien qu'en activité depuis seulement huit mois, nous avons déjà un profil international notable. Nous avons présenté notre travail à plusieurs conférences internationales, notamment au Stats Café on Big Data Governance and Data for Policy (Café statistique sur la gouvernance des mégadonnées et les données pour les politiques) de 2021 de la Commission économique et sociale des Nations Unies pour l'Asie et le Pacifique. Nous sommes également représentés au sein de l'équipe de travail sur le leadership éthique de la CEE-ONU, qui étudie actuellement les questions d'éthique professionnelle et d'éthique des données au niveau organisationnel et qui élabore des ressources pour soutenir les organismes nationaux de statistique (ONS) dans ces domaines, et nous avons dirigé une collaboration internationale sur la production de directives en matière d'éthique dans l'utilisation de l'apprentissage automatique dans les statistiques officielles. Il s'agit d'un volet important du travail international plus général sur l'utilisation de l'apprentissage automatique dans les statistiques officielles qui est coordonné par la Commission économique des Nations Unies pour l'Europe.

Le Centre dirige également un programme de travail visant à partager les apprentissages et les expériences entre ONS afin d'élaborer des solutions concertées aux défis de l'éthique des données. Ainsi, nous avons mené cette année une enquête internationale auprès des organismes nationaux de statistique visant à cerner les défis éthiques communs auxquels ils font face et les façons dont ils y répondent. L'enquête a défini trois sujets, qui pourraient faire l'objet d'une analyse plus approfondie et sont considérés comme des problèmes d'éthique des données communs aux ONS : les politiques en matière d'éthique des données, l'optimisation éthique de l'utilité des données pour la recherche et les statistiques, et la distinction entre les utilisations statistiques et opérationnelles des données. Nous envisageons de diriger des collaborations internationales sur ces trois domaines importants.

5. Synthèse

La réflexion sur l'éthique des données évolue rapidement. L'une des premières publications du Centre for Applied Data Ethics était [un examen du paysage de l'éthique des données au Royaume-Uni](#) (UK Statistics Authority, 2021), qui présentait le volume important de travaux novateurs réalisés sur la question de l'éthique des données au Royaume-Uni par différents organismes. En tant qu'utilisateurs prolifiques de données, souvent de façon toujours plus novatrice et sophistiquée, les analystes de données de l'ensemble du système statistique ont la possibilité de contribuer significativement à la réflexion sur l'éthique des données en l'enrichissant. Le travail du Centre jouera un rôle crucial en donnant à la communauté des analystes le pouvoir d'apporter cette contribution importante.

Notre expérience a démontré que si les analystes ont les moyens de tenir compte de l'éthique des données et de régler les problèmes d'éthique au début de la phase de conception de leur recherche, la plupart des inquiétudes au sujet de l'éthique des données peuvent être apaisées et ne devraient donc plus constituer un obstacle notable à l'utilisation des données dans des analyses radicales et ambitieuses. Le fait de proposer des services efficaces en matière d'éthique des données, une orientation et un soutien aux utilisateurs pour permettre aux analystes de l'ensemble du système statistique de tenir compte de l'éthique de leur travail à l'étape de la conception de la recherche peut donc jouer un rôle important, en aidant le système statistique à produire des recherches et des statistiques de grande qualité pour le bien public.

Bibliographie

UK Statistics Authority, 2020, Statistics for the public good, <https://uksa.statisticsauthority.gov.uk/statistics-for-the-public-good/>

UK Statistics Authority, 2021 Centre for Applied Data Ethics Strategy, <https://uksa.statisticsauthority.gov.uk/publication/centre-for-applied-data-ethics-strategy-enabling-ethically-appropriate-research-and-statistics-for-the-public-good/>

UK Statistics Authority, 2021, Ethical considerations in the use of geospatial data for research and statistics, <https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-geospatial-data-for-research-and-statistics/pages/1/>

UK Statistics Authority, 2021, Considering public good in research and statistics: Ethics guidance, <https://uksa.statisticsauthority.gov.uk/publication/considering-public-good-in-research-and-statistics-ethics-guidance/>

UK Statistics Authority, 2021, Considering public views and engagement regarding the use of data for research and statistics, <https://uksa.statisticsauthority.gov.uk/publication/considering-public-views-and-engagement-regarding-the-use-of-data-for-research-and-statistics/pages/1/>

UK Statistics Authority, 2021, Ethical considerations in the use of machine learning for research and statistics, <https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-machine-learning-for-research-and-statistics/>

UK Statistics Authority, 2021, Using data from third parties for research and statistics: High level ethics checklist, <https://uksa.statisticsauthority.gov.uk/publication/using-data-from-third-parties-for-research-and-statistics-high-level-ethics-checklist/>

UK Statistics Authority, 2021, Identifying gaps, opportunities and priorities in the applied data ethics guidance landscape, <https://uksa.statisticsauthority.gov.uk/publication/identifying-gaps-opportunities-and-priorities-in-the-applied-data-ethics-guidance-landscape/>