



COVID-19 : Les défis reliés à l'ouverture et au partage des données sur la COVID-19

MEGAN WYLIE, ANALYSTE DE RECHERCHE, LABORATOIRE DE RECHERCHE APPLIQUÉE

NABEEL AHMED, AGENT DE PROGRAMME, SERVICE DE CONSULTATION

JEAN-NOÉ LANDRY, DIRECTEUR EXÉCUTIF

SARAH GAGNON-TURCOTTE, DIRECTRICE, LABORATOIRE DE RECHERCHE APPLIQUÉE

11 AOÛT 2020

Les données sont devenues un centre d'attention capital aux yeux des gouvernements et du public durant la pandémie de la COVID-19. En effet, on les a reconnues à titre de facteur crucial dans la lutte contre la propagation de la maladie par le biais d'initiatives telles que les applications de [dépistage des contacts](#) ou le concept d'« [aplatissement de la courbe](#) ». Tout comme la pandémie a eu des répercussions sans précédent sur nos vies, de même a-t-elle fait monter en flèche et de manière inattendue la popularité des données.

Le problème est que les données sur la COVID-19 ne sont pas parfaites. À l'heure actuelle, il existe des problèmes bien documentés autour de la collecte et de la communication des données, tels que des [retards](#) dans la collecte et le partage, des communications [incohérentes](#) entre lieux, ou encore des [changements dans les méthodes](#). De plus, les Canadiens n'ont pas accès aux données dont ils ont besoin pour comprendre la propagation et l'incidence de la COVID-19. Les demandes de renseignements sur la démographie des cas au Canada, [en particulier les variables relatives à la race ou à l'ethnicité](#), n'ont pas encore trouvé entière satisfaction, et les gens continuent [d'exprimer le souhait d'avoir un meilleur accès aux données](#) sur la COVID-19 pour mieux éclairer la prise de décisions.

« Chaque chiffre se lie à une histoire humaine : un voisin, une entreprise familiale... les chiffres nous permettent de comprendre toute une série de situations, et permettent aux décideurs politiques et à tous les Canadiens d'y faire face. »

- M. Anil Arora, conférencier principal, Statisticien en chef du Canada

Étant le principal OBNL au Canada spécialisé dans les données ouvertes depuis 2011, Nord Ouvert connaît bien ces questions. Dans un récent [document de politique](#), nous avons résumé les besoins des praticiens des données durant la pandémie, et avons également recueilli des [cas d'utilisation](#) illustrant la façon dont les gouvernements emboîtent le pas et diffusent de nouvelles données pour améliorer l'intervention et le rétablissement au cours de la pandémie. À l'heure actuelle, tandis que [les gouvernements du monde](#) entier réagissent à la COVID-19 en prenant des décisions rapidement, il est temps plus que jamais de s'appuyer sur les principes de gouvernement ouvert pour alimenter à la fois la confiance du public et la résilience à l'égard des chocs futurs.

LES TROIS PRINCIPAUX DÉFIS CONFRONTANT LES UTILISATEURS DE DONNÉES

Afin de répondre aux besoins des utilisateurs de données, Nord Ouvert et le Conseil canadien des normes ont réuni plus de 100 intervenants le 2 juillet 2020 pour discuter de l'utilisation des données durant la pandémie. Absorbés par l'urgence de la pandémie et la nécessité d'y répondre rapidement, les gouvernements n'ont pas consulté le public au sujet de la transparence de l'information et des données, ni pour savoir si les données disponibles répondaient aux attentes des intervenants. La consultation de Nord Ouvert s'est concentrée sur les données ouvertes et partagées concernant la propagation de la maladie, l'incidence sur les communautés et la réponse des gouvernements, en faisant fond sur [la Charte internationale d'ouverture des données](#) et sur l'appel de [l'Unité des données et du gouvernement numériques de l'Organisation de coopération et de développement économiques \(OCDE\)](#) à [réunir les](#)

[intervenants](#) afin de s'informer sur les besoins locaux en matière de données sur la COVID-19.

Près de la moitié des participants œuvraient dans le secteur public, et plus d'un tiers d'entre eux se sont identifiés comme appartenant soit à la société civile, soit au monde universitaire. Il s'agit de groupes de personnes familières avec les données ouvertes et partagées au Canada, notamment les défis confrontant les producteurs et utilisateurs de données ainsi que la réponse canadienne en matière de données sur la COVID-19.

La consultation *Données ouvertes sur la COVID-19* s'est déroulée en ligne, et a donné le coup d'envoi avec le discours d'ouverture du conférencier principal et statisticien en chef du Canada, monsieur Anil Arora, suivi de celui de quatre autres conférenciers invités. Ensuite, des discussions en petits groupes ont permis aux participants de mettre en lumière les ensembles de données disponibles et désirées. Les participants inscrits à la consultation ont également eu l'occasion de remplir un sondage préalable à l'événement servant à déterminer leurs besoins en matière de données (produisant 40 réponses). Les constats présentés dans le présent rapport résument les apports des participants lors des discussions en petits groupes ainsi que les réponses au sondage. Pour obtenir de plus amples détails sur la consultation *Données ouvertes sur la COVID-19 – Mobiliser les Canadiens autour de leurs besoins en matière de données*, veuillez lire notre [article de cyberchronique](#).

Défi N° 1 : Les participants souhaitent qu'on ouvre davantage de données et veulent savoir quand elles sont disponibles.

Il y a une demande évidente pour une plus grande disponibilité des données et, selon les participants, encore beaucoup de données qu'on pourrait ouvrir et partager. En témoignent plus de 200 contributions que les participants ont ajoutées à notre tableau collaboratif virtuel durant la consultation (voir tableau 1).

Les demandes d'ensembles de données variaient considérablement. Les participants ont demandé de plus amples données sur les conditions de travail à distance, sur les points de vue des gens au sujet de la façon dont nous devrions concevoir nos espaces urbains, et sur les lieux de dépistage. Dans notre analyse, nous avons trouvé

44 demandes uniques d'ensembles de données (liste disponible en annexe A). Le tableau 2 présente un échantillon des ensembles de données ainsi que certaines des questions restantes auxquelles les participants voudraient répondre avec les données ouvertes souhaitées.

Les données démographiques relatives à tous les aspects de la propagation et des retentissements de la COVID-19 se sont avérées les plus fréquemment citées. Ces données, disponibles dans d'autres pays, [comme les États-Unis](#), ont permis au public d'évaluer plus effectivement l'incidence de la COVID-19 sur certaines communautés marginalisées. De plus, d'autres données sociales et démographiques se sont aussi révélées très convoitées, à savoir les déplacements des personnes, leur logement et leur situation professionnelle.

Tableau 1 : Les participants ont affiché plus de 200 contributions sur notre tableau collaboratif

	Propagation de la maladie	Impact communautaire	Action gouvernementale	Total
Disponible	37	20	22	79
Non disponible	55	51	32	138
Total	92	71	54	217

**Ce tableau compte les contributions uniques de feuillets-mémos. Ce tableau ne représente pas des ensembles de données uniques, car les participants ont demandé les mêmes ensembles de données à plusieurs reprises. En outre, nous avons constaté que de nombreuses contributions chevauchaient plusieurs questions à examiner, ce qui indique la possibilité de disposer de quelques ensembles de données bien choisis qui pourraient répondre aux besoins de plusieurs utilisateurs de données.*

Tableau 2 : Un échantillon du large éventail d'ensembles de données dont les participants ont demandé l'ouverture

	Propagation de la maladie	Impact communautaire	Action gouvernementale
Sous-thèmes	- Données révisées sur les cas : faux positifs et négatifs - Dépistage de la maladie chez les personnes asymptomatiques - Dépistage des contacts - Facteurs de vulnérabilité - Interactions entre les personnes - Utilisation et efficacité des masques	- Race/ethnicité - État de personne handicapée - Anxiété des gens - Transport public - Situation de logement (par exemple, sans-abri) - Type de logement (par exemple, logement social) - Domaine de travail (par exemple, le tourisme)	- Bénéficiaire d'un ou plusieurs programmes d'aide du gouvernement (par exemple, PCU*) - Travailleurs essentiels et exposition au travail (par exemple, les travailleurs de l'économie à la demande) - Travail à distance - Approvisionnement (par exemple, type d'entreprise) - Dépenses pour les communautés indigènes
Questions restantes et (ou) cas d'utilisation	- Quels sont les effets à long terme de la COVID-19 sur la santé des personnes? - Les gens se font-ils dépister plusieurs fois? - Quel est le rôle de la culture dans la propagation des maladies? - En quoi le transport pose-t-il un obstacle aux tests? - Quelles sont les répercussions du surpeuplement des logements sur la propagation et l'incidence de la maladie? - Où les gens se réunissent-ils ou interagissent-ils?	- Quelle est l'incidence de la COVID-19 sur les groupes marginalisés? - Comment cette incidence se compare-t-elle entre les différentes maisons de soins infirmiers ou communautés? - Quelle est l'incidence de la présence des enfants à la maison? - Comment l'utilisation des transports publics a-t-elle évolué?	Qui embauche? Où peut-on trouver un emploi?

*Prestation canadienne d'urgence. La liste complète des demandes de données se trouve à l'annexe A.

Défi N° 2 : La pandémie exacerbe des obstacles de longue date à l'utilisation des données.

Les participants ont souligné la nécessité d'une meilleure interopérabilité, de données diffusées à des résolutions spatiales plus faibles, et de données de plus haute qualité. Le tableau 3 présente leurs besoins.

L'interopérabilité se définit comme la capacité à intégrer plusieurs ensembles de données, permettant aux utilisateurs de grouper des données. La communication de données représentant des zones localisées, ou à *des résolutions spatiales plus faibles*, aide à comprendre la situation locale et revêt une plus grande utilité pour la recherche. *L'inconstance de la qualité des données* rend plus difficile pour les participants de faire confiance aux données publiées aux fins de recherche et de prise de décisions.

« [Nous] devons nous mettre d'accord sur des identificateurs communs afin de permettre de grouper des données entre les ensembles de données pour une même personne. »

- **Monsieur Mark Leggott**, conférencier invité, Directeur exécutif de Données de recherche Canada

Tableau 3 : Obstacles de longue date à l'utilisation des données

Interopérabilité	Qualité	Résolution spatiale
• Identificateurs et sémantique communs pour le groupement des données • Catégorisation commune des variables, telles que les tranches d'âge, entre les autorités compétentes • Méthodes et formats cohérents entre les ensembles de données et les autorités compétentes	• Transparence des changements de la qualité des données (par exemple, les faux positifs) • Données opportunes qui répondent à des questions pertinentes • Métadonnées pour évaluer la qualité des données elles-mêmes	• Diffuser des données à des paliers géographiques inférieurs, en particulier pour les zones rurales et à l'échelon des quartiers • Intégrer à l'échelle nationale les unités géographiques recevant la diffusion des données, en vue d'une couverture cohérente • Protéger la vie privée des personnes que les données représentent

La pandémie a mis à découvert l'urgence d'une gouvernance générale des données qui relève ces défis et facilite l'utilisation des données entre les autorités compétentes.

La conférencière Bonnie Healy a évoqué un exemple important de cette situation. La Loi canadienne sur la santé ne mentionne pas les Premières nations; cela rend difficile la création de systèmes de partage de données interopérables entre les Premières nations et d'autres fournisseurs de données sur la santé au Canada.

« Lorsque nous examinons les systèmes de données, nous devons reconnaître les systèmes de santé des Premières nations comme des acteurs à part entière ».

- Madame Bonnie Healy, conférencière invitée, Directrice de la santé à la Confédération des Pieds-Noirs

Défi N° 3 : C'est déroutant d'accéder aux données, surtout au Canada.

Dans de nombreux cas, les intervenants se sont dits confus à propos de la nature des données disponibles et de la manière d'y accéder. Pour atténuer cette confusion, les participants ont suggéré ce qui suit :

- les données devraient être accessibles depuis une source unique, comme [le portail des données ouvertes du Canada](#);
- il devrait exister une carte de l'écosystème indiquant les points où les données sont déjà disponibles;
- les fournisseurs de données devraient lier au préalable les données pour faciliter leur utilisation; et
- les fournisseurs de données devraient organiser les données dans un format qui facilite leur utilisation à des fins d'analyse.

« [Il y a] tant d'information qu'on peine à s'y retrouver. »
« [Il est] difficile d'identifier rapidement les données déjà
accessibles. »

- **Participants**

Tout en reconnaissant les défis inhérents à la coordination de l'accès aux données entre plusieurs gouvernements, les fournisseurs de données doivent communiquer clairement les données, où y accéder, et de la part de qui au Canada elles proviennent.

PROCHAINES ÉTAPES

La consultation *Données ouvertes sur la COVID-19* de Nord Ouvert a mis en évidence trois défis en matière d'accès aux données et de leur utilisation, en passant de la discussion à la mise en lumière de domaines d'action clés. Il est crucial que les gouvernements fournissent ces données et procurent aux communautés la possibilité de mieux comprendre les causes et les retentissements de la COVID-19.

« Il s'agit d'un dialogue avec les Canadiens ... cette conversation doit se poursuivre. »

- **Monsieur Anil Arora**, conférencier principal, Statisticien en chef du
Canada

Nord Ouvert poursuivra cette conversation dans [le cadre de notre projet récemment financé](#) avec [l'Observatoire international sur les impacts sociétaux de l'intelligence artificielle et de la technologie numérique](#) (OBVIA), qui explorera l'application des mécanismes de gouvernance des données à la recherche en santé et à l'organisation sanitaire. Nous nous réjouissons à l'avance de poursuivre ces discussions avec nos partenaires dans les mois à venir par voie de nos travaux de recherche et de

renforcement des capacités. Ces efforts appuient un accès élargi à des données de haute qualité, opportunes et normalisées qui protègent la santé publique et améliorent la coordination des réponses gouvernementales par le biais de pratiques efficaces de gouvernance des données partagées au Canada.

Nord Ouvert est reconnaissante envers les participants qui ont rempli notre sondage et assisté à notre consultation. Vous nous avez fourni des observations, idées et perspectives précieuses reprises ici et dans notre annexe.

À PROPOS DE NORD OUVERT

Nous travaillons avec divers intervenants communautaires et administrations publiques branchés et innovateurs afin de les aider à utiliser les données et les technologies de manière efficace, éthique et collaborative pour résoudre des problèmes complexes. Nous soutenons des réseaux mondiaux d'égal à égal destinés au partage de connaissances qui améliore les pratiques de gouvernance intelligentes et ouvertes pour les citoyens du Canada et du monde entier.

[Abonnez-vous à notre infolettre](#)

Pour accéder à nos plus récentes nouvelles et ressources!

ANNEXE A

Les intervenants ont également évoqué divers autres besoins en matière de données que Nord Ouvert n'a pas pu saisir dans notre document de consultation. Nous énumérons ici les données brutes recueillies lors de notre événement.

Domaines où les intervenants souhaitent avoir accès à davantage de données

1. La propagation des maladies

- Données révisées sur les cas : faux positifs et négatifs
- Dépistage de la maladie chez les personnes asymptomatiques
- Dépistage des contacts
- Facteurs de vulnérabilité
- Interactions entre les personnes
- Utilisation et efficacité des masques
- Encombrement et utilisation des installations
- Immunisation
- Besoins en espace hospitalier
- Temps d'attente des patients
- Emplacements des pharmacies
- Facturation des frais
 - Régie de l'assurance maladie du Québec (RAMQ) ou autres organismes gouvernementaux
 - Installations
 - Prestataires de soins de santé (par exemple, les médecins)
- Renseignements sur les établissements de soins de santé
 - Noms
 - Services offerts
 - Nombre de lits
 - Adresse
 - Coordonnées

2. *Répercussions communautaires*

- Race/ethnicité
- État de personne handicapée
- Anxiété des gens
- Transport public
- Situation de logement (par exemple, sans-abri)
- Type de logement (par exemple, logement social)
- Domaine d'activité (par exemple, le tourisme)
- Retentissements sur les lieux de travail
 - Productivité
- Les priorités des citoyens
 - Que valorisent les gens?
 - Quel aspect veulent-ils que prenne la société à la suite de la pandémie?
 - Quels changements devraient demeurer en permanence?
 - Qu'est-ce qui devrait redevenir comme auparavant?
- L'utilisation des espaces par les citoyens : intérieur et extérieur
- Comportements personnels, perceptions, et évaluation des risques
 - Pour le retour au travail
 - Perception de l'efficacité des masques
- Indicateurs de bien-être communautaire
- Organisations d'aide alimentaire
- Consommation de substances
- Accès à Internet et aux réseaux à large bande

3. *Action gouvernementale*

- Bénéficiaire d'un ou de plusieurs programmes d'aide gouvernementaux (par exemple, PCU)
- Travailleurs essentiels et exposition au travail (par exemple, les travailleurs de l'économie à la demande)
- Travail à distance
- Approvisionnement (par exemple, type d'entreprise)

- Dépenses pour les communautés indigènes
- Nations indigènes
 - Ressources à la disposition des communautés
 - Données sur la population
 - Conditions préalables au contact
- Aides publiques aux entreprises
- Marchés publics
 - Types d'entreprise
 - Caractéristiques du propriétaire
- Soutien gouvernemental aux communautés indigènes, par nation
- Plans et marches à suivre de réouverture en place
- Toilettes publiques
- Procédures d'hygiène et d'entretien
- Respect de la réglementation
- Applications de dépistage des contacts
 - Où servent-ils?
 - Quelle est leur évaluation des risques?
- Des renseignements clairs sur les modalités d'accès aux soins de santé dans chaque province
- Propriété bénéficiaire

Échelles géographiques demandées par les participants

- Rural
- Municipal
- Les données du recensement à l'échelon :
 - de la zone de diffusion
 - de la division de recensement
- Granularité plus faible
- Fichiers nationaux intégrés

Autres besoins exprimés

- Métadonnées avec autorité et qui sont utilisées
- Mesures normalisées imposées par chaque palier de gouvernement
- Information contextuelle
 - Par exemple, des comparaisons historiques à chaque échelon géographique
- Information d'analyse destinée aux décisions gouvernementales
 - Par exemple, analyse des approches d'autres pays
 - Plans de la deuxième vague
 - Quelles données font l'objet de collecte même si on ne les communique pas (par exemple, l'analyse du dépistage des contacts)
- Coordonnées des responsables de la gestion des données
- Code source pour les applications de dépistage des contacts
- Conception du système et de la base de données
- Données produites par la communauté

Comme ces demandes de données sont assez générales, les gouvernements devront remobiliser les intervenants afin de comprendre leurs variables de données particulières, leur échelle géographique et leurs autres besoins.