

**Rapport *Ce que nous avons entendu* : consultations sur l'exigence de dépôt
de données prévues dans la *Politique des trois organismes sur la gestion
des données de recherche***



Conseil de recherches en
sciences humaines du Canada

Social Sciences and Humanities
Research Council of Canada

Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie, 2025

Numéro de catalogue CR22-136/2025F-PDF

ISBN 978-0-660-77876-1

Rapport *Ce que nous avons entendu* : consultations sur l'exigence de dépôt de données prévues dans la *Politique des trois organismes sur la gestion des données de recherche*

Table des matières

1. Résumé.....1

2. Contexte4

3. Modes de consultation5

4. Ce que nous avons entendu6

 4.1 Clarté de la Politique 6

 4.2 Échéancier d'application 8

 4.3 Capacité et besoins 10

 4.4 Répertoires de données..... 11

 4.5 Curation des données 12

 4.6 Partage des données..... 13

 4.7 Données sensibles..... 14

 4.8 Souveraineté des données autochtones..... 16

 4.9 Mégadonnées 17

 4.10 Suivi et conformité 18

 4.11 Autres points à considérer 19

5. Conclusion21

6. Prochaines étapes22

1. Résumé

Le présent rapport résume la rétroaction reçue par les Instituts de recherche en santé du Canada (IRSC), le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) et le Conseil de recherches en sciences humaines du Canada (CRSH) (les « organismes ») concernant l'application de l'exigence de dépôt de données prévue dans la *Politique des trois organismes sur la gestion des données de recherche* (la « Politique »). Des consultations ont eu lieu entre juillet 2023 et janvier 2025 pour recueillir les divers points de vue des parties prenantes ainsi que de l'information sur l'évolution des politiques et pratiques de gestion des données de recherche au pays et à l'étranger; les défis, les besoins, la capacité et l'état de préparation associés au dépôt de données au Canada; l'échéancier qui serait réaliste pour le dépôt des données. Les organismes ont recouru à différents modes de consultation,

notamment des rencontres en ligne et en personne avec les parties prenantes et les organismes partenaires du milieu de la recherche qui ont donné lieu à plusieurs observations écrites.

Dans l'ensemble, les personnes participantes étaient largement d'accord avec l'exigence de dépôt de données et ont convenu de son importance – et de celle de la Politique elle-même – pour améliorer la gouvernance des données de recherche au Canada. Il est jugé nécessaire que les politiques et pratiques de gestion de données de recherche au pays cadrent avec celles d'autres organismes de financement et partenaires de recherche importants. Selon les personnes participantes, l'exigence de dépôt et la simplification de la diffusion des données au Canada amélioreraient la transparence et la reproductibilité des recherches financées par l'État, en plus de faciliter la réutilisation des données, ce qui renforcerait l'efficacité de la recherche. Pour ce qui est de l'exigence de dépôt de données des trois organismes, les gens du milieu ont donné des idées et parlé de points et de besoins essentiels dans l'écosystème de la recherche canadienne dans son ensemble – ce qui comprend les organismes subventionnaires, les chercheuses et chercheurs, les fournisseurs de services, les établissements (bibliothèques et bureaux de la recherche) et les organismes partenaires. En général, les gens du milieu ne sont pas certains d'être prêts à négocier le virage, mais conviennent de l'importance de déposer et de partager les données de recherche canadiennes.

1. **Clarté de la Politique** : Les personnes participantes ont souligné à maintes reprises la nécessité de clarifier certains aspects de l'exigence, notamment les objectifs de reproductibilité et de réutilisation des données; la formulation des attentes concernant le partage des données, qui se distingue du dépôt; les obligations éthiques, juridiques et commerciales; la définition des termes « dépôt de données », « curation » et « accès approprié »; et l'applicabilité du consentement des personnes participantes et des obligations des chercheuses et chercheurs selon l'*Énoncé de politique des trois conseils : éthique de la recherche avec des êtres humains* ([EPTC2](#)).
2. **Échéancier d'application** : Les personnes participantes ont exprimé des opinions différentes à propos d'un échéancier réaliste d'application de l'exigence. Certaines ont parlé d'un manque de préparation chez les parties prenantes et d'adaptation des services et infrastructures pour répondre à la demande prévue à court terme (deux-trois ans). D'autres ont plaidé pour une application rapide afin d'épouser les normes internationales et de catalyser la préparation des établissements. Il a été question des avantages de l'adoption d'une approche graduelle pour tenir compte des différents degrés de préparation des milieux et établissements de recherche. Dans l'ensemble, la préparation perçue relativement au dépôt des données différait entre les membres, mais on reconnaissait la nécessité d'aller de l'avant avec le dépôt et le partage des données au Canada.
3. **Capacité et besoins** : Il a été soulevé que le Canada jouit d'une base solide – expertise, infrastructures et services – qui facilite la gestion et le dépôt des données de recherche. Cependant, il faudra des ressources supplémentaires pour adapter les services et la capacité à la demande prévue lorsque l'exigence prendra effet. Il faudra également répondre aux défis propres à la gestion et au dépôt des données sensibles, des données autochtones et des mégadonnées. Beaucoup d'établissements, surtout les plus petits, n'ont ni la capacité ni les ressources nécessaires pour gérer les données de recherche, ce qui complique l'application généralisée du dépôt des données. De fortes réserves ont été émises concernant l'évolution des services de gestion des données de recherche, notamment leur curation et la formation des chercheuses et chercheurs, ainsi que sur

l'augmentation de la charge de travail des chercheuses et chercheurs et du personnel de soutien, faute de ressources supplémentaires.

4. **Dépôts de données** : Dans l'ensemble, les personnes participantes estiment que les chercheuses et chercheurs au pays ont les infrastructures numériques suffisantes pour déposer les données non sensibles dans les dépôts de données (contrairement au stockage actif des données). Toutefois, les spécialistes et fournisseurs de services œuvrant dans le domaine de la gestion des données de recherche ont souligné l'importance de pouvoir trouver et consulter les données dans ces dépôts de données et non seulement de les conserver. Certaines personnes participantes ont demandé des conseils pour choisir les bons dépôts de données. D'autres ont plutôt demandé des recommandations, et d'autres encore, une liste des caractéristiques recommandées.
5. **Curation des données** : Il a été question de l'importance d'avoir des normes de curation minimale pour le dépôt des données, normes qui respectent autant que possible les principes [FAIR](#) (repérabilité, accessibilité, interopérabilité et réutilisation). Il a été suggéré que le niveau de curation dépende de la nature et de l'utilisation prévue des données. La curation était généralement perçue comme une responsabilité partagée des chercheuses et chercheurs, des établissements et des responsables de dépôt de données. A été souligné à maintes reprises le besoin crucial de financement supplémentaire pour les établissements, les régions et le système si l'on veut subventionner les services de curation, puisque la curation est indispensable à une gestion des données respectueuse des principes FAIR.
6. **Partage des données** : Les personnes participantes ont largement adhéré à l'idée que les données déposées doivent être « aussi accessibles que possible et aussi inaccessibles que nécessaire ». La sécurité de la recherche et le contrôle des accès sont essentiels si l'on veut protéger les données sensibles, les données autochtones et les données protégées par la propriété intellectuelle. Selon les personnes participantes, les métadonnées ouvertes favorisent la repérabilité à accès restreint.
7. **Données sensibles** : L'absence d'infrastructure nationale et institutionnelle pour déposer les données sensibles et en contrôler l'accès est revenue souvent durant les consultations. Les politiques institutionnelles sur la garde et la curation des données, ainsi que le déroulement du travail associé aux données sensibles, sont mal définies et appliquées dans la plupart des établissements. Les personnes participantes ont parlé de la nécessité de clarifier ce qu'on entend par « données sensibles » et demandé des conseils sur la gestion, la curation et la consultation de ces données en vue de la préparation des options de dépôt de données. Elles ont exprimé des réserves quant à la variation des normes disciplinaires, à la nécessité de consulter les comités d'éthique de la recherche sur le dépôt et la consultation des données éthiques, et au besoin de souplesse pour les données commerciales sensibles.
8. **Souveraineté des données autochtones** : L'exigence de dépôt de données cadre avec le droit à l'autodétermination des peuples autochtones. Elle reconnaît que les données émanant des travaux faits par et avec les Premières Nations, les Métis ou les Inuits dont les territoires ancestraux et traditionnels se trouvent au Canada doivent être gérées selon des principes de gestion élaborés et approuvés par ces communautés, sur la base d'un consentement préalable, libre et éclairé. Les personnes participantes ont fortement recommandé aux organismes de produire des lignes directrices et options pour le dépôt des données avec les détentrices et détenteurs de droits autochtones et les organismes de gouvernance autochtones, tout en reconnaissant que les

gouvernements autochtones sont des entités souveraines qui ont le contrôle sur leurs données.

9. **Mégadonnées** : Les personnes participantes ont recommandé de clarifier la définition de « mégadonnées » et parlé de la nécessité de collaborer avec des sites d'hébergements nationaux pour ce qui touche l'informatique de pointe dans la recherche afin de préciser comment l'exigence de dépôt de données s'appliquerait aux ensembles de données extrêmement volumineux.
10. **Suivi et conformité** : Les personnes participantes ont voulu savoir si la conformité serait évaluée uniquement sur la base du dépôt des données ou aussi sur la réutilisabilité des données, selon les principes FAIR. Plusieurs ont avancé que les organismes devraient s'occuper de la surveillance de la conformité, par exemple en faisant des vérifications aléatoires et en utilisant des identifiants pérennes (PID). Il a été suggéré d'avoir des plans de gestion des données (PGD) comme point de départ à la surveillance et d'intégrer la conformité aux rapports de fin de subvention et aux futures demandes de financement. Une autre suggestion était d'harmoniser la conformité aux politiques sur le libre accès aux publications scientifiques.
11. **Autres points à considérer** : Veiller à ce que l'exigence de dépôt de données n'exacerbe pas les inégalités entre les chercheuses et chercheurs et les établissements de recherche; clarifier la relation entre le dépôt, la curation et la préservation des données; parler des défis associés à la garde à long terme de données à accès restreint; traiter de l'éthique de la réutilisation des données; et parler de la complexité du dépôt de code.

2. Contexte

En mars 2021, les Instituts de recherche en santé du Canada ([IRSC](#)), le Conseil de recherches en sciences naturelles et en génie du Canada ([CRSNG](#)) et le Conseil de recherches en sciences humaines du Canada ([CRSH](#)) (les « organismes ») ont adopté la [Politique des trois organismes sur la gestion des données de recherche](#) (la « Politique ») pour favoriser l'excellence en recherche au pays par des pratiques rigoureuses de gestion des données de recherche et de gestion des autres données.

La Politique énonce trois exigences. La première est que chaque établissement postsecondaire et hôpital de recherche pouvant administrer des fonds des IRSC, du CRSNG ou du CRSH doit avoir une [stratégie de gestion des données de recherche](#). La deuxième est que pour certaines occasions de financement, les personnes candidates doivent soumettre un [plan de gestion des données \(PGD\)](#) à l'organisme approprié au moment de la demande. La troisième est que les titulaires de subvention sont tenus de déposer dans un dépôt numérique les données de recherche, les métadonnées et les codes qui appuient directement les conclusions publiées dans des revues et des préimpressions découlant de la recherche financée par les organismes.

Conformément aux échanges avec le milieu de la recherche et d'autres parties prenantes, les organismes veulent appliquer la politique graduellement, au rythme du perfectionnement des pratiques et de la capacité de gestion des données de recherche au Canada et sur la scène internationale. La dernière exigence prévue dans la Politique, soit le dépôt des données, sera introduite après une évaluation des stratégies des établissements et de l'état de préparation du milieu de la recherche canadien.

Les données d'une étude transversale récente sur les stratégies de gestion des données de recherche de

plus de 200 établissements canadiens se trouve dans ce [document](#)¹ (en anglais). Selon cette analyse, l'adoption de stratégies de gestion des données de recherche dénote un progrès important parmi les établissements pour ce qui est de se conformer à la Politique dans leurs activités de recherche. Cependant, il faut redoubler d'efforts concernant les lacunes de consultation, les contraintes relatives aux ressources, les plans de gestion des données (PGD) et le soutien et les outils de dépôt de données afin que tous les établissements respectent la souveraineté des données autochtones. Plus précisément, beaucoup d'établissements ont fait état de problèmes de capacité pour l'adaptation des infrastructures et services à la hausse potentielle de la demande en gestion de données de recherche lorsque les exigences de PGD et de dépôt de données des trois organismes prendront effet.

De juillet 2023 à janvier 2025, les organismes ont vastement consulté les gens du milieu de la recherche au Canada et à l'étranger pour mieux comprendre :

- les divers points de vue des parties touchées par l'exigence de dépôt des données;
- l'évolution des politiques et pratiques de gestion des données de recherche sur les scènes nationale et internationale;
- les défis, les besoins, la capacité et l'état de préparation associés au dépôt de données au Canada;
- la définition d'un échéancier réaliste pour exiger le dépôt des données.

3. Modes de consultation

Que ce soit en tant qu'organismes ou que participants, les organismes ont pris part à de nombreuses consultations avec des partenaires et des membres du milieu de la recherche, en ligne et en personne. Figuraient notamment au nombre des partenaires la direction et plusieurs groupes de spécialistes de la gestion des données de recherche de l'Alliance de recherche numérique du Canada, des fournisseurs de services canadiens (Boréal, le Dépôt fédéré de données de recherche, SciNet, Calcul Québec), le Bureau de la Conseillère scientifique en chef du Canada, le Réseau canadien des Centres de données de recherche, le Secrétariat sur la conduite responsable de la recherche, des membres des organes consultatifs des trois organismes et du milieu de la recherche autochtone, des organismes subventionnaires nationaux et internationaux, des ministères et organismes à vocation scientifique du gouvernement du Canada, des organismes intergouvernementaux, et des organismes des secteurs sans but lucratif et privé. Quant aux membres du milieu de la recherche, on retrouvait des spécialistes de la gestion des données de recherche, des bibliothécaires, des chercheuses et chercheurs, des administratrices et administrateurs de la recherche, des membres du réseau d'administration de la recherche autochtone (Network of Indigenous Research Administrators), des spécialistes de la sécurité de la recherche et des TI, des membres de comités d'éthique de la recherche, et des conseillères et conseillers juridiques en recherche. En plus de ce qui s'est dit lors des consultations, les organismes ont reçu des observations écrites de plusieurs des parties prenantes. Les déléguées et délégués universitaires des IRSC et les leaders pour le CRSNG et le CRSH ont aussi donné leurs avis et de la rétroaction.

¹ Ripp C., Armond A.C., Albert M., Apavaloae A., Cooper A., Costanzo L., Dearborn D., Franks A., Khan S., Lartey E., MacKinnon K., Moher D., Perpalaj F., Roche D.G., Ramos J., Steeleworthy M., Wang M. et Cobey K. (2024), « 39 Schématisation des stratégies institutionnelles de gestion des 603 données de recherche canadiennes: une étude transversale », prépublication OSF <https://doi.org/10.31219/osf.io/tknkp9>.

Il y a eu des séances et rencontres de consultation plus générales avec des particuliers ou des petits groupes de partenaires et de parties prenantes. Entre juillet et septembre 2023, quatre rencontres virtuelles ont eu lieu avec le Conseil des présidents du réseau d'experts de l'Alliance de recherche numérique du Canada. Deux rencontres en personne se sont déroulées au printemps 2024, une à l'University of Calgary après la Conférence canadienne des administratrices et administrateurs de recherche (46 personnes participantes), l'autre à Halifax pendant la conférence annuelle de l'International Association for Social Science Information Service and Technology (environ 50 personnes participantes). En 2024, trois séances virtuelles d'échange d'information et rencontres générales ont été organisées à la demande des universités. Trois séances virtuelles (deux en anglais et une en français) ont également eu lieu en janvier 2025. Ces activités, largement publicisées, ont attiré plus de 350 personnes de plus de 200 établissements au pays, qui ont d'ailleurs été invitées à s'exprimer durant une période de questions et de discussion, et à répondre à un sondage anonyme en ligne. Les commentaires cités dans le présent rapport ont été anonymisés.

4. Ce que nous avons entendu

Se trouve ci-dessous un aperçu des commentaires et idées reçus par les organismes à propos des principaux enjeux associés à l'exigence de dépôt de données. Même s'il se veut exhaustif, le présent rapport peut ne pas refléter tout ce qui a été soulevé. Cependant, les organismes ont noté tous les commentaires et les intégreront autant que possible.

Dans l'ensemble, les personnes participantes étaient fortement d'accord avec le dépôt des données et qualifiaient l'exigence et la Politique de très bénéfiques à la gouvernance des données de recherche et aux pratiques de gestion des données de recherche au Canada. Dans sa rétroaction, la communauté a fait la lumière sur les idées, besoins et questions essentiels en lien avec l'écosystème national de recherche, notamment les organismes subventionnaires, les chercheuses et chercheurs, les fournisseurs de services, les établissements et les autres organisations qui gèrent des données de recherche au pays. Cette rétroaction vise directement l'application de l'exigence de dépôt de données, y compris les questions dépassant le champ d'application de la Politique, mais demeurant pertinentes pour la communauté élargie et les organismes partenaires.

4.1 Clarté de la Politique

Les personnes participantes ont demandé à maintes reprises la clarification des objectifs et avantages de l'exigence de dépôt de données. La Politique devrait expliquer les différentes fins du dépôt, comme la validation des résultats de recherche (transparence et reproductibilité) pour renforcer la fiabilité des résultats de recherche publiés, et comment le partage de données peut accroître la réutilisation et l'efficacité des travaux de recherche canadiens.

Plusieurs personnes, y compris des spécialistes de la gestion des données de recherche, ont aussi indiqué des passages de la Politique à éclaircir. Par exemple, on mentionne dans une phrase que les titulaires de subvention « sont tenus de déposer dans un dépôt numérique les données de recherche,

les métadonnées et les codes qui appuient directement les conclusions de la recherche publiées dans des revues de même que les préimpressions découlant de la recherche financée par les organismes subventionnaires ». Pourtant, dans un autre passage, on dit ceci : « Les titulaires de subvention ne sont pas tenus de partager leurs données, mais les organismes subventionnaires s'attendent à ce que les chercheurs donnent un accès convenable aux données pour autant que les exigences éthiques, culturelles, juridiques et commerciales le permettent, et conformément aux [principes FAIR](#) et aux normes propres à leurs disciplines. » Ces formulations créent de la confusion quant à la nécessité de publier ou non les données et pourraient mener à une mauvaise interprétation par les chercheuses et chercheurs et les autres parties prenantes, ce qui nuirait aux objectifs de la Politique. Il a été recommandé d'utiliser une affirmation forte, comme celle des National Institutes of Health sur le fait de maximiser la publication des données scientifiques appropriées pour insister sur l'importance de partager les données, sans pour autant écarter les restrictions relatives aux données sensibles, aux données autochtones et aux données liées au partenariats public-privé, par exemple.

Les personnes participantes ont aussi mentionné que l'ambiguïté entre « partage de données » et « dépôt de données » a généré de la confusion dans la communauté. Les organismes devraient clarifier ce qu'ils entendent par « partage » et préciser quelles « exigences éthiques, juridiques et commerciales » s'appliquent à la mise à disposition des données pour réutilisation. Les personnes participantes ont demandé si l'exigence par défaut était le partage public des données lorsque possible ou si l'exigence était respectée simplement par le dépôt des données dans un dépôt numérique, sans forcément les rendre publiques. Il a été proposé de faire la distinction entre l'obligation de partager des données et l'obligation de partager des métadonnées, ainsi que d'indiquer que certaines métadonnées doivent toujours être consultables publiquement, peu importe l'accessibilité des données, afin que les données de recherche soient stockées aux bons endroits et restent trouvables ou découvrables peu importe les exigences de partage.

Beaucoup ont recommandé de définir les mots clés de la Politique, car l'absence de définitions claires nuirait à l'application de l'exigence dans les établissements. Un glossaire exhaustif permettrait à toutes les parties prenantes d'avoir la même compréhension des termes clefs. Les personnes participantes ont souligné que plusieurs termes pouvaient être interprétés de façons différentes, comme « dépôt », « stockage sécuritaire », « préservation » et « curation », tandis que d'autres étaient ambigus, comme « accès approprié », « partage » et « normes disciplinaires ». En outre, il est difficile de savoir comment les normes disciplinaires sont établies, qui les fixe et ce qui se passe si elles contredisent la Politique. Ce manque de clarté crée de l'incertitude et risque d'entraîner des frictions dans l'application de la Politique ; la Politique devrait clarifier comment des différences entre la Politique et certaines normes disciplinaires seront réconciliées.

Il a été mentionné que la Politique devrait préciser le seuil de conformité minimal de l'exigence de dépôt des données. Figuraient parmi les recommandations des exemples et scénarios pour aider les chercheuses et chercheurs à comprendre les répercussions pratiques de l'exigence. Des conseils précis pourraient par ailleurs permettre d'éviter une application inégale entre disciplines et entre établissements.

Certaines personnes participantes ont laissé entendre que les « données de recherche, les métadonnées et les codes » devraient être considérés comme des catégories distinctes, chacune s'accompagnant de ses propres exigences. Cette distinction simplifierait l'application de la Politique en permettant de traiter ces éléments différemment, par exemple en facilitant la publication des métadonnées pour améliorer la trouvabilité, même si les données ou les codes sont soumis à des restrictions.

Il a été dit que la Politique devrait traiter du lien entre l'exigence de dépôt de données et l'Énoncé de politique des trois conseils : éthique de la recherche avec des êtres humains (EPTC2). Plus précisément, il faudrait mentionner comment l'exigence s'inscrit dans le nouvel article de l'EPTC2 sur le consentement général, soit le fait que les personnes participantes doivent consentir de manière distincte à ce que les données soient déposées soit à des fins de publication des résultats, soit afin de faciliter d'autres recherches. La Politique devrait préciser les conditions et les manières dont les chercheuses et chercheurs qui déposent des données destinées à accompagner des résultats publiés (consentement précis) restreindront le partage et la réutilisation des données par les gens n'ayant pas donné leur consentement à l'utilisation des données pour de futures recherches (consentement général).

Les personnes qui font des recherches sur des êtres humains bénéficieraient de conseils sur l'harmonisation entre la Politique et l'EPTC2 (2022).

(Groupe de spécialistes de la gestion des données de recherche)

4.2 Échéancier d'application

Les personnes participantes ont exprimé différents points de vue sur l'échéancier approprié pour l'application de l'exigence de dépôt des données (figure 1). Pour certaines, un délai serré ne donnerait pas suffisamment de temps aux chercheuses et chercheurs, aux établissements et aux responsables de dépôt de données pour se préparer, surtout vu la complexité de l'entreprise (inquiétudes entourant les données sensibles, les données autochtones, les mégadonnées, la curation et le soutien au dépôt). Inversement, d'autres ont fait remarquer le retard du Canada par rapport à d'autres organismes de financement de la recherche, qui exigent déjà le dépôt de données ainsi que leur partage. Selon ces personnes, il serait impossible d'atteindre un état de préparation qui permettrait une application sans faille de l'exigence dans le milieu de la recherche canadien. Elles ont insisté sur l'importance d'agir tôt plutôt que tard afin d'encourager la communauté de recherche à s'adapter. Certaines personnes participantes pensent que l'annonce d'une date de prise d'effet inciterait les établissements à agir et les pousserait à améliorer leur état de préparation une fois la date fixée. Leur raisonnement est que les directions des établissements hésitent souvent à investir dans une initiative qui n'est pas encore active. Il a été souligné que la transparence concernant l'échéancier et le plan des organismes est nécessaire pour instaurer la confiance et favoriser l'acceptation par la communauté.

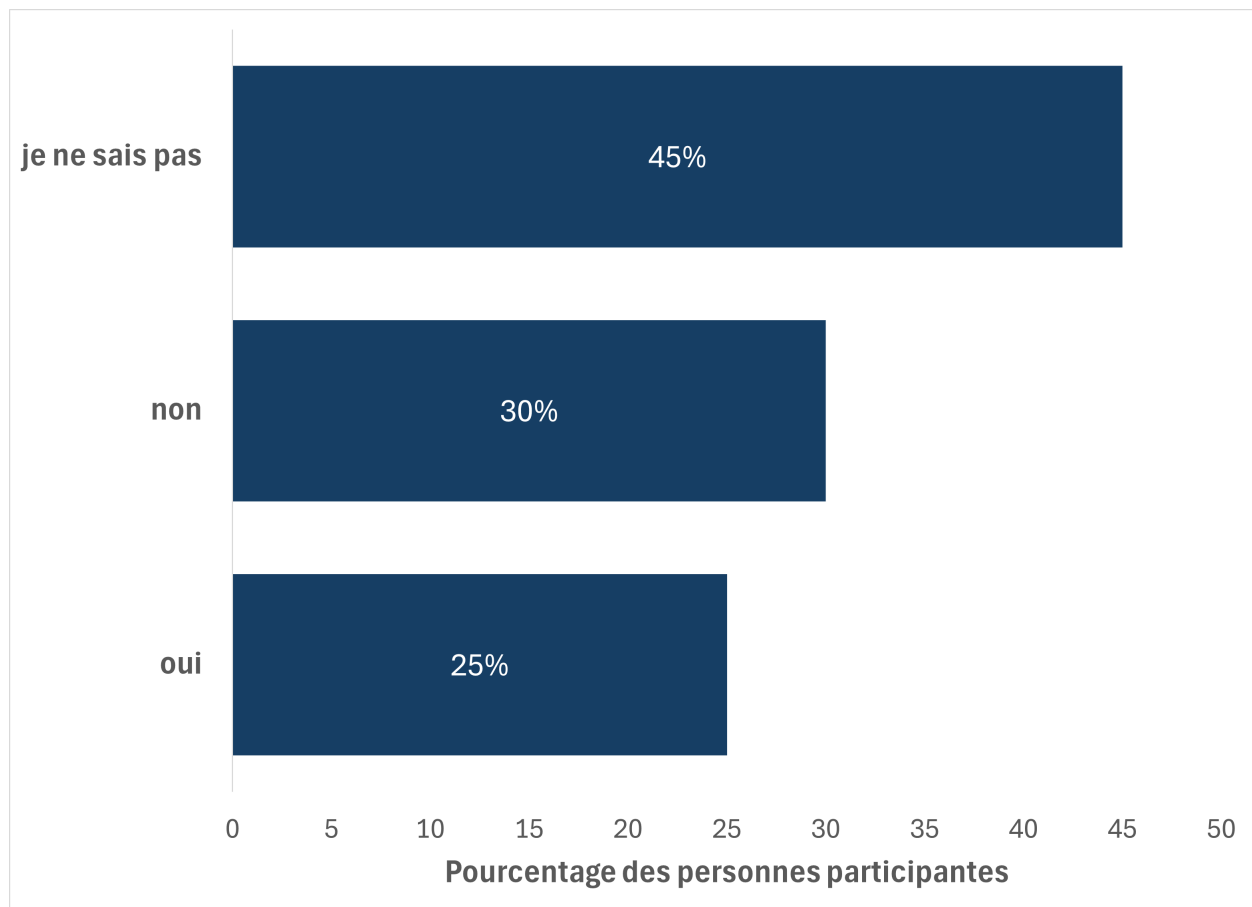
Beaucoup de personnes participantes ont proposé une approche graduelle pour gérer les difficultés et tenir compte des différents états de préparation des communautés et établissements de recherche. Une approche graduelle permettrait une intégration par étapes de l'exigence, en allant des cas les plus simples aux cas les plus complexes (p. ex. les données sensibles ou les mégadonnées). Toutefois, des

doutes subsistent : une approche graduelle calquée sur le déploiement du PGD pourrait prendre trop de temps et retarder considérablement le projet. L'importance des PGD pour aider les chercheuses et chercheurs à planifier le dépôt des données a également été soulignée, puisque les PGD visent à préciser où et comment les données de recherche seront gérées, stockées, déposées et partagées, le cas échéant.

Ce ne sont pas tous les établissements qui sont prêts, mais la barre est assez basse pour qu'ils puissent l'être quand une date sera officiellement fixée. Le Canada est en retard par rapport aux États-Unis, à l'Europe, au Japon et à d'autres pays. Il est important de fixer une date pour donner aux établissements une échéance concrète.

(Personne participante)

Figure 1. Cette figure montre le pourcentage de personnes participantes (N = 138) qui ont répondu « Oui », « Non » ou « Je ne sais pas » à la question « Compte tenu du délai entre le financement et la publication, pensez-vous que le milieu de la recherche canadien sera prêt à respecter l'exigence de dépôt de données applicable aux subventions accordées après le 1^{er} janvier 2026? ». La question a été posée en janvier 2025 dans un sondage en direct, pendant une séance virtuelle sur l'application de l'exigence de dépôt des données.



4.3 Capacité et besoins

Beaucoup de personnes participantes étaient préoccupées par le manque de capacité, de ressources et d'infrastructures pour faciliter la conformité à l'exigence. Les représentantes et représentants d'établissements, surtout les plus petits comme les collèges et CÉGEPs, doutent de leur capacité à respecter l'exigence, car elles et ils ont moins de ressources, d'accès à de l'expertise et de moyens que les grands établissements pour appliquer des politiques complexes. Les personnes participantes ont mentionné que, malgré les stratégies de gestion de données de recherche publiées, les établissements doivent élaborer des stratégies à long terme et des plans d'application pour soutenir la gestion et le dépôt des données.

Les demandes de financement pour les services de curation des données ainsi que la formation des chercheuses et chercheurs et du personnel de soutien ont été mentionnées à maintes reprises pendant les consultations. Sans ressources destinées à la formation et au soutien, beaucoup d'établissements auront du mal à faire appliquer l'exigence. Certaines personnes ont insisté sur la nécessité d'avoir du parrainage et une direction dans les établissements pour prioriser l'application des pratiques de gestion des données de recherche et y participer.

Beaucoup ont parlé des contraintes de temps imposées aux chercheuses et chercheurs ainsi qu'aux établissements : le temps nécessaire pour bien gérer, conserver et déposer les données est une préoccupation importante. Il ressort des consultations que la gestion des données de recherche est souvent perçue comme une tâche supplémentaire plutôt que comme faisant partie intégrante du processus de recherche. Les administratrices et administrateurs de la recherche et les bibliothécaires sont déjà surchargés, et l'application de l'exigence de dépôt sans nouvelles ressources pourrait leur ajouter beaucoup de travail.

Il faut plus d'argent pour soutenir cette initiative et une meilleure formation des chercheuses et chercheurs, ce qui nécessite aussi du financement. Je pense qu'il est déraisonnable de s'attendre à ce que les établissements utilisent le FSR [Fonds de soutien à la recherche] et l'allocation de Sécurité de la recherche. Ces fonds sont déjà pleinement utilisés à des fins spécifiques.

Beaucoup d'entre nous dirigent des équipes de recherche avec de petites subventions et un budget fixe, mais font face à l'inflation galopante. Les coûts supplémentaires liés à la gestion des données seront un défi pour un grand nombre d'entre nous.

(Personne participante)

4.4 Répertoires de données

Les fournisseurs de services, les chercheuses et chercheurs et les spécialistes de la gestion des données de recherche ont souligné l'ampleur et la résilience de l'infrastructure et des services existants de dépôt de données au Canada. Le [Dépôt fédéré de données de recherche](#) (DFDR) et [Boréal](#), [entre autres](#) (site en anglais), offrent aux chercheuses et chercheurs une infrastructure de dépôt hébergée au Canada en accès libre. La découverte de données est rendue possible par [lunaris](#), un service national bilingue de découverte de données provenant d'établissements de recherche à travers le pays. Bien qu'il y ait des restrictions quant aux types de données que les dépôts de données canadiens acceptent actuellement, des projets sont en cours pour élargir leur champ d'application et fournir des solutions pour contrôler l'accès aux données sensibles, lorsque le dépôt est possible et approprié.

Les personnes participantes ont souligné l'importance de définir clairement en quoi consiste un dépôt de données approprié pour aider les chercheuses et chercheurs à les différencier des autres solutions d'entreposage de données, tel que le stockage actif. Beaucoup de chercheuses et chercheurs stockent leurs données sur les serveurs de leur établissement ou des plateformes infonuagiques, mais il n'est pas clair si ces pratiques répondent à l'exigence de dépôt des données, ce qui sème la confusion.

Des directives sur les « caractéristiques souhaitables des dépôts de données », comme celles fournies par les National Institutes of Health et l'Office of Science and Technology Policy (OSTP) des États-Unis, aideraient les chercheuses et chercheurs à faire le bon choix sans restreindre leurs options. Il a également été dit que les organismes ont un rôle à jouer dans la promotion de l'utilisation d'une infrastructure canadienne, particulièrement pour stocker, déposer et protéger les données sensibles.

Beaucoup de personnes participantes ont parlé de la nécessité d'établir des normes minimales pour les dépôts de données, notamment des plans de durabilité, la possibilité d'émettre des identifiants pérennes (p. ex. identificateurs d'objets numériques [DOI]) et de définir les périodes de curation, pour garantir un stockage et une curation sécuritaires des données. Par exemple, les dépôts de données devraient avoir un mandat et un plan de curation et/ou de préservation des données; fournir de l'information permettant aux gens de découvrir les données et de se renseigner sur leur dépôt; donner aux utilisatrices et utilisateurs potentiels un accès direct aux données ou à l'information sur les conditions d'accès; veiller à ce que chaque ensemble de données ait un identifiant pérenne (PID) permettant de le définir et de le citer facilement dans les publications.

Il a été dit qu'une exigence pour les dépôts de données d'obtenir une certification (p. ex. CoreTrustSeal) pourrait être trop restrictive et limiter fortement les options offertes au milieu de la recherche canadienne. Même s'ils n'ont pas les ressources nécessaires pour obtenir la certification CoreTrustSeal, beaucoup de dépôts de données généraux, dont ceux financés en totalité ou en partie par le gouvernement fédéral, sont considérés par les spécialistes de la gestion des données de recherche comme des options viables au dépôt des données. La certification n'est pas la seule façon de prouver la fiabilité d'un dépôt de données.

On a fait remarquer que l'une des principales missions des dépôt de données est la trouvabilité et l'accessibilité des données de recherche par le public. Il est donc peu probable que des fournisseurs de services de dépôt – y compris canadiens – cautionnent la soumission de données non sensibles, mais à accès restreint, par les chercheuses et chercheurs (c.-à-d. en limitant les données qui sont accessibles et peuvent être publiées).

Si un ensemble de données est toujours inaccessible, peu importe les circonstances, il faut se demander si cela vaut la peine de le rendre repérable. Les politiques des dépôts de données pourraient limiter ce genre de dépôt.

(Groupe de spécialistes de la gestion des données de recherche)

Nous connaissons Boréal et le DFDR, mais il est difficile d'essayer de démêler tous les dépôts de données par sujet ou discipline. Il faut pouvoir prodiguer des conseils sur la question.

(Personne participant à l'administration de la recherche)

4.5 Curation des données

La curation des données est le fait d'organiser, de gérer et de préserver des données pour en assurer la qualité, la facilité d'utilisation et l'accessibilité à long terme. Ce processus relève des personnes ayant généré les données et des spécialistes de la gestion des données de recherche. Différents avis ont été exprimés concernant le degré de curation requis avant le dépôt des ensembles de données. Le niveau de curation variera selon le type de données, l'utilisation prévue et la durée de stockage. Certaines personnes participantes ont soutenu l'adoption d'une norme minimale de curation des données afin d'en garantir l'accès, la découverte, la réutilisation et la préservation à long terme. Cette norme pourrait servir de point de départ pour que les ensembles de données ne soient pas simplement déposés, mais suivent autant que possible les principes FAIR (repérabilité, accessibilité, interopérabilité et réutilisation). Les personnes participantes ont parlé de l'importance de la curation pour l'interopérabilité et la réutilisation des données, mais aussi de la nécessité de définir des normes pour les métadonnées, dont la normalisation est essentielle pour que les données déposées puissent être trouvées et réutilisées.

Selon les personnes participantes, beaucoup d'établissements, même grands, n'ont pas la capacité pour aider les chercheuses et chercheurs avec la curation de données. Cependant, ce processus ne doit pas être vu comme une responsabilité uniquement des chercheuses et chercheurs, mais devrait faire partie des efforts conjugués des gestionnaires de données, des bibliothécaires et d'autres professionnelles et professionnels de l'information, qui ont l'expertise nécessaire à la bonne curation des données. Beaucoup ont toutefois parlé de l'importance cruciale d'un financement systémique supplémentaire pour soutenir les services de curation : les dépôts de données comme les établissements (bibliothèques) ont besoin de ressources pour que les données déposées soient bien consignées et gérées à des fins de réutilisation.

Maintes personnes ont aussi dit que si le degré de curation était insuffisant, beaucoup d'ensembles de données déposés seraient inutilisables ou contreviendraient aux principes FAIR, ce qui irait à l'encontre de l'exigence de dépôt des données. Les demandes d'aide à la curation présentée par le milieu de la recherche concernant les dépôts de données et les services des établissements sont appelés à augmenter rapidement après l'entrée en vigueur de l'exigence. Il est pourtant probable que la curation d'un grand nombre d'ensembles de données déposés ne soit pas adéquate si l'on n'investit pas de façon systémique.

Un problème majeur pour les petits établissements et les établissements connexes est l'inégalité des ressources pour la curation des données, même chez Boréal.

J'aimerais vraiment avoir des lignes directrices sur la manière de conserver les données et de les rendre utilisables. La NLM [National Library of Medicine aux États-Unis] a un moteur de recherche pour les dépôts de données, qui comprennent ceux de l'ICPSR [Inter-university Consortium for Political and Social Research], de Dataverse et des sciences sociales. Si on ne guide pas les gens, ils opteront pour la facilité!

(Personnes participantes)

4.6 Partage des données

Beaucoup de personnes participantes ont parlé du fait que la Politique exige le dépôt de données, mais pas leur partage en tant que données ouvertes. En effet, bien que la Politique encourage les données ouvertes et le respect des principes FAIR, elle n'impose pas de données ouvertes. Selon les personnes participantes, il vaudrait mieux définir les attentes des organismes quant au dépôt des données et préciser ce qu'il convient de faire au-delà du simple stockage numérique. On a fait remarquer qu'il est contraire aux principes FAIR sur la trouvabilité de déposer des données sans publier d'information à leur sujet, surtout des métadonnées et des identifiants pérennes (PID). Plusieurs spécialistes de la gestion des données de recherche craignent que le dépôt des données devienne un acte performatif si l'on ne fait que déposer les données sans les rendre trouvables, accessibles ou réutilisables.

Pour plusieurs personnes participantes, les données ouvertes devraient être l'attente par défaut en partage de données et en leur absence, il faudrait justifier la non-partage. Idéalement, ces justifications figureraient dans un PGD au début d'un projet de recherche. Plus de 90 % des répondantes et répondants à un sondage réalisé à ce sujet lors de séances virtuelles de consultation en janvier 2025 adhéraient à l'idée que les données déposées soient « aussi accessibles que possible et aussi inaccessibles que nécessaire ». En général, selon les personnes participantes, l'ouverture peut améliorer la découvrabilité et la réutilisation des données, même si l'accès à beaucoup d'entre elles (comme les données sensibles) doit être restreint (par le stockage dans une infrastructure sécurisée avec des gestionnaires attitrés plutôt que sur des disques durs personnels ou des plateformes nuagiques). Le fait d'exiger des métadonnées ouvertes pour les données sensibles à accès restreint, et une description claire du processus pour en demander l'accès, aiderait les organismes à promouvoir les principes FAIR.

pour les données sensibles au Canada, en particulier par l'amélioration de leur trouvabilité et de leur accessibilité.

Il y avait un fort consensus sur le fait qu'un des objectifs principaux de l'exigence est d'encourager la transparence et la réutilisation des données à des fins de nouvelles recherches. En ce sens, la Politique devrait exiger non seulement le dépôt des données relatives aux publications, mais aussi leur disponibilité pour réutilisation (le cas échéant) et leur consignation soigneuse pour qu'elles soient réellement réutilisables.

Il faut préciser comment le partage ou le dépôt de données peut servir à d'autres fins, comme la reproduction, la réutilisation, etc. Peut-être ajouter dans une FAQ que le partage de données pour reproduction et utilisation dans une autre recherche peut avoir des conséquences éthiques (EPTC2, chapitre 3, section E, article 3.13).

(Groupe de spécialistes de la gestion des données de recherche)

La responsabilité des chercheuses et chercheurs est relativement bien définie, mais qu'en est-il de celle des établissements? Les taux de départ des chercheuses et chercheurs sont élevés ces dernières années. Qui s'assure que les données restent accessibles après un départ à la retraite ou un départ du pays?

(Personne participante)

4.7 Données sensibles

L'absence d'une infrastructure nationale et institutionnelle pour déposer les données sensibles et en contrôler l'accès a été mentionnée à plusieurs reprises. Les options sont aussi limitées en dehors des dépôts de données canadiens. Selon les lois provinciales et territoriales sur les données sensibles, des solutions locales et régionales doivent être instaurées. Pour respecter les exigences de la Politique sans ignorer la réalité nationale, il pourrait être utile de créer un comité ou un conseil chargé de discuter des irritants, des défis et des occasions.

Selon certaines personnes participantes, il faudrait définir de manière exhaustive les données sensibles afin que les chercheuses et chercheurs sachent clairement comment respecter l'exigence de dépôt des données. Cette définition doit ratisser assez large pour tenir compte des différents besoins des disciplines, tout en faisant état des divers risques. On constate un manque criant de directives sur la gestion des données sensibles, notamment les données découlant de recherches avec des êtres humains (en général) et des communautés autochtones (en particulier), ainsi que sur les recherches menées par ou en collaboration avec les partenaires du secteur. Ces directives devraient porter sur la sécurité des données, la propriété intellectuelle, le contrôle de l'accès et la préservation à long terme.

Il existe également des variations dans les normes et les niveaux de risque acceptables pour les données sensibles entre les disciplines. Certaines personnes ont laissé entendre que les organismes devraient collaborer avec différentes disciplines pour établir les types de données qui sont sensibles ainsi que les niveaux de risque acceptables. D'autres ont soutenu que ces décisions incombaient à chaque domaine

et discipline. On a également avancé que les organismes ne peuvent pas donner de directives précises et exhaustives et doivent plutôt transmettre des axes généraux à suivre, surtout concernant la préparation des données en vue de leur dépôt et partage (le cas échéant), y compris des conseils sur la dépersonnalisation et l'anonymisation. Des inquiétudes ont aussi été soulevées concernant la sécurité des données et le manque de normalisation pour les données sensibles, car les méthodes standards du domaine des TI pourraient ne pas s'appliquer au dépôt de données.

Certaines personnes participantes ont recommandé une application progressive de l'exigence de dépôt pour les données sensibles afin de permettre la préparation d'infrastructures et de pratiques appropriées au pays. Entre-temps, les organismes pourraient envisager la prise de mesures comme l'obligation de publier les métadonnées avec un DOI, de rendre ces métadonnées accessibles dans les principaux outils de découverte, dont [lunaris](#) (le service national de découverte de données), et d'assortir les données d'un processus clair de demande si elles sont à accès restreint. Il a toutefois été signalé que les dépôts de données – surtout canadiens – refusent généralement les dépôts de métadonnées sans lien direct aux données, vu la difficulté de garantir la permanence des données primaires stockées ailleurs et leur accessibilité. À noter que les solutions devraient être conçues en collaboration avec les fournisseurs de services, dont l'Alliance de recherche numérique du Canada.

Les personnes participantes ont indiqué que la dépersonnalisation des données est souvent nécessaire à un accès éthique aux données sensibles. Il faudrait préciser dans la Politique que ne constitue pas une infraction le fait de donner accès à des données dépersonnalisées en vue de leur réutilisation tant qu'on respecte la vie privée des gens ayant participé à la recherche et que ceux-ci ont consenti au partage public des données. Il a aussi été dit que les métadonnées (même en l'absence de données primaires) pouvaient représenter un risque de réidentification, d'où la nécessité de donner des directives sur l'évaluation et l'atténuation des risques.

Plusieurs personnes participantes ont recommandé de faire appel à des comités d'éthique de la recherche pour la gestion des données sensibles, car ils ont les outils nécessaires pour déterminer les exigences du dépôt éthique des données et du contrôle de l'accès. Les écarts potentiels entre l'exigence de dépôt et le mandat de consentement large de l'EPTC2 ont aussi été mentionnés (voir la section 4.1. *Clarté de la Politique*). Il a été recommandé de clarifier la nécessité de veiller à ce que les obligations éthiques et juridiques soient respectées au moment du dépôt de données et du partage des données ou métadonnées anonymisées d'êtres humains.

Les organismes ont appris qu'il y avait des raisons valables de ne pas déposer ou partager certaines données sensibles de nature commerciale ou autre. Ces raisons doivent être reconnues et intégrées de manière flexible dans la Politique. Si les données ne peuvent pas être déposées dans un dépôt de données reconnu, les chercheuses et chercheurs doivent expliquer pourquoi et justifier une autre option pour leur curation et préservation, s'il y a lieu.

Il faudrait fournir aux chercheuses et chercheurs ainsi qu'aux responsables de dépôt de données une liste des types de données sensibles viables et des exigences de protection minimale.

Il n'existe ni infrastructure nationale ni infrastructure intégrée pour accueillir les métadonnées ouvertes

pour les données sensibles fermées ou à accès restreint.
(Spécialistes de la gestion des données de recherche)

4.8 Souveraineté des données autochtones

Il y avait un large consensus sur le caractère essentiel de la souveraineté des données autochtones pour l'autodétermination des Autochtones, les efforts de réconciliation et une bonne collaboration entre les chercheuses et chercheurs et les détentrices et détenteurs de droits autochtones. Pendant les consultations, les organismes ont fait remarquer que la Politique insiste sur l'importance pour les détentrices et détenteurs de droits autochtones de contrôler la collecte, la propriété, la protection, l'utilisation et le partage des données qui les concernent, conformément aux principes de gouvernance des données élaborés et approuvés. Les personnes participantes ont affirmé que l'engagement à l'égard de la souveraineté des données autochtones est crucial si l'on veut éviter que les chercheuses et chercheurs déposent des données sans le consentement des détentrices et détenteurs de droits autochtones.

Beaucoup ont parlé de la nécessité de préparer des directives et options de dépôt de données avec les communautés et les organismes de gouvernance autochtones. Il faut discuter avec les détentrices et détenteurs de droits autochtones afin qu'elles et ils puissent décider si et dans quelles conditions leurs données peuvent être déposées. Il a été question de ressources conçues par des Autochtones pour aider les chercheuses et chercheurs, les administratrices et administrateurs de la recherche, les organismes de financement et les autres parties à maintenir des niveaux adéquats de participation et d'élaboration conjointe pour les recherches impliquant des détentrices et détenteurs de droits autochtones. Ont notamment été citées les ressources suivantes : l'Indigenous Research Level of Engagement Tool ([IRLET](#)) et les Circumpolar Inuit Protocols for Equitable and Ethical Engagement ([EEE-Protocols-LR-WEB](#)) du Conseil circumpolaire inuit. Les personnes participantes ont également souligné l'importance des principes PCAP^{MD} (propriété, contrôle, accès et gestion) comme cadre pour la souveraineté des données des Premières Nations. Toutefois, il est reconnu dans la Politique que ce modèle ne correspond pas toujours aux besoins et valeurs des communautés, regroupements collectifs et organisations des Premières Nations, métis et inuits, et qu'une approche fondée sur les distinctions est nécessaire à la reconnaissance et au respect des droits, intérêts et réalités uniques de ces communautés.

Les organismes ont appris que certains responsables de dépôt de données pourraient devoir élaborer des politiques et procédures sur le respect de la souveraineté des données autochtones pour aider les communautés autochtones à gérer leurs données et à protéger leurs droits. Il a été recommandé de définir, dans la Politique, ce qu'on entend par « communauté » et de veiller à ce que les gouvernements autochtones souverains jouissent d'un pouvoir décisionnel sur les recherches et données concernant leurs citoyennes et citoyens. En outre, la Politique devrait explicitement reconnaître et soutenir les protocoles d'entente et accords sur les données couramment utilisées dans les recherches réalisées par et en collaboration avec des Autochtones afin de préciser les modalités de gestion, de dépôt et d'accessibilité des données.

Les recommandations suivantes ont été faites pour aider les organismes à promouvoir la souveraineté et la gouvernance des données autochtones : clarifier le concept de données autochtones en fournissant une définition dans les documents d'application de la Politique; rédiger les documents, les PGD et les propositions de recherche des organismes en langage simple afin qu'ils puissent être compris par toutes les parties concernées; rendre les modèles de PGD suffisamment flexibles pour qu'ils puissent être adaptés aux besoins et caractéristiques uniques des communautés; exiger que les demandeurs de subvention expliquent dans leur PGD comment la gestion des données a été élaborée conjointement avec les détentrices et détenteurs de droits autochtones; verser des subventions pour encourager et favoriser la participation précoce des détentrices et détenteurs de droits autochtones, avant même la rédaction d'une proposition de recherche; et favoriser la rédaction de directives sur la participation des chercheuses et chercheurs venant de la communauté autochtone visée par la recherche.

Les membres du Cercle de leadership autochtone en recherche des trois organismes et du Groupe de référence sur les bonnes pratiques d'évaluation par les pairs pour la recherche autochtone ont insisté sur l'importance d'investir dans les infrastructures essentielles de gestion et de stockage des données pour les communautés, regroupements collectifs et établissements autochtones. Elles et ils ont aussi parlé de la nécessité de lier les connaissances à la communauté, indiquant que certains concepts, comme les « données » et la « propriété intellectuelle », ne s'inscrivent pas dans beaucoup de façons dont les Autochtones conçoivent le monde. Elles et ils ont encouragé la mise en place d'un cadre pour le dépôt des données qui met l'accent sur la préservation, la protection et la transmission du savoir traditionnel aux prochaines générations.

Quand on parle de PGD et de dépôt de données, rien ne devrait se faire sur nous sans nous. Quand une communauté n'a pas le temps, la capacité ou l'intention de participer à un projet de recherche, la recherche ne devrait pas se faire.

(Personne participant à l'administration de la recherche autochtone)

Les gouvernements autochtones sont des détenteurs de droit; ils doivent être consultés sur les PGD et le dépôt des données.

Personne faisant de la recherche

4.9 Mégadonnées

Les organismes ont appris que l'application de la Politique devrait tenir compte des considérations liées aux immenses ensembles de données, ou « mégadonnées ». Selon les personnes participantes, les pratiques de gestion, de stockage, de curation et de partage des mégadonnées découlant de projets utilisant le calcul informatique de pointe (CIP) diffèrent considérablement de celles qui s'appliquent aux plus petits ensembles de données, qui peuvent facilement être déposés dans des dépôt de données.

Plusieurs personnes ont constaté que beaucoup de dépôt de données n'ont pas l'infrastructure nécessaire aux mégadonnées. Les études utilisant le CIP génèrent des volumes de données importants,

données qui doivent être manipulées et stockées par des gens spécialisés, ce dont manquent beaucoup de dépôt de données actuellement. Les systèmes de CIP ont pourtant cette capacité. Par ailleurs, le transfert d'immenses ensembles de données (qui atteignent de centaines de téraoctets ou de pétaoctets) vers des dépôts de données pose beaucoup de défis, même si le dépôt de données à l'infrastructure adéquate. Par conséquent, l'exigence de dépôt devrait prévoir que les mégadonnées puissent demeurer dans les systèmes appropriés, même si ceux-ci ne cadrent pas parfaitement avec les principes FAIR (p. ex. les données ne sont ni trouvable ni assorties d'un PID, les normes communes aux métadonnées ne sont pas appliquées, le stockage n'est pas garanti ou exige une nouvelle demande).

Il faudrait donc consulter et collaborer étroitement avec les [sites d'hébergement nationaux en CIP](#) pour que l'exigence de dépôt des données soit cohérente et réaliste pour les mégadonnées. Il faudrait aussi approfondir la question du partage des métadonnées, de leur accès direct et de leur contrôle sur les sites d'hébergements nationaux.

Sur le plan de la communauté du CIP ou du CHP [calcul de haute performance] (les administratrices et administrateurs de systèmes, les analystes, les gestionnaires et les chercheuses et chercheurs mêmes), nous gérons et contrôlons déjà les données ainsi que leur accessibilité quand nous utilisons les sites d'hébergement nationaux. Les données se trouvent dans nos systèmes de stockage. Il n'y a pas à faire de « dépôt » autre part. Ce dont nous avons besoin, c'est un processus pour assurer la conformité des données aux pratiques exemplaires de gestion des données de recherche, avec le consentement des chercheuses et chercheurs.

(Analyste du stockage, CIP)

4.10 Suivi et conformité

Certaines personnes participantes ont demandé la clarification des objectifs de surveillance de la conformité des chercheuses et chercheurs à l'exigence de dépôt des données. Il faudrait préciser la conformité que les organismes entendent assurer : le simple dépôt des ensembles de données ou aussi le respect des principes FAIR pour que les données soient vérifiables et réutilisables? Beaucoup de personnes ont fait valoir que les organismes devraient réfléchir à une façon de garantir la réutilisabilité des données déposées : la surveillance de la conformité devrait déterminer si les données sont présentes, certes, mais aussi si elles sont bien consignées et peuvent être réutilisées. Il a toutefois été reconnu que, comme les organismes ne peuvent pas contrôler chaque ensemble de données, la surveillance de la conformité doit porter avant tout sur des critères gérables et efficaces, tels que les métadonnées essentielles, des procédures claires pour les données à accès restreint et des formats de fichiers non exclusifs, entre autres. Les organismes pourraient évaluer les données déposées aléatoirement, se servir des PID pour lier les articles publiés aux ensembles de données déposés et aux subventions, et rappeler aux chercheuses et chercheurs leurs obligations, au besoin.

Plusieurs personnes participantes ont parlé des PGD comme point de départ pour la surveillance de la conformité. Les PGD pourraient détailler les résultats attendus et être modifiés au fil de l'étude. Ils pourraient aussi servir à déterminer si les chercheuses et chercheurs ont respecté les exigences de

dépôt et de partage public des données, le cas échéant, ou l'obligation d'inclure les métadonnées et de donner un accès contrôlé aux données à usage restreint. En outre, la conformité à l'exigence de dépôt de données pourrait être intégrée aux rapports de fin de subvention ou dans les autres demandes de financement. Ces rapports et demandes pourraient être liés aux PID des ensembles de données et codes déposés afin d'en améliorer la découvrabilité.

Plusieurs personnes ont suggéré de lier la surveillance de la conformité à la [Politique des trois organismes sur le libre accès aux publications](#) et d'harmoniser les mandats de conformité pour le libre accès et le dépôt des données. L'évaluation responsable de la recherche devrait être au cœur des mécanismes de conformité et des incitatifs au dépôt et au partage des données, conformément aux recommandations de la [Déclaration de San Francisco sur l'évaluation de la recherche](#) (site en anglais).

Des inquiétudes ont été exprimées quant à la manière dont la conformité sera surveillée si les données sont déposées ailleurs que dans des dépôts de données, par exemple quand des données sensibles sont stockées sur les serveurs d'un établissement ou dans une solution infonuagique. Les personnes participantes ont déclaré que les organismes devraient préciser comment ils vérifieront le dépôt des données à usage restreint sur des plateformes fermées, puisque cette vérification pourrait être difficile. Un mécanisme de surveillance de la conformité pour les ensembles de données à accès restreint doit veiller à ce que les données et métadonnées déposées restent trouvables, faciles à consulter et bien gérées.

L'admissibilité aux prochaines subventions pourrait-elle dépendre de la conformité démontrée précédemment?

Voici quelques incitatifs au partage des données : réformes d'évaluation de la recherche pour tenir compte d'une variété de résultats (ensembles de données, logiciels, contributions sociétales, etc.), accès ultérieur à plus de financement et promotion des résultats de recherche.

(Personnes participantes)

4.11 Autres points à considérer

Impartialité et équité en recherche : Beaucoup de personnes participantes ont dit que l'exigence de dépôt des données ne devrait pas engendrer d'inégalités, ou exacerber celles existantes, entre les chercheuses et chercheurs, les établissements et les groupes de recherche. L'application devrait tenir compte du fait que certains groupes auront plus de mal que d'autres à gérer la surcharge de travail, donc à prévoir les ressources et le soutien nécessaires.

Rétention et préservation des données : Plusieurs personnes participantes ont recommandé de clarifier le lien entre la rétention, la préservation et le dépôt des données. Les politiques de rétention et de préservation changent d'un dépôt de données à l'autre. Les mesures à prendre dépendent du type de données déposées et du niveau de granularité de la préservation. Les organismes pourraient collaborer avec la communauté pour proposer des recommandations ou exigences communes de rétention. Les personnes participantes ont dit qu'il y avait une différence entre stocker des données sur un serveur sans les supprimer (rétention) et appliquer des pratiques numériques de préservation qui garantissent

l'accessibilité continue du contenu. Il faudrait également tenir compte de la durée de rétention des ensembles de données et des codes à des fins de vérification et de reproductibilité par opposition à la durée de rétention en vue d'une réutilisation et d'une intégration à d'autres recherches.

Garde des données : Les organismes ont appris qu'actuellement, il n'y a ni directives ni principes pour déterminer qui doit gérer l'accessibilité aux données à usage restreint et aux données sensibles à long terme. A également été soulevée la question des droits de licence et des droits d'auteur pour toutes les données : les responsables de dépôt de données et les établissements auront du mal à gérer les données au fil du temps sans la participation des chercheuses et chercheurs si les données ne sont pas sous licence libre. Selon les personnes participantes, il faudrait bonifier la définition de « garde des données » et indiquer pourquoi cette activité est souhaitable ainsi que ce qu'elle apporte aux établissements, aux organismes et à leur personnel sur les plans de la propriété, des responsabilités, des mesures et des besoins en ressources. Il a été dit que la quasi-totalité des responsables de dépôt de données, ou presque, refuseront d'assumer la garde des données et que les établissements auront du mal à définir et à assumer la garde d'ensembles de données éparpillés dans différents dépôts de données. Peu importe la définition donnée à la garde des données et aux responsabilités connexes, il faudra réfléchir à ces deux points et les noter dans un PGD dès le début d'un projet. Pour la garde des données, on pourrait fixer la durée durant laquelle une chercheuse ou un chercheur doit partager ses données ou encore la durée maximale de stockage des données dans un dépôt de données.

Éthique de la réutilisation des données : Certaines personnes participantes ont parlé de la nécessité d'inclure dans la Politique les gens qui réutilisent des données ainsi que leurs responsabilités, notamment sur le plan éthique et lors de la citation de la source appropriée. Une des recommandations était de tenir compte dans la Politique des exigences éthiques et juridiques associées à la réutilisation des données de recherche.

Dépôt des codes : Les organismes ont appris que la quantité des ressources et de l'expertise nécessaire au dépôt et au partage des codes selon les principes FAIR varie grandement, notamment parce que les codes en recherche vont de simples scripts d'analyse à des progiciels et bibliothèques complexes qui nécessitent un environnement de traitement spécifique pour être lancés. Le niveau d'utilisation des codes dans les analyses des chercheuses et chercheurs – contrairement à celui des logiciels avec interface utilisateur graphique – varie aussi grandement. Il y a un fort consensus dans la communauté de la curation canadienne : les connaissances, aptitudes et capacités nécessaires à la curation des codes sont bien loin derrière celles nécessaires à la curation des données, ce qui pourrait constituer un défi pour l'application d'une exigence de dépôt de codes respectueux des principes FAIR, en plus des données. En outre, peu de dépôts de données existants sont adaptés au dépôt des codes : les versions que l'on retrouve sur les plateformes de contrôle, comme GitHub, sont couramment utilisées pour partager des codes, mais n'ont pas les fonctionnalités clés des dépôts de données pour que le code soit accessible, interopérable et préservé.

Les pratiques exemplaires de curation des codes n'ont pas la portée et la maturité de celles pour la curation des données. Des formations et un savoir collectif sont nécessaires pour répondre aux besoins anticipés. Les services commerciaux comme GitHub jouent un rôle très précieux dans le paysage des

logiciels de recherche et des codes. Il convient aussi de noter que leurs mandats ne concordent pas entièrement avec ceux des dépôts de données sans but lucratif. Les politiques, offres et prix pourraient changer à tout moment en raison des besoins d'affaires.

(Groupe de spécialistes de la gestion des données de recherche)

5. Conclusion

Les échanges ont fait ressortir d'importants points à prendre en compte pour l'exigence de dépôt des données de la *Politique des trois organismes sur la gestion des données de recherche*, notamment les **différents points de vue des parties prenantes**, y compris les chercheuses et chercheurs, les bibliothécaires, les spécialistes de la gestion des données de recherche et des TI, les administratrices et administrateurs de la recherche, et les représentantes et représentants de groupes autochtones et de diverses organisations. Les thèmes récurrents étaient la volonté d'avoir des directives claires et pratiques, l'importance de soutenir les chercheuses et chercheurs, les inquiétudes concernant la hausse de la charge de travail et des coûts pour les établissements, et l'obligation de respecter la souveraineté des données autochtones.

Concernant l'évolution des **politiques et pratiques de gestion des données de recherche sur les scènes nationale et internationale**, les personnes participantes ont fait observer que le Canada est en retard par rapport aux autres grands organismes de financement de la recherche pour ce qui est du dépôt et du partage obligatoires des données. A été soulignée l'importance de respecter les pratiques exemplaires internationales, comme les principes FAIR. La plupart des personnes participantes estimaient que les données de recherches doivent être « aussi accessibles que possible et aussi inaccessibles que nécessaire ». Il a aussi été question du paysage évolutif des dépôts de données et des défis associés à la gestion, au dépôt et au partage de certains types de données, surtout les données sensibles et les mégadonnées.

Pour ce qui est **des défis, des besoins, de la capacité et de l'état de préparation**, les personnes participantes ont insisté sur l'importance d'avoir des ressources supplémentaires pour soutenir les services et infrastructures de gestion des données de recherche, surtout dans les plus petits établissements. Voici les besoins qui ont été souvent mentionnés pour l'application : définition des termes clés; directives concernant les données sensibles, commerciales et autochtones; et financement des activités essentielles, comme la curation et la formation.

Il y a eu de grandes divergences d'opinions concernant un **échancier réaliste d'application de l'exigence de dépôt de données**. Certaines personnes étaient en faveur d'une adoption rapide pour opérer le changement, tandis que d'autres ont émis des réserves parce que les services et infrastructures nécessaires ne sont pas encore prêts. L'adoption d'une approche graduelle a reçu beaucoup de soutien, parce qu'elle permettrait de tenir compte des différents degrés de préparation et des défis associés aux divers types de données de recherche. Dans l'ensemble, les consultations ont confirmé la nécessité d'éclaircir la Politique, d'ajouter des ressources et de poursuivre le dialogue pour que l'application de l'exigence de dépôt de données soit fructueuse et équitable.

6. Prochaines étapes

Les organismes rédigeront un plan d'application de l'exigence de dépôt des données qui tiendra compte des précisions demandées durant les consultations. Ce plan sera ensuite transmis au milieu de la recherche pour rétroaction.

Nos coordonnées

IRSC : ResearchData-Donneesderecherche@cihr-irsc.gc.ca

CRSNG : ResearchData-Donneesderecherche@nserc-crsng.gc.ca

CRSH : ResearchData-Donneesderecherche@sshrc-crsh.gc.ca

Pour en savoir plus sur la *Politique des trois organismes sur la gestion des données de recherche* :
<https://science.gc.ca/site/science/fr/financement-interorganismes-recherche/politiques-lignes-directrices/gestion-donnees-recherche/politique-trois-organismes-gestion-donnees-recherche>