

# Conceptualisation d'une approche nationale quant aux services de curation de données au Canada

Version 2.0

Juin 2020

Réseau Portage  
Association des bibliothèques de recherche du Canada  
[portage@carl-abrc.ca](mailto:portage@carl-abrc.ca)

[www.carl-abrc.ca](http://www.carl-abrc.ca)

**portage**

SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE  
SHARED STEWARDSHIP OF RESEARCH DATA

**CARL ABRC**  
CANADIAN ASSOCIATION OF RESEARCH LIBRARIES  
ASSOCIATION DES BIBLIOTHÈQUES DE RECHERCHE DU CANADA

## Auteurs

Ce rapport synthétise les idées et les réflexions des participants au Forum canadien de renforcement de la communauté de curation de données en octobre 2019. Parmi les auteurs du rapport, il y a :

**Erin Clary**

Réseau Portage  
[erin.clary@carl-abrc.ca](mailto:erin.clary@carl-abrc.ca)

**Jason Brodeur**

Université McMaster  
[brodeujj@mcmaster.ca](mailto:brodeujj@mcmaster.ca)

**Lee Wilson**

Réseau Portage  
[lee.wilson@ace-net.ca](mailto:lee.wilson@ace-net.ca)

**Jeff Moon**

Réseau Portage  
[moonj@queensu.ca](mailto:moonj@queensu.ca)

**Shahira Khair**

Université de Victoria  
[skhair@uvic.ca](mailto:skhair@uvic.ca)

Les versions provisoires ont été examinées et révisées par :

**Jennifer Abel**

Réseau Portage

**Ève Paquette-Bigras**

Université de Montréal

**Corey Davis**

Université de Victoria

**Laure Perrier**

Université de Toronto

**Reyna Jenkyns**

Ocean Networks Canada

**Kevin Read**

Université de la Saskatchewan

**Lisa Johnston**

Université du Minnesota

**Felicity Tayler**

Université d'Ottawa

**Caitilín O'Hare**

Réseau Portage

**Minglu Wang**

Université de York

Participants au Forum canadien sur la curation de données et membres de la communauté élargie de gestion de données de recherche au Canada.

## Remerciements

L'Association des bibliothèques de recherche du Canada (ABRC), le Réseau Portage, la bibliothèque de l'Université McMaster et le Conseil de recherche en sciences humaines (CRSH) ont contribué au financement du Forum canadien sur la curation de données. Nous souhaitons aussi remercier le Groupe d'experts sur la curation des données de Portage ainsi que les comités de programmation et de logistique locale qui ont permis la tenue du Forum canadien sur la curation des données.

SSHRC  CRSH



Library



## Table des matières

<b>AUTEURS</b> .....	2
<b>REMERCIEMENTS</b> .....	3
<b>RÉSUMÉ</b> .....	5
<b>INTRODUCTION</b> .....	7
CURATION DES DONNÉES DANS LE CONTEXTE DE LA GDR.....	9
APPUYER UNE CURATION RIGOUREUSE DES DONNÉES À L'ÉCHELLE NATIONALE .....	11
<b>RÉSUMÉ DE L'ÉVÈNEMENT</b> .....	14
SÉANCES DE FORMATION SUR LA CURATION DES DONNÉES .....	14
FORUM SUR LE DÉVELOPPEMENT COMMUNAUTAIRE .....	16
<b>RÉSULTATS DU FORUM SUR LE DÉVELOPPEMENT COMMUNAUTAIRE</b> .....	17
OBSTACLES ET DÉFIS .....	17
CONCEPTUALISATION D'UNE APPROCHE NATIONALE DES SERVICES DE CURATION .....	19
ACTIVITÉS RECOMMANDÉES.....	20
<b>PROCHAINES ÉTAPES ET DISCUSSION</b> .....	25
FEUILLE DE ROUTE .....	25
AUTRES CONSIDÉRATIONS.....	28
<b>RÉSUMÉ DES RECOMMANDATIONS RELATIVES À LA NOIRN</b> .....	31
<b>BIBLIOGRAPHIE</b> .....	33
<b>ANNEXE</b> .....	37
LISTE DES SIGLES .....	37
RÉPARTITION DES PARTICIPANTS AU FORUM SUR LE DÉVELOPPEMENT COMMUNAUTAIRE DU FCCD .....	38
APERÇU DES DÉPÔTS NATIONAUX DE DONNÉES PRÉSENTÉ AU FCCD .....	39

## Résumé

La curation de données est une gestion active de données de recherche à chaque étape de production, soit la création, la sauvegarde, l'utilisation, l'archivage, le partage et la réutilisation des données. C'est un processus itératif qui donne de la valeur à l'érudition en optimisant les ensembles de données pour leur usage actuel ainsi que leur découvrabilité et réutilisation future ; il assure que les ensembles de données que développent et déposent les chercheurs incarnent les principes d'équité. Les tâches liées à la curation de données nécessitent des compétences diversifiées, notamment des connaissances de la discipline, du processus de recherche, des normes de métadonnées et des pratiques exemplaires de gestion de données ainsi qu'une aisance à travailler dans un environnement très technologique. Étant données les complexités intrinsèques de la curation de données, il y a différentes lacunes sur le plan du savoir, des compétences, des outils et de la coordination qui limitent la valeur ajoutée des activités de gestion de données de recherche. Un besoin de formation dépassant les capacités des curateur de données et nécessitant une coordination élargie des services de curation de données dans l'ensemble des organisations existe.

Ce résumé présente les conclusions du Forum canadien sur la curation de données, un événement national de trois jours tenu en octobre 2019 visant à trouver les obstacles et les défis ressentis dans le travail de curation et à articuler la manière dont une approche nationale de curation de données au Canada pourrait surmonter ces défis. Parmi les participants au forum, il y avait des chercheurs, des professionnels en gestion de données et en informatique, des décideurs et des administrateurs représentant une grande diversité d'instituts de recherche et d'organisations, de bailleurs de fonds et d'organismes de soutien de partout au Canada. Les participants ont identifié les éléments suivants comme obstacles au travail de curation de données :

- Capacité de curation limitée étant donné les priorités concurrentes ;
- Aucun service formel de gestion de données de recherche ou aucune ressource dédiée pour aider les producteurs de données avec la curation ;
- Absence d'un message communiquant la valeur de la curation à différents intervenants ;
- Besoin de normes, de flux de travail, de pratiques exemplaires et d'opportunités de formation ;
- Besoin de clarification des rôles et des responsabilités et des données probantes en appui aux services les plus utiles pour les chercheurs et les producteurs de données.

En réponse à ces enjeux, les participants ont envisagé une approche nationale concertée pour la curation de données qui mettrait à contribution le travail des organisations actuelles, dont la nouvelle Infrastructure de recherche numérique. Les prochaines étapes du développement d'un service national de curation de données sont présentées dans le rapport (voir la feuille de route à

la page 25). Celles-ci suscitent l'engagement des chercheurs, des producteurs de données et autres intervenants pour faire en sorte que les services soient pertinents ; l'investissement en infrastructure, notamment l'« intervention humaine » nécessaire à la curation ; le renforcement de la communauté de pratique en curation de données ; et les ressources et les services de soutien qui aideront la capacité de curation locale. Pour jeter les bases d'une stratégie de curation nationale élargie, les actions suivantes seraient nécessaires :

- Reconnaître l'importance de la curation des données pour donner accès aux produits de données FAIR et au savoir libre ainsi que du rôle des curateurs de données dans ce processus.
- Reconnaître que la curation des données est une activité complexe et intensive qui exige un investissement important dans l'embauche et le perfectionnement de personnel hautement qualifié.
- Collaborer activement avec les chercheurs et les producteurs de données du Canada afin de mieux comprendre leurs besoins en matière de curation et les défis auxquels ils font face, et travailler avec ces communautés pour veiller à ce que les services de curation que nous développons les soutiennent.
- Contribuer à la création d'un cadre national pour les services de curation qui appuie les besoins de curation à grande échelle et répond à la diversité des besoins des divers organismes.
- Appuyer l'élaboration et la prestation de possibilités d'éducation, de ressources d'apprentissage et de documents d'orientation dans les deux langues officielles afin de favoriser le perfectionnement des curateurs de données partout au Canada.
- Collaborer avec les communautés autochtones pour veiller à ce que les services et les ressources de curation que nous développons soient en harmonie avec leurs efforts de gestion des données.
- Appuyer l'établissement et le maintien d'une communauté de pratique nationale sur la curation.
- Collaborer avec les communautés de la curation, de la conservation numérique et de l'éthique pour aborder des questions complexes liées à la curation et à la gestion à long terme des données biomédicales et d'autres données sur la santé, des données anciennes et des données récupérées, des données multilingues et des données générées à l'extérieur des établissements universitaires.
- Collaborer avec des spécialistes de la conservation numérique pour déterminer les mesures de gestion et de curation des données qui appuieront la conservation à long terme.
- Élaborer un mécanisme pour la communauté de la curation afin d'éclairer les décisions en matière de gouvernance et de politiques liées à la gestion des données de recherche.
- Appui direct ou indirect à la création de postes de curation nécessaires pour l'excellence des ensembles de données versés dans les dépôts de données canadiens.

## Introduction

Les données de recherche sont les principales sources d'information utilisées pour appuyer et valider les résultats de la recherche. Les données peuvent être générées ou recueillies à partir d'études d'observation, d'expériences ou de simulation, ou compilées, établies ou traitées à partir de données existantes<sup>1</sup>. De nombreuses données, saisies à un moment particulier, sont irremplaçables, et d'autres peuvent être trop coûteuses à recréer. Le partage des données catalyse l'innovation et la découverte, car il permet, d'une part, la réutilisation ou la transformation des produits de la recherche pour répondre à de nouvelles questions de recherche et, d'autre part, la répétition et la vérification des analyses, contribuant ainsi à résoudre les problèmes liés à la transparence et à la reproductibilité de la recherche<sup>2</sup>.

Les gouvernements, les bailleurs de fonds et les chercheurs reconnaissent la valeur du partage des données comme véhicule de progrès scientifique et de collaboration internationale, susceptible de stimuler la croissance économique et d'améliorer le bien-être social<sup>3</sup>, et un nombre croissant d'organismes de financement et d'éditeurs ont adopté des politiques qui exigent des chercheurs qu'ils assurent la communication des données sur demande ou leur archivage dans un dépôt désigné<sup>4</sup>. La *Feuille de route pour la science ouverte* du gouvernement du Canada recommande que les résultats de la recherche scientifique nouvellement financée soient « ouverts par conception et par défaut », dans la mesure du possible<sup>5</sup>, mais reconnaît également que le simple partage de données ne suffit pas pour réaliser le plein potentiel de la science ouverte. Pour ce faire, les données doivent être gérées correctement tout au long de leur cycle de vie<sup>6</sup> et respecter les

---

<sup>1</sup> CASRAI, s.v. « Research Data »; consulté le 2 février 2020. <https://casrai.org/research-data/>

<sup>2</sup> Open Science Collaboration, « Estimating the reproducibility of psychological science », *Science*, vol. 349, n° 6251, 28 août 2015, aac4716. <https://doi.org/10.1126/science.aac4716>

<sup>3</sup> Organisation de coopération et de développement économiques, « Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics », 2007. <https://doi.org/10.1787/9789264034020-fr>; CODATA et Paul Uhlir, « The Value of Open Data Sharing: A CODATA Report for the Group on Earth Observations », Zenodo, 2015. <http://doi.org/10.5281/zenodo.33830>; Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Bureau du sous-ministre, Direction générale des communications et du marketing, « Déclaration de principes des trois organismes sur la gestion des données numériques », 2016. [https://www.science.gc.ca/eic/site/063.nsf/fra/h\\_83F7624E.html](https://www.science.gc.ca/eic/site/063.nsf/fra/h_83F7624E.html)

<sup>4</sup> Voir, par exemple, 130 politiques archivées sur le site de FAIRSharing. <https://fairsharing.org/policies/>

<sup>5</sup> Gouvernement du Canada, Bureau de la conseillère scientifique en chef du Canada, « Feuille de route pour la science ouverte », 2020. [https://www.ic.gc.ca/eic/site/063.nsf/fra/h\\_97992.html](https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97992.html)

<sup>6</sup> Pour une description du cycle de vie des données, voir, par exemple, DataONE, « Data Life Cycle ». <https://www.dataone.org/data-life-cycle>; Conseil du leadership sur l'infrastructure de recherche numérique, « Data Management Position Paper For Innovation, Science, and Economic Development Canada », 2017; figure 2, p. 6 : « Data related Activities During the Research Process ». <https://wiki.queensu.ca/display/QDSRWG>

« principes FAIR » – facilement repérables, accessibles, interopérables et réutilisables<sup>7</sup>.

L'importance de saines pratiques de gestion des données pour maximiser la valeur de la recherche est largement reconnue<sup>8</sup> et la gestion des données est l'une des trois principales responsabilités de la nouvelle organisation d'infrastructure de recherche numérique (NOIRN)<sup>9</sup>. En 2020, de nombreux chercheurs devront inclure des plans de gestion des données (PGD) dans leurs propositions de subvention, et les établissements qui administrent les fonds des trois organismes subventionnaires devront élaborer une stratégie de gestion des données de recherche (GDR) qui décrit quels seront les mécanismes de soutien des efforts de gestion des données de leurs chercheurs<sup>10</sup>.

Afin de soutenir pleinement les efforts des chercheurs visant à gérer les données tout au long du cycle de vie d'un projet et à les partager à l'issue d'un projet, il faut surmonter un certain nombre d'obstacles et de défis. Un éventail de décisions et de mesures en amont peuvent avoir une incidence sur l'utilité et la valeur globales des données et déterminer si et comment les données peuvent être partagées. Le manque de temps et de ressources pour bien décrire et organiser les données en vue de l'archivage<sup>11</sup>, l'incertitude quant au lieu et à la façon de partager les données, les préoccupations concernant la réutilisation éthique des données et le manque de clarté au sujet de concepts comme la propriété des données<sup>12</sup>, le droit d'auteur et l'octroi de licences sont au nombre des inquiétudes soulevées qui peuvent limiter le partage de données. Si les données sont partagées, le potentiel de réutilisation dépend de l'intégrité des données, démontrée en partie par leur qualité et leur exhaustivité ainsi que par la documentation connexe, mais il peut être long et difficile de produire les documents appropriés, et il y a actuellement peu d'incitation à le faire<sup>13</sup>.

---

<sup>7</sup> Wilkinson et coll., « The FAIR Guiding Principles for scientific data management and stewardship », *Scientific Data*, vol. 3, 2016. <https://doi.org/10.1038/sdata.2016.18>

<sup>8</sup> Gouvernement du Canada, « Feuille de route pour la science ouverte »; Gouvernement du Canada, « Déclaration de principes des trois organismes sur la gestion des données numériques »; Baker et coll., « Research data management in Canada »; Groupe de travail sur la stratégie des données de recherche, « Sommet 2011 sur les données de recherche canadiennes : Portrait de la situation des données », 2011. <https://www.rdc-drc.ca/wp-content/uploads/f-Background-to-the-2011-Canadian-Research-Data-Summit.pdf>

<sup>9</sup> Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Bureau du sous-ministre, « Infrastructure de recherche numérique », 2019. <https://www.ic.gc.ca/eic/site/136.nsf/fra/accueil>

<sup>10</sup> Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Bureau du sous-ministre, Direction générale des communications et du marketing, « ÉBAUCHE – Politique des trois organismes sur la gestion des données de recherche – Aux fins de consultation », 2018. [https://www.science.gc.ca/eic/site/063.nsf/fra/h\\_97610.html](https://www.science.gc.ca/eic/site/063.nsf/fra/h_97610.html).

<sup>11</sup> Perrier, Laure, Blondal, Erik et Heather MacDonald, « The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis », *PLoS ONE*, vol. 15, n° 2, 2020. <https://doi.org/10.1371/journal.pone.0229182>; Stuart et coll., « Whitepaper: Practical challenges for researchers in data sharing » (version 1), Figshare, 2018. <https://doi.org/10.6084/m9.figshare.5975011.v1>

<sup>12</sup> Perrier, Blondal et MacDonald, « The views, perspectives, and experiences of academic researchers ».

<sup>13</sup> Perrier, Blondal et MacDonald, « The views, perspectives, and experiences of academic researchers ».



Étant donné que le partage de données est une exigence relativement nouvelle, les chercheurs font état également du besoin de formation, de ressources et de services qui appuieront leurs initiatives de partage de données<sup>14</sup>.

## Curation des données dans le contexte de la GDR

La curation des données est une composante importante de la GDR qui comprend l'examen itératif ainsi que la révision des données et des métadonnées afin d'améliorer leur conformité aux principes FAIR -- en les optimisant tant pour l'utilisation actuelle que pour la découverte et la réutilisation futures<sup>15</sup>. Bien que certains services et logiciels offrent des fonctionnalités qui simplifient ou automatisent partiellement certaines activités de curation (p. ex. [Dataverse, outil de curation des données](#) de Scholars Portal et [Bitcurator](#) pour le traitement des formats de fichiers et la détection d'identificateurs personnels), la nature hétérogène et multidisciplinaire des données de recherche exige habituellement l'intervention de curateurs qui ont pour mandat d'évaluer la qualité des données et des métadonnées, de définir des normes FAIR appropriées et de collaborer avec les chercheurs pour bonifier les dépôts. La curation des données ne se fait pas en un seul endroit; par ailleurs, les curateurs de données assument divers rôles et interagissent avec les données dans un éventail de capacités<sup>16</sup>. Les chercheurs, les gestionnaires de données, les intendants des données, les bibliothécaires universitaires et d'autres praticiens des données peuvent tous organiser les éléments d'un ensemble de données à mesure qu'ils cheminent dans le cycle de vie des données de recherche. Les activités de curation des données pertinentes comprennent :

- l'élaboration de stratégies de collecte et de documentation des données ;
- la création de documentation et de métadonnées pour expliquer et contextualiser les données ;
- la mise en œuvre d'un plan de stockage sécurisé, de systèmes de contrôle des versions pour les fichiers et les codes et la mise en place de mesures de protection comme les sauvegardes de fichiers et les vérifications de l'intégrité des fichiers ;
- l'assurance de la qualité au cours de l'analyse des données par l'inspection des métadonnées, la vérification des fichiers et la révision des codes ;

---

<sup>14</sup> Perrier, Blondal et MacDonald, « The views, perspectives, and experiences of academic researchers ».

<sup>15</sup> Groupe d'experts sur la curation des données du réseau Portage, « Introduction à la curation des données », Réseau Portage, Ottawa, 2019. [https://portagenetwork.ca/wp-content/uploads/2019/09/Curation\\_Primer\\_Aug2019\\_FR.pdf](https://portagenetwork.ca/wp-content/uploads/2019/09/Curation_Primer_Aug2019_FR.pdf)

<sup>16</sup> Une longue liste d'activités de curation des données et de définitions se trouve à l'annexe de Johnston et coll., « How Important Are Data Curation Activities to Researchers? Gaps and opportunities for academic libraries », *Journal of Librarianship and Scholarly Communication*, vol. 6, n° 1, 2018. <https://doi.org/10.7710/2162-3309.2198>; voir aussi le site Web du Data Curation Network. <https://datacurationnetwork.org/data-curation-activities/>

- la préparation des ensembles de données à verser dans un dépôt de données ou des archives, ce qui comprend l'application de mesures appropriées d'anonymisation ou de contrôle d'accès ;
- la transformation des formats de fichiers afin d'améliorer la réutilisation et la conservation à long terme ;
- l'augmentation des métadonnées et la mise en relation des ensembles de données pour accroître la découvrabilité<sup>17</sup>.

Considérée dans son ensemble, la curation des données profitera à toute une gamme de collectivités et de groupes d'intervenants :

*Table 1: Groupe d'intervenants et Avantages potentiels découlant de la curation des données*

<b>Groupe d'intervenants</b>	<b>Avantages potentiels découlant de la curation des données<sup>18</sup></b>
Bailleurs de fonds, gouvernement, public	Amélioration du rendement des investissements lorsque les données peuvent être trouvées, consultées, comprises et transformées pour répondre à de nouvelles questions de recherche.
Établissements	Amélioration de la conformité des chercheurs aux mandats des bailleurs de fonds et du gouvernement. Protection de la réputation associée à l'assurance de la qualité et à une plus grande transparence de la recherche.
Chercheurs et communautés de recherche	Amélioration de la transparence du processus scientifique. Amélioration de l'efficacité résultant d'une gestion adéquate des données. Amélioration de la qualité des données qui encourage les initiatives de recherche collaborative et interdisciplinaire. Citations pour les ensembles de données publiées.
Exploitants d'infrastructures (p. ex. gestionnaires de dépôts)	Amélioration de la qualité des produits grâce à une meilleure préparation pour la conservation à long terme. Application de normes visant à accroître la découvrabilité et l'interopérabilité. Probabilité accrue de satisfaire aux exigences de certification des dépôts comme le Coretrustseal <sup>19</sup> .

<sup>17</sup>Groupe d'experts sur la curation des données du réseau Portage, « Introduction à la curation des données ».

<sup>18</sup>Groupe d'experts sur la curation des données du réseau Portage, « Introduction à la curation des données; Digital Curation Centre, « Digital curation 101: Incentives for digital curation », Edinburgh, Digital Curation Centre, 2011. <http://www.dcc.ac.uk/sites/default/files/documents/DC%20101%20Incentives%20for%20Curation.pdf>

<sup>19</sup> Coretrustseal Standards and Certification Board, « Coretrustseal Trustworthy Data Repositories Requirements 2020–2022 », Zenodo, 2019. <http://doi.org/10.5281/zenodo.3638211>.

## Appuyer une curation rigoureuse des données à l'échelle nationale

La profession émergente qu'est la curation des données exige un éventail de compétences et d'expertises, or les rôles et les responsabilités confiées à un curateur peuvent ne pas être uniformes d'un établissement à l'autre. Les responsables des dépôts de données peuvent embaucher des curateurs à temps plein pour travailler avec les chercheurs à la fin d'un projet de recherche afin de préparer les ensembles de données devant être versées dans le dépôt et stockées à long terme, tandis que les bibliothécaires chargés de la gestion des données de recherche et les intendants des données peuvent estimer que certains aspects de la curation des données constituent une des nombreuses tâches adjacentes à la recherche qu'on leur demande d'accomplir tout au long du cycle de vie du projet de recherche<sup>20</sup>. Les chercheurs eux-mêmes entreprennent de nombreuses tâches de curation en gérant leurs propres données<sup>21</sup>, bien qu'ils ne se considèrent probablement pas comme des curateurs de données.

Pour appuyer les efforts de gestion et de partage des données des chercheurs, les organismes s'efforcent souvent de trouver des curateurs qui connaissent bien le processus et les méthodes de recherche, et qui peuvent chercher à acquérir de l'expérience en recherche ou des compétences spécialisées sous forme de diplôme d'études supérieures<sup>22</sup>. Les connaissances disciplinaires et la compréhension du processus de recherche peuvent être particulièrement importantes pour la prise en charge de données complexes, par exemple dans la gestion des données sensibles, auquel cas les curateurs doivent comprendre le rôle de l'éthique dans la gestion des données, le risque de divulgation, le consentement éclairé et connaître les procédures de désidentification. Dans le contexte canadien, le respect et la compréhension des principes de PCAP<sup>®23</sup> et d'autres initiatives de gouvernance de l'information dans les collectivités autochtones sont également importants. Les curateurs doivent connaître entre autres normes celles qui sont relatives aux métadonnées ainsi que les caractéristiques, les tendances et les pratiques exemplaires existantes et émergentes en matière de gestion des données et de communications savantes. Ils doivent être au courant des

---

<sup>20</sup> Kim, Jeonghyun, Warga, Edward et William Moen, « Competencies Required for Digital Curation : An Analysis of Job Advertising », *The International Journal of Digital Curation*, vol. 8, n° 1, 2013. <https://doi.org/10.2218/ijdc.v8i1.242>; Tamaro, Anna Maria, Matusiak, Krystyna K., Sposito, Frank Andreas et Vittore Casarosa, « Data Curator's Roles and Responsibilities : An International Perspective », *Libri*, vol. 69, n° 2, 2019. <https://doi.org/10.1515/libri-2018-0090>;

Mons, Barend, *Data Stewardship for Open Science: Implementing FAIR Principles*, CRC Press, New York, 2018, p. 28. <https://doi.org/10.1201/9781315380711>

<sup>21</sup> Johnston et coll., « How Important Are Data Curation Activities to Researchers? »

<sup>22</sup> Tamaro, Matusiak, Sposito et Casarosa, « Data Curator's Roles and Responsibilities ».

<sup>23</sup> Centre de gouvernance de l'information des Premières Nations, « Ownership, Control, Access and Possession (OCAP™): The Path to First Nations Information Governance », Centre de gouvernance de l'information des Premières Nations, Ottawa, 2014.

[https://fnigc.ca/sites/default/files/docs/ocap\\_path\\_to\\_fn\\_information\\_governance\\_final.pdf](https://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance_final.pdf).

diverses licences, des lois sur le droit d'auteur, des politiques des bailleurs de fonds et des politiques institutionnelles<sup>24</sup>. Les curateurs doivent être à l'aise dans un environnement à forte composante technologique, et ils peuvent être tenus de connaître divers systèmes d'exploitation, langages de script et de programmation, ainsi que des logiciels et des outils qui appuient les décisions et les mesures de curation, la conservation et l'analyse des données, la visualisation et la transformation des données<sup>25</sup>. Les curateurs peuvent également être appelés à participer à des activités de sensibilisation, à fournir un soutien technique pour divers outils et logiciels, à promouvoir les pratiques exemplaires en matière de gestion des données et à aider à élaborer des politiques de gestion des données et des flux de travail au sein de leur organisme<sup>26</sup>.

Compte tenu des complexités inhérentes à la profession, il existe diverses lacunes dans les connaissances, les compétences, les outils et la coordination de la curation des données qui peuvent limiter la valeur du rendement des activités de gestion des données de recherche<sup>27</sup>. Il est peu probable que les curateurs de chaque groupe de recherche, organisme et établissement possèdent toutes les compétences et l'expertise nécessaires pour gérer un ensemble diversifié de données tout au long de son cycle de vie<sup>28</sup>. Il y a un grand besoin de possibilités de formation pour renforcer les capacités des curateurs de données et d'occasions de coordination de haut niveau afin d'assurer la curation des données et de parfaire les pratiques dans l'ensemble des organismes. Les réseaux de curation fédérés aux États-Unis et au Royaume-Uni (le Data Curation Network et le Digital Curation Centre, respectivement) illustrent bien la façon dont les services, les ressources et la formation peuvent être coordonnés à l'échelle nationale.

Il est essentiel de combler ces lacunes, car il existe une demande croissante de services de curation. Les organismes de certification de référentiels, comme Coretrustseal, exigent un haut niveau avéré de curation des données<sup>29</sup>. Ces exigences, ainsi que d'autres normes publiées relativement aux dépôts de données fiables, comme la norme ISO 16363<sup>30</sup>, éclaireront les éditeurs et les bailleurs de fonds lorsqu'ils établiront des critères pour recommander des dépôts de données. Un ensemble de critères récemment publiés par l'Office of Science and Technology Policy

---

<sup>24</sup> Kim, Warga et Moen, « Competencies Required for Digital Curation ».

<sup>25</sup> Kim, Warga et Moen, « Competencies Required for Digital Curation ».

<sup>26</sup> Kim, Warga et Moen, « Competencies Required for Digital Curation »; Tammara, Matusiak, Sposito et Casarosa, « Data Curator's Roles and Responsibilities ».

<sup>27</sup> Johnston et coll., « How Important Are Data Curation Activities to Researchers? ».

<sup>28</sup> Johnston et coll., « Data Curation Network: A Cross-Institutional Staffing Model for Curating Research Data », *The International Journal of Digital Curation*, vol. 13, n° 1, 2018. <https://doi.org/10.2218/ijdc.v13i1.616>.

<sup>29</sup> Coretrustseal Standards and Certification Board, « Coretrustseal Requirements ».

<sup>30</sup> Organisation internationale de normalisation (ISO), « ISO 16363:2012 – Audit et certification des référentiels numériques de confiance », 2012.

des États-Unis<sup>31</sup> et une autre série de critères publiés par FAIRsharing, Datacite et un groupe de représentants de maisons d'édition<sup>32</sup> font valoir que la curation est une fonction de référentiel souhaitable qui sera requise comme pratique exemplaire minimale à l'avenir. De même, la *Déclaration de principes des trois organismes sur la gestion des données numériques* affirme que les données de recherche devraient, dans la mesure du possible, être conservées dans un dépôt ouvert et structuré<sup>33</sup> et la *Feuille de route pour la science ouverte* reconnaît que la curation est l'une des activités de gestion des données nécessaires à la mise en œuvre des principes FAIR<sup>34</sup>. Bien que la curation ne soit pas mentionnée directement, Innovation, Sciences et Développement économique Canada (ISDE) reconnaît que la nouvelle infrastructure de recherche numérique (IRN) du Canada nécessitera du personnel hautement qualifié (PHQ) possédant l'expertise nécessaire pour appuyer les chercheurs<sup>35</sup>, et un certain nombre d'exposés de position<sup>36</sup> et de documents d'information<sup>37</sup> qui ont éclairé le processus font état précisément de l'importance de la curation dans la réorganisation de l'IRN.

Afin de mieux comprendre l'état actuel de la pratique de la curation des données au Canada et les besoins à combler à chapitre, et pour exposer comment une approche nationale de la curation des données pourrait aider à relever ces défis, le réseau Portage de l'Association des bibliothèques de recherche du Canada (ABRC), en collaboration avec la bibliothèque de l'Université McMaster et le Conseil de recherches en sciences humaines (CRSH), a organisé le Forum canadien sur la curation des données à l'Université McMaster, à Hamilton, du 16 au 18 octobre 2019. Le rapport qui résume l'événement présente les résultats du forum et propose les prochaines étapes pour élargir les services de curation des données et le soutien de ces services au Canada.

---

<sup>31</sup> États-Unis, Office of Science and Technology Policy, « Request for Public Comment on Draft Desirable Characteristics of Repositories for Managing and Sharing Data Resulting From Federal Funded Research », Federal Register, Doc. 2020-00689, 85 FR 3085, 2020.

<sup>32</sup> McQuilton et coll., « FAIRsharing Collaboration with Datacite and Publishers : Data Repository Selection, Criteria That Matter », Open Science Framework, 2019. <https://doi.org/10.17605/OSF.IO/N9QJ7>

<sup>33</sup> Canada, *Déclaration de principes des trois organismes sur la gestion des données numériques* ».

<sup>34</sup> Gouvernement du Canada, « Feuille de route pour la science ouverte ».

<sup>35</sup> Canada, « Infrastructure de recherche numérique ».

<sup>36</sup> Conseil du leadership sur l'infrastructure de recherche numérique, « Data Management Position Paper »; Castle et coll., « Data Management Roadmap – 2019-2024: For Submission to Innovation, Science and Economic Development (ISED) », 2019. <https://www.rdc-drc.ca/download/position-paper-data-management-roadmap-2019-2024/?wpdmdl=2482>.

<sup>37</sup> Baker et coll., « Research data management in Canada ».

## Résumé de l'événement

Le Forum canadien sur la curation des données (FCCD) a réuni des représentants de divers établissements et organismes<sup>38</sup> dans le but de favoriser une communauté de pratique nationale pour catalyser l'élaboration et l'adoption de normes, de pratiques et d'outils de curation des données, ainsi que le perfectionnement et l'acquisition de compétences dans l'ensemble des disciplines et des établissements. L'événement s'est déroulé en deux parties. Pendant la première journée et demie, les participants ont assisté à des ateliers axés sur le partage des connaissances et l'acquisition de compétences pratiques. A suivi un Forum sur le développement communautaire dont l'objet était de mieux comprendre les capacités et les défis actuels du Canada en matière de curation. Les participants du Forum sur le développement communautaire ont contribué activement à l'élaboration d'une vision d'une approche nationale de la curation des données au Canada qui permettrait de relever ces défis.

## Séances de formation sur la curation des données

La formation sur la curation des données offerte lors du FCCD était une occasion de perfectionnement professionnel qui a permis aux participants de faire du réseautage et d'accroître leurs compétences en gestion et curation des données dans le cadre de divers ateliers. Chaque participant a assisté à trois des six séances simultanées d'une demi-journée. Les pratiques exemplaires et des ressources importantes ont été mises en commun pendant les ateliers et les praticiens ont pu acquérir des connaissances et de l'expérience grâce à des activités de formation pratique.

---

<sup>38</sup> Voir à l'annexe B la répartition des participants au Forum de développement communautaire par type d'organisme, région géographique et poste.

Table 2: Atelier et Description

Atelier	N°	Description
Introduction à la programmation R pour les spécialistes en sciences sociales	1a	Présentation de la syntaxe R et l'interface RStudio, importation des données, bases de données, facteurs, ajout ou suppression de lignes et de colonnes, calcul des statistiques sommaires et tracés graphiques.
Pour que la gestion éthique et efficace des données fonctionne pour vous!	1b	Présentation des mécanismes de soutien et des services de GDR, comme l'outil Assistant PGD de Portage, et discussion des pratiques exemplaires en matière de gestion des données pour gérer les données de façon sécuritaire et éthique tout au long de leur cycle de vie.
Préparation préalable (data wrangling) et versionnage des données avec Openrefine et GitHub	2a	Présentation de la plateforme GitHub comme un outil d'acquisition, de documentation et de versionnage des données et des codes, et Openrefine, un logiciel d'exploration, de nettoyage et de mise en forme des données à des fins d'interopérabilité.
Curation des données pour la reproductibilité	2b	Présentation des pratiques exemplaires en matière de recherche reproductible, des défis de la création et de la curation des données de recherche pour appuyer la reproductibilité et des outils comme OSF (Open Science Framework), les blocs-notes Jupyter et l'application Binder qui peuvent être utilisés pour organiser, documenter et partager des données, des codes et des environnements informatiques.
Modèle de curation des données du Digital Curation Network (DCN)	3a	Présentation de diverses étapes de la curation et de listes de contrôle normalisées qui aident à assurer que tous les ensembles de données reçoivent un traitement uniforme. Introduction à la curation des données et explication détaillée de la mise en œuvre pratique de disciplines ou de types de fichiers particuliers.
Curation des données dans les dépôts	3b	Présentation des principes FAIR et discussion des pratiques actuelles en matière de curation dans les dépôts généraux et les dépôts par domaine, partout au Canada, qui aident à améliorer la repérabilité, l'accessibilité, l'interopérabilité et la reproductibilité des données de recherche.

Les documents du forum, y compris les diapositives et le matériel de l'atelier (en anglais), peuvent être consultés au <https://data-curation.github.io/materials/>.

## Forum sur le développement communautaire

Un forum d'une journée et demie sur le développement communautaire a suivi les séances de formation. Les participants comprenaient des chercheurs, des spécialistes des données et des professionnels de l'information, des fonctionnaires et des administrateurs représentant un large éventail d'établissements de recherche, d'organisations non gouvernementales et d'organismes de recherche, de gouvernements et d'organismes de financement de partout au Canada. Le Forum avait pour objectifs de favoriser une compréhension collective du rôle de la curation des données dans le contexte plus vaste de la gestion et de l'intendance des données, d'exposer les défis de la curation, de déterminer les ressources qui appuieraient les pratiques exemplaires en matière de curation des données et de préciser les façons dont la coordination nationale des efforts de curation des données pourrait appuyer les besoins locaux dans ce domaine. Les participants ont assisté à trois séances plénières et, dans le cadre d'une série de séances en petits groupes, ont collaboré activement à l'élaboration d'une vision d'une approche nationale de la curation des données au Canada.

- Jeff Moon, directeur du réseau Portage, a donné un aperçu de l'évolution du paysage de l'infrastructure de recherche numérique (IRN) au Canada et de l'état du soutien national à la GDR, et décrit le rôle important de la curation des données dans cet écosystème. Le rôle de la curation des données dans les établissements a fait l'objet de discussions plus poussées par les membres du groupe d'experts Jeff Moon; Alexander Clark, vice-recteur associé à la recherche, Faculté des sciences infirmières de l'University of Alberta; et Mark Leggott, directeur général de Données de recherche Canada (DRC). Au cours de la séance de questions et réponses qui a suivi, les membres du groupe d'experts ont fait valoir le rôle complémentaire que les dépôts de domaine, les dépôts institutionnels et les dépôts nationaux jouent tous dans une approche nationale de la gestion des données, et ont reconnu l'importance d'intégrer les initiatives et les perspectives autochtones dans le paysage de la GDR.
- Lisa Johnston et Cynthia Hudson-Vitale ont donné aux participants un aperçu du développement et de la croissance du Data Curation Network (DCN) des États-Unis, un réseau interinstitutionnel coordonné de curateurs de données et de services de curation de données. Les partenaires du réseau collaborent pour créer et partager des procédures et des pratiques exemplaires en matière de curation des données et fournissent des services de curation en nature. Les jeux de données soumises au réseau sont confiées à un curateur possédant une expertise dans le type de fichier ou la discipline de recherche sur lesquels



portent les données en question, et celles-ci font l'objet d'une curation à l'aide d'un ensemble normalisé d'étapes (conformément aux listes de contrôle C-U-R-A-T-E-D)<sup>39</sup>. Les recommandations sont transmises à l'établissement local, lequel coordonne les changements nécessaires avec le chercheur. L'origine, la croissance et la viabilité du DCN ont été particulièrement pertinentes pour les activités de visualisation qui ont suivi, car elles fournissaient un modèle prospectif tangible pour l'établissement et l'évolution d'un réseau de curation des données au Canada.

- Des représentants de dépôts nationaux de données triés sur le volet ont fait le point sur leurs services et discuté des défis liés à la curation auxquels ils font face lorsqu'ils fournissent des données conformes aux principes FAIR sur leurs plateformes. Au cours de la séance, on a présenté des exposés tant sur les dépôts généraux [Dépôt fédéré de données de recherche (DFDR)] et Scholars Portal Dataverse (SPDV)] que sur les dépôts de domaine [Système intégré d'observation des océans du Canada (SIOOC)] qui visent à intégrer les données entre les établissements, les organismes et les programmes de recherche. Les discussions ont porté sur les différences et les similitudes importantes dans les approches de curation des données et les besoins des dépôts de données canadiens. Voir [en annexe](#) un aperçu des dépôts dont les participants ont discuté.

## Résultats du Forum sur le développement communautaire

Au cours des séances en petits groupes, les participants ont discuté des principaux obstacles et défis aux activités de curation des données au sein de leur organisme et dans leur travail. On a demandé à chaque groupe de concevoir une approche nationale de la curation des données qui permettrait de résoudre les problèmes soulevés. Les groupes ont décrit les ressources et le soutien nécessaires à la mise en œuvre de leurs modèles et ont déterminé les composantes qui devraient être coordonnées à l'échelle nationale, régionale ou locale. Les résultats de cet exercice sont résumés ci-dessous, ainsi que les recommandations pour un réseau canadien de curation des données.

### Obstacles et défis

#### *Capacités limitées*

Les capacités limitées pour ce qui est des travaux de curation préoccupent au plus haut point et tous les groupes estiment qu'il s'agit d'un obstacle. On a souligné le manque de temps en raison de

---

<sup>39</sup> Étapes et listes de contrôle C-U-R-A-T-E-D. <https://datacurationnetwork.org/resources/resources-2/>

priorités concurrentes et le manque de financement pour les travaux liés à la curation. De nombreux organismes n'ont pas de service officiel de GDR ou de personnel spécialisé pour effectuer les travaux de curation, et certains participants ont dit craindre que, si la demande de services de curation augmente dans leurs établissements locaux, on ne puisse pas y satisfaire.

### *Besoin de formation et d'orientation*

Le besoin de formation et de ressources pour appuyer les travaux de curation est une autre préoccupation pressante. Les curateurs sont en contact des données provenant d'un large éventail de disciplines et de sources, non seulement des données de recherche, mais aussi des données sur les actifs acquis et les ressources culturelles, des données relatives aux Autochtones, des données biomédicales et d'autres données sur la santé, des données anciennes et des données récupérées, des données et des métadonnées en français et en anglais, ainsi que des données générées à l'extérieur des établissements universitaires. Les participants ne se sentent pas encore prêts à faire face à l'ensemble unique d'exigences et de défis que présente chaque type de données. Il y a un besoin urgent de formation et d'autres possibilités d'éducation pour renforcer l'expertise dans le domaine de la curation; les participants ont aussi mentionné qu'il fallait élaborer des trousseaux d'outils, des flux de travail, des normes et des politiques pour appuyer les travaux de curation et éclairer les pratiques exemplaires.

### *Besoin intrinsèque d'une approche multilingue/internationale*

Le Canada est un pays multilingue qui compte de nombreux locuteurs du français, de l'anglais et des langues autochtones. Les participants reconnaissent le besoin de services de curation, de ressources et de possibilités de formation dans plusieurs langues et conviennent que, même s'il peut être difficile de répondre aux besoins en matière de gestion et de curation des données dans un pays multilingue, il est impératif que nous le fassions.

### *Manque de clarté au sujet des services existants, des rôles et des attentes en matière de rigueur*

Les participants aimeraient avoir une idée juste de la situation actuelle en ce qui a trait à la curation. Ils veulent mieux comprendre quels sont les services, les ressources et les possibilités qui existent et la façon de les exploiter à l'échelle locale. Compte tenu des ressources limitées dans de nombreux établissements, certains participants se demandent comment établir l'ordre de priorité des ensembles de données liées à la curation et si tous les jeux de données nécessitent le même niveau de curation. D'autres se demandent qui devrait assumer la responsabilité des travaux de curation et quelle devrait être la répartition des tâches à ce chapitre. Bien que la réponse puisse différer selon les établissements, les organismes, les archives et les dépôts, et dépendre du

nombre de membres du personnel disponibles et de leurs compétences individuelles, certaines précisions sur les rôles et les responsabilités des chercheurs ou de l'équipe de recherche, ainsi que sur les responsabilités de l'intendant des données ou de l'administrateur des données, peuvent être utiles.

### *Difficulté de communiquer la valeur*

Les participants ont également fait remarquer que la valeur de la curation peut être difficile à communiquer. Les défenseurs n'ont pas de message clair pour communiquer la valeur de la curation à divers groupes d'intervenants, entre autres aux chercheurs, aux administrateurs locaux et aux organismes de financement. Il y a peu de paramètres pour évaluer et démontrer l'incidence de la curation, de sorte qu'il est difficile d'encourager l'adhésion, compte tenu des priorités concurrentes pour obtenir des fonds et mettre à profit le temps du personnel.

### *Collaboration avec les chercheurs et d'autres curateurs à l'extérieur de notre communauté de pratique autodéclarée*

Les participants reconnaissent également la nécessité d'établir des ponts entre les curateurs et les bibliothécaires chargés de la gestion des données de recherche dans les bibliothèques universitaires, ainsi qu'entre les curateurs et les intendants des données des dépôts de domaine, des centres de données et d'autres organismes qui gèrent les données. Le premier défi pourrait consister à trouver divers intervenants susceptibles de contribuer à un réseau de curation et d'en tirer profit. Les chercheurs et les autres producteurs de données sont un groupe d'intervenants qui, selon les participants, doivent prendre une part active aux efforts de curation des données. Leurs commentaires seront précieux pour élaborer et offrir des services et des ressources qui répondront à leurs besoins en matière de gestion et d'archivage des données.

## Conceptualisation d'une approche nationale des services de curation

Pour relever ces défis, les participants ont envisagé une approche de curation des données coordonnée à l'échelle nationale qui tirerait parti du travail des organismes et des systèmes existants, comme Portage de l'ABRC, Données de recherche Canada (DRC), le Centre de gouvernance de l'information des Premières Nations (CGIPN), le SIOOC, les consortiums régionaux établis et la nouvelle organisation d'infrastructure de recherche numérique (NOIRN)<sup>40</sup>. Un réseau national de curation doit représenter un large éventail d'organismes, de points de vue et de voix. Ses travaux devraient être axés sur la collectivité et son objectif devrait être d'accroître et de renforcer les capacités de curation à l'échelle régionale et locale. Les participants ont demandé que

---

<sup>40</sup> <https://engagedri.ca/?lang=fr>

des mesures soient prises le plus rapidement possible et ont affirmé qu'une bonne partie du travail préparatoire peut être faite dès maintenant, même si un réseau de curation plus officiel est en voie d'être établi. À cette fin, les premières étapes de l'élaboration d'une stratégie nationale qui donnera aux particuliers et aux institutions les moyens de bâtir et d'accroître les capacités de curation sont proposées ci-dessous. Il est important de noter que chacun des objectifs proposés éclaire les autres et les relie entre eux, et qu'ils doivent être pris en compte de façon globale.

## Activités recommandées

### *Renforcer les capacités de curation à l'échelle locale*

- Investir dans le PHQ pour appuyer la gestion et la curation des données de recherche.
- Créer des occasions de formation et d'acquisition de compétences.
- Soutenir le perfectionnement professionnel et l'éducation.
- Élaborer et partager des normes, des flux de travail, des pratiques exemplaires, des politiques et d'autres directives pour éclairer la pratique à l'échelle locale.

Les capacités limitées en matière de travaux de curation en raison des priorités concurrentes sont considérées comme une préoccupation majeure. Il ne sera pas possible d'offrir des services de curation qui répondent aux besoins des chercheurs et d'autres intervenants sans un PHQ spécialisé. Pour veiller à ce que la main-d'œuvre soit prête à répondre à ces besoins, un service national devrait élaborer des possibilités de formation pour aider les PHQ existants à parfaire leurs compétences et inciter les établissements locaux à investir dans les PHQ. Des possibilités de perfectionnement professionnel et d'autres occasions de réseautage et d'acquisition de compétences pratiques en curation sont nécessaires à court terme. À plus long terme, les programmes de certificat et l'amélioration des programmes d'études pourraient offrir aux professionnels de l'information, aux spécialistes des données, aux intendants des données, aux nouveaux étudiants et aux chercheurs en début de carrière des possibilités de formation officielle qui leur donne les outils et les connaissances nécessaires pour relever les défis de la recherche savante contemporaine.

Les curateurs sont en contact avec une grande variété de données et de ressources numériques, qu'il s'agisse de données sur les actifs acquis et les ressources culturelles, de données relatives aux Autochtones, de données sur la santé et d'autres données recueillies auprès des participants humains, de données anciennes et de données récupérées, de descriptions et de données multilingues ou de données générées à l'extérieur des établissements universitaires. Il faut également mettre en commun des ressources comme des normes, des flux de travail et des politiques pour guider les curateurs et éclairer les pratiques exemplaires à l'échelle locale. Bien que cet ensemble de ressources puisse être géré à l'échelle nationale, il faudrait mettre en place des

mécanismes qui permettent aux curateurs et aux spécialistes des données locaux de participer au processus d'élaboration et de formuler des commentaires. Les ressources doivent être élaborées avec la participation d'une vaste communauté de pratique afin de répondre aux besoins de la collectivité. Les curateurs des établissements et des organismes locaux peuvent mettre à profit les pratiques exemplaires et les flux de travail pour éclairer et normaliser leurs travaux, mais ce sont eux qui ont le dernier mot sur les décisions prises à l'égard des ensembles de données qu'ils consomment.

### *Favoriser une communauté de pratique de la curation*

- Créer un cadre qui encourage les particuliers et les établissements à faire du réseautage.
- Mettre sur pied des mécanismes formels et informels permettant aux curateurs de communiquer entre eux et de participer à l'exploitation des ressources.
- Tenir compte de la diversité des types de données, des langues, des organismes et des points de vue.

Les curateurs, les gestionnaires de données de recherche et les intendants des données bénéficieraient d'occasions de réseautage formelles et informelles pour faire progresser la pratique de la curation. Une approche nationale de la curation exigera un cadre qui favorise la collaboration et une communauté de pratique qui va au-delà des établissements universitaires classiques pour mobiliser le groupe diversifié de personnes et d'établissements qui travaillent avec les données. Les serveurs de liste de diffusion, les plateformes de travail collaboratif comme Slack, les rencontres virtuelles et d'autres moyens de communication informels permettront à toute personne, à tout dépôt, à toute organisme ou à toute institution qui souhaitent communiquer, de poser des questions, d'échanger des idées et de participer au réseau. Les groupes de travail offrent aux particuliers et aux établissements un moyen plus officiel d'établir des relations et de collaborer pour trouver des solutions à des questions complexes, et des événements organisés, semblables à la formation en curation offerte lors du FCCD, permettront aux curateurs de mettre sur pied des réseaux en personne et de partager et d'acquérir des compétences pratiques en curation des données.

Un réseau national de curation des données doit reposer sur les principes d'équité et d'inclusion. Étant donné que les diverses données gérées seront variées, il doit y avoir une grande diversité dans les points de vue et les voix qui contribuent à l'élaboration des politiques et à l'exploitation des ressources. Les groupes d'intervenants habituels comprennent les établissements d'enseignement, les bailleurs de fonds, les administrateurs, les chercheurs, les dépôts de domaine et les dépôts institutionnels. Toutefois, un large éventail de collectivités et d'organismes qui participent à la curation, à la recherche, à la gestion ou à l'intendance des données, ne font pas

partie du milieu de la recherche universitaire classique, notamment les communautés autochtones, les organismes sans but lucratif, les ministères, les bibliothèques publiques et d'autres groupes communautaires. Un réseau national de curation devrait établir des liens et mobiliser ces communautés de parties prenantes afin que les nouvelles politiques et les ressources en matière de curation soient représentatives des divers établissements et organismes et puissent s'y appliquer. Ces groupes d'intervenants doivent également disposer de ressources qui leur permettent d'apporter une contribution égale et de participer pleinement au réseau.

Un réseau national devrait également explorer des façons de contribuer aux travaux des organismes internationaux comme la Research Data Alliance, le Digital Curation Network (DCN), le Digital Curation Centre (DCC), l'Australian Research Data Commons (ARDC), le Centre national de coordination de la gestion des données de recherche des Pays-Bas (LCRDM) et l'European Open Science Cloud (EOSC – nuage européen pour la science ouverte), par exemple, et d'en tirer parti. Une association professionnelle de curateurs de données donnerait à ces derniers une voix unifiée pour communiquer avec des entités comme la NOIRN et se faire entendre aux réunions nationales de GDR, et pourrait diriger l'élaboration des programmes d'études en plus de créer des occasions d'interagir et de réseauter avec des professionnels et des organismes internationaux du domaine de la curation. La mise en place de ressources et de services multilingues à l'échelle nationale offre également une occasion unique de tirer parti de nos travaux à l'échelle mondiale et de collaborer avec les communautés de pratique internationales.

### *Communiquer la valeur de la curation*

- Établir des paramètres pour mesurer la valeur ajoutée par la curation.
- Promouvoir la valeur de la curation et défendre la profession.
- Mobiliser les chercheurs, les producteurs de données et d'autres groupes d'intervenants qui bénéficieront de services de curation efficaces.

Pour renforcer les capacités de curation et susciter l'adhésion à l'échelle locale, il faut démontrer clairement la valeur de la curation. Il est essentiel d'obtenir l'adhésion des chercheurs, des producteurs de données et d'autres intervenants qui bénéficieront des services de curation et de consulter activement ces groupes d'intervenants à mesure que de nouveaux services sont mis au point pour faire en sorte que ceux-ci soient utiles et répondent à leurs besoins. Des messages clairs permettront aux curateurs, aux professionnels de l'information et aux intendants des données, entre autres, de communiquer la valeur de la curation à divers groupes d'intervenants locaux, dont les chercheurs et d'autres fournisseurs de données, les utilisateurs finaux, les services et les professionnels des technologies de l'information, les bailleurs de fonds, les administrateurs et les bureaux des services de recherche des établissements. Les professionnels de la gestion des

données pourraient recruter et engager des champions au sein de ces divers groupes d'intervenants pour aider à promouvoir plus largement la valeur de la curation.

Il faudrait évaluer et adopter des paramètres afin de mesurer l'incidence de la curation et établir une base de données probantes démontrant l'importance de la curation des données dans la GDR. Les résultats de ces efforts peuvent servir à défendre la profession et à faire valoir que la curation doit occuper une place primordiale dans le volet gestion des données de la NOIRN. Outre les paramètres qui mesurent la valeur ajoutée par la curation, les indicateurs de l'utilisation des données, les études d'impact et les histoires de réussites seront toutes importantes pour encourager le renforcement des capacités de curation dans les établissements locaux.

### *Développer et entretenir l'infrastructure partagée*

- Élaborer et tenir à jour des logiciels et du matériel essentiels à la curation et à la conservation des données.
- Établir un cadre dans lequel un réseau de curateurs peut communiquer et fonctionner.
- Investir dans le « facteur humain » nécessaire à une gestion efficace des données.
- Élaborer et coordonner les services de curation.

À l'échelle nationale, il est important de continuer à élaborer et à tenir à jour des options de dépôt qui fournissent des mécanismes de curation des données, comme Dataverse et le DFDR. Les ressources devraient être consacrées à des plateformes, outils, logiciels, environnements informatiques avancés, solutions de connectivité de réseau et services techniques partagés comme la production de DOI, qui peuvent tous être mis à profit par les établissements et les organismes locaux. Les plateformes nationales de dépôt devraient continuer de chercher du financement pour l'élaboration d'outils automatisés afin d'aider les curateurs dans leur travail, de créer des outils qui permettront de créer des liens avec l'infrastructure et les services de curation numérique, et développer l'infrastructure nécessaire pour soutenir les métadonnées multilingues pour les systèmes de découvertes.

Les curateurs et les intendants des données doivent être en mesure d'appliquer et de gérer les règles d'accès aux données. Il faudra investir dans une infrastructure de dépôt qui appuie la capacité de contrôler ou de restreindre l'accès aux données conformément aux lois sur la protection des renseignements personnels (données sur la santé) ou aux normes culturelles (données sur la propriété autochtone). Si la curation doit se faire dans l'ensemble des établissements et des organismes, les curateurs auront également besoin d'une méthode sécurisée pour accéder aux données et aux métadonnées et les partager à partir d'emplacements distribués.

La coordination nationale des pratiques exemplaires, des normes, des flux de travail et de l'élaboration des politiques peut également profiter aux établissements locaux. Même si l'élaboration de ces ressources doit être axée sur la collectivité et éclairée par les besoins des divers intervenants locaux, un service national pourrait servir de base de connaissances pour héberger et gérer non seulement ces ressources, mais aussi d'autres logiciels et outils de curation. Le SIOOC fournit un exemple de ce type de coordination, car dans le cadre de ce système ce sont les représentants régionaux qui déterminent les pratiques exemplaires et les normes en matière de métadonnées et veillent à leur conformité avec les décisions de la communauté océanographique internationale à ce chapitre<sup>41</sup>.

Il sera tout aussi important d'investir dans la composante humaine de la gestion des données. Ce sont des personnes qui bâtiront les systèmes et les réseaux, en assureront la connexion et l'entretien, mais les établissements auront diverses capacités de contribution et de nombreux établissements manquent actuellement de personnel spécialisé en GDR. Il faut en premier lieu tout mettre en œuvre pour renforcer les capacités locales et créer un cadre souple et évolutif dans lequel les curateurs peuvent fonctionner et communiquer. Les participants étaient ouverts à l'idée d'un réseau interinstitutionnel de curateurs de données afin de partager l'expertise, mais les établissements doivent y adhérer et il faut trouver un moyen précis d'assurer l'équilibre entre les besoins locaux et ceux du réseau avant d'envisager un modèle plus avancé de dotation partagée.

### *Représenter les curateurs et éclairer la gouvernance*

- Établir et maintenir un groupe consultatif national sur la curation qui interagit avec la communauté de pratique et les organismes nationaux sur les questions de gouvernance et de soutien de la curation des données.
- Collaborer avec les organismes de financement nationaux et les organisations d'infrastructure de recherche numérique pour élaborer des politiques et des lignes directrices en matière de curation dans des domaines complexes comme l'éthique, la gestion des droits relatifs aux données et la conservation à long terme.
- Élaborer des modèles de financement et une infrastructure qui assurent la viabilité du réseau.

Un lien avec la gouvernance (par l'entremise d'un comité ou d'un conseil consultatif) devrait être établi pour fournir une orientation stratégique et déterminer la portée des questions de curation, et faire valoir qu'il s'agit d'une composante essentielle de la gestion efficace des données. Un conseil consultatif ou une entité semblable permettra à la communauté de pratique de la curation

---

<sup>41</sup> Stewart et coll., « The Development of a Canadian Integrated Ocean Observing System (CIOOS) », *Frontiers in Marine Science*, vol. 6, 2019. <https://doi.org/10.3389/fmars.2019.00431>.



de communiquer d'une seule voix avec la NOIRN et d'autres organismes de gestion des données à l'échelle nationale et internationale. Comme dans le cas d'autres initiatives de soutien à la recherche, le travail du conseil consultatif devrait être axé sur la collectivité et ses membres devraient être représentatifs des divers intervenants de l'écosystème de recherche, y compris les chercheurs, les bibliothécaires et les intendants des données de divers dépôts de domaine, les dépôts nationaux et institutionnels et les centres de gouvernance de l'information des Premières Nations.

Une stratégie et des objectifs clairs seront utiles aux organismes lors de la mise en œuvre des services de curation. Le conseil consultatif pourrait assurer le leadership sur des sujets complexes comme l'éthique, les données sensibles, les protections de responsabilité, le droit d'auteur, les licences, les accords complexes de partage de données et la conservation numérique, et aider à élaborer des politiques pour éclairer les pratiques locales. Cette entité pourrait contribuer à promouvoir la valeur de la curation et proposer un modèle de réseau de curation viable à long terme.

Le gouvernement fédéral devra investir dans des services nationaux de curation des données, mais un réseau viable ne peut pas compter uniquement sur les fonds fédéraux. Le personnel chargé de la curation devra aider les chercheurs qui versent des données dans des dépôts nationaux comme le DFDR, mais il faudra aussi consacrer des ressources au renforcement des capacités de curation dans les établissements locaux. La mise en réseau d'une main-d'œuvre qualifiée et répartie permettrait de partager l'expertise et atténuer les effets de la réduction du financement ou la perte d'effectif à l'échelle nationale. D'autres travaux seront nécessaires pour définir la portée et l'orientation stratégique d'un réseau de curation des données et pour déterminer la façon dont un service national de curation des données pourrait être maintenu à long terme.

## Prochaines étapes et discussion

### Feuille de route

Le tableau suivant propose des activités précises à entreprendre à court, à moyen et à long terme pour répondre à chacun des objectifs décrits dans la section précédente. Il est à noter que les parties responsables de chaque activité ne sont pas définies, mais on s'attend à ce que ces activités soient coordonnées et exécutées par les composantes de la NOIRN.

*Objectif 1: Renforcer les capacités de curation des données (locales et nationales)*

<b>Court terme (1 an)</b>	<b>Moyen terme (2 ans)</b>	<b>Long terme (5 ans et plus)</b>
<p>Analyser les lacunes afin d'évaluer les capacités de curation actuelles, les ressources existantes ainsi que les besoins des chercheurs et des autres intervenants.</p> <p>Élaborer des flux de travail général en curation conformément aux pratiques exemplaires et en sollicitant les commentaires de la communauté.</p> <p>Solliciter les commentaires de la communauté sur les types de formation et de ressources en curation qui seraient utiles, et chercher des possibilités de financement.</p>	<p>Créer des contenus et organiser des ateliers afin d'offrir plus d'occasions d'acquisition de compétences pratiques et de réseautage.</p> <p>Développer et alimenter une base de connaissances sur la curation avec l'apport de la communauté.</p> <p>Embaucher et coordonner un PHQ pour organiser les dépôts nationaux et répondre aux besoins locaux en matière de curation.</p> <p>Explorer les options de formation avancée, comme la formation des formateurs, en vue de renforcer les capacités de façon exponentielle.</p>	<p>Élaborer et élargir les possibilités de formation officielle, comme l'amélioration des programmes d'études pour les professionnels de l'information et les chercheurs, et les programmes de certificat pour les professionnels en exercice.</p>

*Objectif 2: Favoriser une communauté de pratique*

<b>Court terme (1 an)</b>	<b>Moyen terme (2 ans)</b>	<b>Long terme (5 ans et plus)</b>
<p>Déterminer les façons dont les curateurs peuvent établir des liens informels (p. ex. rencontres virtuelles mensuelles, bureau de référence des curateurs, serveur de liste de diffusion, canal Slack).</p> <p>Créer des occasions pour les curateurs de collaborer, de travailler et d'apprendre dans leur langue maternelle.</p> <p>Chercher à trouver des intendants des données, des gestionnaires de données, des chercheurs et d'autres personnes participant aux travaux de curation qui peuvent travailler à l'extérieur de la bibliothèque ou du dépôt, ou à l'extérieur des milieux de recherche universitaires habituels.</p> <p>Établir des liens avec des communautés de pratiques connexes, comme la communauté de pratique de la conservation.</p>	<p>Créer des occasions de réseautage pour les curateurs, tant virtuelles qu'en personne.</p> <p>Sensibiliser et mobiliser les intendants des données, les gestionnaires des données et d'autres personnes qui exercent leurs activités à l'extérieur des milieux de recherche universitaires habituels.</p> <p>Établir des groupes de travail ou veiller à ce que la voix des curateurs soit représentée dans les groupes de travail existants afin d'aborder les questions de curation difficiles dans des domaines comme les données sensibles, l'éthique et la responsabilité, la sensibilisation et l'inclusion.</p>	<p>Créer une association professionnelle pour les curateurs de données ou un sous-groupe de curateurs dans une association professionnelle existante.</p> <p>Continuer d'appuyer les occasions de réseautage pour la communauté de la curation.</p> <p>Poursuivre la sensibilisation et élargir le réseau des curateurs.</p>

### Objectif 3: Communiquer la valeur de la curation

Court terme (1 an)	Moyen terme (2 ans)	Long terme (5 ans et plus)
<p>Commencer à constituer une base de données probantes sur la valeur ajoutée par la curation. Déterminer et évaluer les paramètres existants.</p> <p>Trouver des exemples de réussite et des champions des données dans le milieu de la recherche qui pourraient aider à faire valoir l'importance des services de curation des données.</p>	<p>Élaborer un ensemble de mesures d'évaluation et mettre à l'essai leur application aux processus de curation dans un dépôt existant comme le DFDR.</p> <p>Élaborer et promouvoir un message clair pouvant être utilisé pour faire valoir l'importance de la curation auprès de divers groupes d'intervenants.</p> <p>Collaborer activement avec les chercheurs et les créateurs de données pour promouvoir la valeur des services de curation et veiller à ce que les services fournis répondent aux besoins actuels et changeants.</p>	<p>Mettre en œuvre un ensemble normalisé de mesures pour illustrer la valeur de la curation dans les dépôts locaux, nationaux, institutionnels et de domaine.</p>

### Objectif 4: Développer et entretenir l'infrastructure

Court terme (1 an)	Moyen terme (2 ans)	Long terme (5 ans et plus)
<p>Continuer de chercher du financement pour l'élaboration d'outils de curation automatisés.</p> <p>Mettre en place une infrastructure pour gérer les canaux de communication informels des curateurs.</p> <p>Évaluer les étapes <u>C-U-R-A-T-E-D</u> et d'autres modèles de curation des données.</p> <p>Établir des points de connexion entre les flux de travail de curation et de conservation.</p> <p>Élaborer une base de connaissances centrale comme point de départ pour permettre aux curateurs d'accéder aux outils, aux ressources, aux documents, aux lignes directrices, aux liens vers des ressources externes, etc.</p>	<p>Encourager l'investissement local dans la curation des données par le PHQ.</p> <p>Explorer davantage les modèles de dotation interinstitutionnels ou régionaux pour les curateurs.</p> <p>Réfléchir à l'infrastructure la plus appropriée pour le partage des données entre les curateurs.</p> <p>Examiner l'infrastructure qui permettra aux curateurs de gérer divers types de données, comme les données à accès restreint ou les données visées par des politiques complexes de gestion des droits.</p> <p>Assurer l'intégration des dépôts à l'infrastructure et aux services de curation numérique.</p> <p>Trouver des méthodes pour appuyer les métadonnées multilingues dans les dépôts partout au Canada.</p>	<p>Mettre à l'essai un modèle de service interinstitutionnel de curation.</p>

## Objectif 5: Représenter les curateurs et éclairer la gouvernance

Court terme (1 an)	Moyen terme (2 ans)	Long terme (5 ans et plus)
Mettre sur pied un conseil consultatif ou un groupe de travail, représentatif de la multitude d'organismes qui œuvrent dans le domaine de la curation des données et gèrent les ressources numériques, afin d'assurer la surveillance et l'orientation stratégique.	<p>Le conseil consultatif fera valoir que la curation est un élément essentiel de la GDR et fournira une orientation stratégique dans des domaines clés comme l'élaboration de politiques, l'éthique, la responsabilité et la viabilité.</p> <p>Collaborer avec la communauté en vue d'élaborer des politiques de curation qui peuvent être utilisées pour éclairer les pratiques locales.</p> <p>Proposer un modèle de viabilité des services de curation en réseau.</p> <p>Chercher des possibilités de financement pour élargir les services et les ressources de curation.</p>	<p>Interagir avec les organismes internationaux de gestion des données dont les programmes comportent un volet curation et rechercher des occasions de collaboration.</p> <p>Veiller à ce qu'il y ait en place un mécanisme visant à examiner et mettre à jour les politiques liées à la curation afin de refléter les changements dans la culture, les technologies et les pratiques exemplaires.</p>

### Autres considérations

Les services de curation font partie intégrante de la GDR. Les établissements doivent être en mesure d'appuyer les efforts de gestion des données des chercheurs et doivent soit offrir des solutions pour la curation et le stockage des données elles-mêmes, soit diriger les chercheurs vers des organismes et des dépôts qui peuvent fournir ces solutions. L'investissement dans les services et le personnel de curation des données améliorera la conformité des résultats de la recherche au Canada aux principes FAIR, ce qui aura pour effet d'accélérer la recherche, de réduire le fardeau des chercheurs et d'augmenter le rendement global des investissements dans la recherche financée par l'État.

La mise en œuvre et l'expansion des services de curation posent un certain nombre de difficultés, comme en témoignent pour une bonne part les défis de la gestion des données de recherche dans son ensemble. Les rôles et les responsabilités ne sont pas clairement définis et peuvent varier d'un établissement à l'autre. Il faut offrir plus d'occasions de formation et

d'acquisition de compétences, élaborer des services, des outils et des ressources partagés et adopter des politiques et des normes pour orienter les pratiques exemplaires. De plus, un changement de culture qui met l'accent sur l'importance et la valeur des données de recherche et de leur gestion efficace est nécessaire<sup>42</sup>. Dans le cadre des séances en petits groupes au Forum et de discussions subséquentes, la communauté a souligné les nombreuses composantes et fonctions nécessaires à une approche nationale de la curation et déterminé un certain nombre de domaines où il faut approfondir le développement et les discussions. Comme cela a été le cas lors de la récente réorganisation nationale de la GDR et de l'IRN, la vaste participation de divers groupes d'intervenants a été essentielle à l'élaboration d'une approche rigoureuse et inclusive de la curation des données au Canada. Cette participation sera tout aussi importante à l'avenir.

Un réseau national de curation des données bien structuré doit s'efforcer de soutenir l'infrastructure fédérée tout en renforçant les capacités de curation et en soutenant les besoins en matière de curation à l'échelle locale. En même temps, les commentaires et la participation des membres des collectivités locales sont essentiels au maintien d'un réseau d'expertise en curation au Canada. Il faut donc une stratégie nationale bien définie pour former de nouveaux curateurs et parfaire les compétences des PHQ existants, tout en appuyant leur travail dans les établissements, les organismes et les dépôts et en leur permettant de participer à la création de politiques de curation, de normes et de flux de travail pour divers types de données et contextes. Un réseau d'experts de la curation, bien informés, hautement qualifiés et adéquatement répartis permettra à l'infrastructure de recherche numérique du Canada de répondre aux besoins des chercheurs dans leurs travaux actuels et futurs, et pourra aider à renforcer la résilience face à toute réduction éventuelle du financement (à l'échelle nationale ou locale) de la gestion des données de recherche. Un certain degré d'autonomie locale est également important. Bien que les pratiques exemplaires et les normes doivent servir à orienter les décisions locales, les établissements devraient conserver un certain pouvoir sur la façon dont elles sont mises en œuvre lors de la curation des ensembles de données qu'ils consomment ou gèrent.

Même si à l'heure actuelle on ne sait pas comment un tel réseau national sera géré et soutenu à long terme, on peut d'ores et déjà prendre des mesures pour éclairer les décisions futures. Quoique certains aspects de ce service puissent être appuyés par la NOIRN, il y a probablement moyen d'adopter des démarches susceptibles de fournir un soutien à l'échelle régionale et locale (au niveau institutionnel/organisationnel). Il faudrait s'efforcer dès maintenant de comparer les capacités et les besoins actuels du Canada en matière de curation — notamment les services, la

---

<sup>42</sup> Baker et coll., « Research data management in Canada ».

formation, le personnel et les ressources de soutien — et de proposer un modèle de partage des coûts, des services et des travaux de développement entre les intervenants. Il faut consulter activement les producteurs de données, les chercheurs et la communauté scientifique et faire en sorte qu'ils participent tout au long du processus. On avoir recours à des sondages et à des consultations individuelles ou en petits groupes pour évaluer la façon dont les services de curation actuels répondent à leurs besoins. Les résultats de ces mesures et les commentaires à leur sujet peuvent servir à éclairer l'expansion des services existants et l'élaboration de nouveaux services et de nouvelles ressources. Une analyse contextuelle qui donne une idée exacte de la conjoncture précisera les liens possibles, cernera les lacunes dans les ressources et les services et veillera à ce que les ressources limitées soient utilisées judicieusement. Les travaux en cours pour relier les établissements et les données dans les dépôts nationaux de données tels que le DFDR, SPDV et les dépôts de domaine comme le SIOOC fourniront également des renseignements précieux, à l'instar des efforts internationaux visant à créer des réseaux de ressources et de services de curation, comme le DCN, le DCC, et l'ARDC.

Comme une approche unifiée est en cours d'élaboration, il est important de tenir compte de la façon dont elle favorise l'inclusion, établit des liens entre divers établissements et organismes et assure l'équité entre les établissements dont les capacités de participation diffèrent. Par exemple, de nombreux dépôts institutionnels qui sont exploités partout au Canada reconnaissent le besoin actuel ou imminent de curateurs de données, mais manquent de ressources pour soutenir une équipe de curation locale entièrement fonctionnelle. Dans de tels cas, des modèles de dotation partagés pour l'expertise en curation, semblables à ceux du DCN, pourraient offrir une solution rentable. Des possibilités similaires existent à l'échelle nationale pour les dépôts comme le DFDR et la fédération de données dans Dataverse de Scholars Portal. Les options vont d'un modèle de dotation partagée selon lequel les curateurs basés dans un établissement sont détachés pour une petite partie de leur temps (p. ex. 5 à 10 %) au sein d'une équipe de soutien à la curation, financée par l'administration centrale, qui fournit des services à tous les établissements, à un modèle hybride regroupant les deux approches. Des approches de dotation partagées permettraient au réseau de prendre de l'expansion à mesure que les capacités locales se renforcent et que les dépôts de données augmentent, mais seraient probablement compliquées à financer et à soutenir. Il y a encore beaucoup à faire pour évaluer les capacités actuelles au Canada, explorer les options de financement et déterminer la façon dont les établissements locaux, y compris ceux qui n'ont pas de personnel spécialisé en GDR, pourraient participer à un tel réseau avant la mise en œuvre d'un projet pilote.

Des décisions importantes doivent également être prises en ce qui concerne les niveaux prévus de curation des ensembles de données, des métadonnées et des codes, et déterminer si toutes

les collections et tous les dépôts nécessitent un niveau égal d'attention et de rigueur. Si ce n'est pas le cas, comment devrait-on établir la priorité et qui devrait prendre ces décisions? Pour répondre à de telles questions, il faut tenir des discussions approfondies entre les groupes d'intervenants plus vastes de l'IRD afin de déterminer où se situent les besoins, de délimiter les capacités existantes et de définir les attentes des intervenants.

Enfin, des questions demeurent sans réponse au sujet de la structure organisationnelle et de la gouvernance en ce qui a trait à une approche nationale de la curation des données. À l'heure actuelle, il n'est pas clair comment un service national de curation s'intégrerait dans le paysage émergent de l'IRN, et comment il pourrait être géré à l'intérieur ou à l'extérieur de la NOIRN. Bien que ce document ait suggéré certains moyens par lesquels la communauté canadienne de la curation pourrait contribuer à la gouvernance et à l'élaboration de politiques en matière de GDR, une grande partie de cette contribution dépendra des structures et des mécanismes élaborés au sein de la NOIRN.

## Résumé des recommandations relatives à la NOIRN

La curation des données est une composante importante d'une saine gestion des données de recherche. La curation améliore la conformité des données aux principes FAIR, favorise la transparence du processus de recherche et augmente le rendement des investissements des bailleurs de fonds dans la recherche<sup>43</sup>. Bien que certains processus de curation puissent être automatisés, en raison de la nature complexe et hétérogène des données de recherche, la curation exige souvent l'intervention d'une personne possédant une combinaison unique de connaissances disciplinaires, de compétences en gestion de l'information, de capacités de collaboration et d'expertise en matière de logiciels. Bien que ce rapport présente des recommandations relatives à diverses activités qui pourraient être entreprises à tout moment à l'avenir, il y a un fort sentiment d'urgence au sein de la communauté. Étant donné que la production, la gestion et l'archivage des données sont en cours, la curation s'impose dès maintenant. Alors que nous travaillons en collaboration pour redéfinir l'écosystème de l'IRN du Canada, des mesures concrètes et efficaces peuvent être prises aujourd'hui pour répondre à un certain nombre de besoins actuels et jeter les bases d'une stratégie nationale de curation plus vaste. Ces mesures sont décrites ci-dessous :

---

<sup>43</sup> Canada, « Déclaration de principes des trois organismes sur la gestion des données numériques »; Innovation, Sciences et Développement économique Canada, « Un moment à saisir pour le Canada : Aller de l'avant dans le domaine des sciences, de la technologie et de l'innovation 2014 », 2014. [http://www.ic.gc.ca/eic/site/113.nsf/fra/h\\_07655.html](http://www.ic.gc.ca/eic/site/113.nsf/fra/h_07655.html)

- Reconnaître l'importance de la curation des données pour donner accès aux produits de données FAIR et au savoir libre ainsi que du rôle des curateurs de données dans ce processus.
- Reconnaître que la curation des données est une activité complexe et intensive qui exige un investissement important dans l'embauche et le perfectionnement de personnel hautement qualifié.
- Collaborer activement avec les chercheurs et les producteurs de données du Canada afin de mieux comprendre leurs besoins en matière de curation et les défis auxquels ils font face, et travailler avec ces communautés pour veiller à ce que les services de curation que nous développons les soutiennent.
- Contribuer à la création d'un cadre national pour les services de curation qui appuie les besoins de curation à grande échelle et répond à la diversité des besoins des divers organismes.
- Appuyer l'élaboration et la prestation de possibilités d'éducation, de ressources d'apprentissage et de documents d'orientation dans les deux langues officielles afin de favoriser le perfectionnement des curateurs de données partout au Canada.
- Collaborer avec les communautés autochtones pour veiller à ce que les services et les ressources de curation que nous développons soient en harmonie avec leurs efforts de gérance des données.
- Appuyer l'établissement et le maintien d'une communauté de pratique nationale sur la curation.
- Collaborer avec les communautés de la curation, de la conservation numérique et de l'éthique pour aborder des questions complexes liées à la curation et à la gestion à long terme des données biomédicales et d'autres données sur la santé, des données anciennes et des données récupérées, des données multilingues et des données générées à l'extérieur des établissements universitaires.
- Collaborer avec des spécialistes de la conservation numérique pour déterminer les mesures de gestion et de curation des données qui appuieront la conservation à long terme.
- Élaborer un mécanisme pour la communauté de la curation afin d'éclairer les décisions en matière de gouvernance et de politiques liées à la gestion des données de recherche.
- Appui direct ou indirect à la création de postes de curation nécessaires pour l'excellence des ensembles de données versés dans les dépôts de données canadiens.



## Bibliographie

Baker, David, Bourne-Tyson, Donna, Gerlitz, Laura, Haigh, Susan, Khair, Shahira, Leggott, Mark, Moon, Jeff et coll., *La gestion des données de recherche au Canada : état des lieux*, Zenodo, 13 décembre 2019. <https://zenodo.org/record/3574712#.Xrr8p0RKiUk>

CASRAI, s.v. « Data curation »; consulté le 2 février 2020. <https://casrai.org/term/data-curation/>

CASRAI, s.v. « Research Data »; consulté le 2 février 2020. <https://casrai.org/research-data/>.

Castle, David, Gerlitz, Laure, Haigh, Susan, Leggott, Mark, MacMullen, Lori, Moon, Jeff, Pirenne, Benoît, Tourangeau, Robbin et Bo Wandschneider, « Data Management Roadmap – 2019-2024 : For Submission to Innovation, Science and Economic Development (ISDE) », 29 mars 2019; consulté le 2 février 2020, <https://www.rdc-drc.ca/download/position-paper-data-management-roadmap-2019-2024/?wpdmdl=2482>.

CODATA et Paul Uhler, « The Value of Open Data Sharing : A CODATA Report for the Group on Earth Observations », Zenodo, 16 novembre 2015. <http://doi.org/10.5281/zenodo.33830>.

Coretrustseal Standards and Certification Board, « Coretrustseal Trustworthy Data Repositories Requirements 2020–2022 », Zenodo, 20 novembre 2019. <http://doi.org/10.5281/zenodo.3638211>

Digital Curation Centre, « Digital curation 101: Incentives for digital curation », Edinburgh, Digital Curation Centre, 2011; consulté le 10 janvier 2020, <http://www.dcc.ac.uk/sites/default/files/documents/DC%20101%20Incentives%20for%20Curati%20on.pdf>.

FAIRsharing, « FAIRsharing policies: A catalogue of data preservation, management and sharing policies from international funding agencies, regulators and journals »; consulté le 13 janvier 2020. <https://fairsharing.org/policies/>.

Centre de gouvernance de l'information des Premières Nations, « Ownership, Control, Access and Possession (OCAP™): The Path to First Nations Information Governance », Centre de gouvernance de l'information des Premières Nations, Ottawa, 2014. ISBN : 978-0-9879882-8-7; consulté le 17 février 2020. [https://fnigc.ca/sites/default/files/docs/ocap\\_path\\_to\\_fn\\_information\\_governance\\_en\\_final.pdf](https://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance_en_final.pdf)

Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, « Un moment à saisir pour le Canada : Aller de l'avant dans le domaine des sciences, de la technologie et de l'innovation 2014 », Innovation, Sciences et Développement économique Canada, 2014, Ottawa; modifié le 5 juillet 2018; consulté le 26 février 2020.

[http://www.ic.gc.ca/eic/site/113.nsf/fra/h\\_07655.html](http://www.ic.gc.ca/eic/site/113.nsf/fra/h_07655.html)

Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Bureau du sous-ministre, Direction générale des communications et du marketing, « Déclaration de principes des trois organismes sur la gestion des données numériques », Innovation, Sciences et Développement économique Canada, Ottawa, 21 décembre 2016; consulté le 10 janvier 2020.

[https://www.science.gc.ca/eic/site/063.nsf/fra/h\\_83F7624E.html](https://www.science.gc.ca/eic/site/063.nsf/fra/h_83F7624E.html)

Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Bureau du sous-ministre, Direction générale des communications et du marketing, « ÉBAUCHE – Politique des trois organismes sur la gestion des données de recherche – Aux fins de consultation », Innovation, Sciences et Développement économique Canada, Ottawa, 25 mai 2018; consulté le 10 janvier 2020.

[https://www.science.gc.ca/eic/site/063.nsf/fra/h\\_97610.html](https://www.science.gc.ca/eic/site/063.nsf/fra/h_97610.html)

Gouvernement du Canada, Innovation, Sciences et Développement économique Canada, Ministre de l'Innovation, des Sciences et du Développement économique, « Infrastructure de recherche numérique », Innovation, Sciences et Développement économique Canada, Ottawa, 7 août 2019; consulté le 10 janvier 2020. <https://www.ic.gc.ca/eic/site/136.nsf/fra/accueil>

Gouvernement du Canada, Bureau de la conseillère scientifique en chef du Canada, « Feuille de route pour la science ouverte », Bureau de la conseillère scientifique en chef du Canada, Ottawa, 26 février 2020; consulté le 22 avril 2020.

[https://www.ic.gc.ca/eic/site/063.nsf/fra/h\\_97992.html](https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97992.html)

Organisation internationale de normalisation (ISO), ISO 16363:2012, « Systèmes de transfert des informations et données spatiales — Audit et certification des référentiels numériques de confiance », 2012. <https://www.iso.org/fr/standard/56510.html>.

Johnston, Lisa R., Carlson, Jacob, Hudson-Vitale, Cynthia, Imker, Heidi, Kozlowski, Wendy, Olendorf, Robert et Claire Stewart, « How Important is Data Curation? Gaps and Opportunities

for Academic Libraries », *Journal of Librarianship and Scholarly Communication*, vol. 6, n° 1, 26 avril 2018, p. eP2198. <https://doi.org/10.7710/2162-3309.2198>

Johnston, Lisa R., Carlson, Jake, Hudson-Vitale, Cynthia, Imker, Heidi, Kozlowski, Wendy, Olendorf, Robert et Claire Stewart, « Data Curation Network: A Cross-Institutional Staffing Model for Curating Research Data », *International Journal of Digital Curation*, vol. 13, n° 1, 26 décembre 2018, p. 125-140. <https://doi.org/10.2218/ijdc.v13i1.616>.

Kim, Jeonghyun, Warga, Edward et William Moen, « Competencies Required for Digital Curation: An Analysis of Job Advertisements », *International Journal of Digital Curation*, vol. 8, n° 1, 14 juin 2013, p. 66-83. <https://doi.org/10.2218/ijdc.v8i1.242>.

Conseil du leadership sur l'infrastructure de recherche numérique, « Data Management Position Paper For Innovation, Science, and Economic Development Canada », 31 août 2017; consulté le 26 février 2020, <https://wiki.queensu.ca/display/QDSRWG>.

McQuilton, Peter, Sansone, Susanna-Assunta, Cousijn, Helen, Cannon, Matthew, Chan, Wei Mun, Carnevale, Ilaria, Cranston, Imogen et coll., « FAIRsharing Collaboration with Datacite and Publishers : Data Repository Selection, Criteria That Matter », Open Science Framework, 19 octobre 2019; consulté le 6 janvier 2020. <https://doi.org/10.17605/OSF.IO/N9QJ7>

Mons, Barend, *Data Stewardship for Open Science: Implementing FAIR Principles*, CRC Press, New York, 2018. <https://doi.org/10.1201/9781315380711>

Open Science Collaboration, « Estimating the reproducibility of psychological science », *Science*, vol. 349, n° 6251, 28 août 2015, aac4716. <https://doi.org/10.1126/science.aac4716>

Organisation de coopération et de développement économiques, « Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics », Éditions OCDE, Paris, 12 avril 2007. <https://doi.org/10.1787/9789264034020-en-fr>

Perrier, Laure, Blondal, Erik et Heather MacDonald, « The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis ». *PLoS ONE*, vol. 15, n° 2, 27 février 2020, e0229182. <https://doi.org/10.1371/journal.pone.0229182>

Groupe d'experts sur la curation des données de Portage, « Introduction à la curation des données », Bureau de la conseillère scientifique en chef du Canada, Réseau Portage, 2019;

consulté le 28 janvier 2020. [https://portagenetwork.ca/wp-content/uploads/2019/09/Curation\\_Primer\\_Aug2019\\_FR.pdf](https://portagenetwork.ca/wp-content/uploads/2019/09/Curation_Primer_Aug2019_FR.pdf).

Groupe de travail sur la stratégie des données de recherche, « Sommet 2011 sur les données de recherche canadiennes : Portrait de la situation des données », décembre 2011; consulté le 26 février 2020. <https://www.rdc-drc.ca/wp-content/uploads/f-Background-to-the-2011-Canadian-Research-Data-Summit.pdf>

Stewart, Andrew, deYoung, Brad, Smit, Mike, Donaldson, Karen, Reedman, Alexa, Bastien, Andréane, Ray Brunsting, Ray et coll., « The Development of a Canadian Integrated Ocean Observing System (CIOOS) », *Frontiers in Marine Science*, vol. 6, 9 août 2019, p. 431. <https://doi.org/10.3389/fmars.2019.00431>.

Stuart, David, Grace Baynes, Grace, Hrynaszkiewicz, Iain, Allin, Katie, Penny, Dan, Lucraft, Mithu et Mathias Astell. « Whitepaper: Practical challenges for researchers in data sharing » (version 1), Figshare, 21 mars 2018. <https://doi.org/10.6084/m9.figshare.5975011.v1>.

Tammaro, Anna Maria, Matusiak, Krystyna K., Sposito Frank Andreas et Vittore Casarosa. « Data Curator's Roles and Responsibilities: An International Perspective », *Libri*, vol. 69, n° 2, 31 mai 2019, p. 89-104. <https://doi.org/10.1515/libri-2018-0090>.

États-Unis, Office of Science and Technology Policy, « Request for Public Comment on Draft Desirable Characteristics of Repositories for Managing and Sharing Data Resulting From Federally Funded Research », Federal Register, Doc. 2020-00689, 85 FR 3085, 17 janvier 2020; consulté le 26 février 2020. <https://www.federalregister.gov/documents/2020/01/17/2020-00689/request-for-public-comment-on-draft-desirable-characteristics-of-repositories-for-managing-and>

Wilkinson, Mark D., Dumontier, Michael, Jan Aalsbersberg, IJsbrand, Appleton, Gabrielle Axton, Myles, Baak, Arie, Blomberg, Niklas et coll., « The FAIR Guiding Principles for scientific data management and stewardship », *Scientific Data*, vol. 3, 15 mars 2016, 160018. <https://doi.org/10.1038/sdata.2016.18>

# Annexe

## Liste des sigles

<b>Sigle</b>	<b>Appellation</b>
ARDC	Australian Research Data Commons
ABRC	Association des bibliothèques de recherche du Canada
FCCD	Forum canadien sur la curation des données
SIOOC	Système intégré d'observation des océans du Canada
DCC	Digital Curation Centre
DCN	Data Curation Network
PGD	Plan de gestion des données
DOI	Identifiant numérique d'objet
IRN	Infrastructure de recherche numérique
FAIR	Facilement repérables, accessibles, interopérables et réutilisables
CGIPN	Centre de gouvernance de l'information des Premières Nations
DFDR	Dépôt fédéré de données de recherche
LCRDM	Centre national de coordination de la gestion des données de recherche (Pays-Bas)
PHQ	Personnel hautement qualifié
ISDE	Innovation, Sciences et Développement économique Canada
NOIRN	Nouvelle organisation d'infrastructure de recherche numérique
PCAP <sup>®</sup>	Propriété, contrôle, accès et possession
CDR	Données de recherche Canada
GDR	Gestion des données de recherche
CRSH	Conseil de recherches en sciences humaines
SPDV	Scholars Portal Dataverse

## Répartition des participants au Forum sur le développement communautaire du FCCD

**Tableau 1 : Répartition par type d'organisme**

Type d'organisme	Nombre
Établissement d'enseignement universitaire - U15 (Canada)	21
Établissement d'enseignement universitaire (États-Unis)	13
Organisations non gouvernementales	10
Fournisseur de services de recherche nationaux et provinciaux <sup>‡</sup>	9
Organisme gouvernemental	5
Établissement d'enseignement universitaire (Canada; non membre d'U15)	2
Bibliothèque publique	1

‡ Comprend ABRC-Portage, Calcul Canada, Scholars Portal, Données de recherche Canada.

**Tableau 2 : Répartition par région géographique**

Région	Nombre
Ontario	29
Canada - Ouest	15
Canada - Est	11
Québec	4
International	2

**Tableau 3 : Répartition par poste**

Région	Nombre
Bibliothécaire/personnel des services de bibliothèque	31
Directeur	5
Gestionnaire	5
Analyste	4
Bibliothécaire universitaire adjoint/directeur associé	3
Gestionnaire/coordonnateur de l'éthique de la recherche	3
Gestionnaire/agent de projet	3
Curateur/spécialiste de la curation des données	3
Vice-président associé, Recherche	1
Bibliométricien	1
Spécialiste en informatique	1
Responsable des archives numériques	1

## Aperçu des dépôts nationaux de données présenté au FCCD

### *Systeme intégré d'observation des océans du Canada (SIOOC)*

Le SIOOC intégrera les données océanographiques canadiennes provenant d'associations régionales afin de rendre les ensembles de données interopérables accessibles aux utilisateurs finaux au moyen d'une plateforme Web nationale. Les associations régionales collaboreront avec les fournisseurs de données et les utilisateurs régionaux pour adopter des normes et des outils internationaux afin d'assurer la gestion et la diffusion des données interopérables. Les défis liés à la création du réseau du SIOOC comprenaient l'intégration de données complexes et diversifiées des trois associations régionales; la prise en compte des écarts dans les politiques, les procédures et les infrastructures locales; et l'établissement d'un cadre en vue de centraliser la gestion de projets répartis et les activités des groupes de travail. Ces défis et les solutions mises en œuvre par le SIOOC fourniront des renseignements précieux si les services de curation sont mis en réseau.

### *Dépôt fédéré de données de recherche (DFDR)*

Créé dans le cadre d'un partenariat entre le réseau Portage de l'ABRC et la Fédération de Calcul Canada, le DFDR favorise la repérabilité en recueillant des métadonnées de plus de 45 dépôts à l'échelle nationale et en exposant les résultats dans une interface de recherche fédérée. Il s'agit également d'un référentiel organisé pour les grands ensembles de données qui tire parti du service de transfert de fichiers Globus pour le téléversement et le téléchargement de données. Actuellement, la production du DFDR est limitée, mais on prévoit passer à la pleine production à l'automne 2020. À ce moment-là, tous les membres du corps professoral des établissements d'enseignement postsecondaire canadiens qui sont admissibles, ou leurs représentants, se verront allouer par défaut un important espace de stockage dans un dépôt structuré et auront accès à un processus pour demander plus d'espace de stockage au besoin.

### *Scholars Portal Dataverse (SPDV)*

Lancé en 2012, SPDV compte actuellement plus de 50 établissements membres partout au Canada et le nombre d'ensembles de données publiés par l'intermédiaire du service a augmenté au cours des deux dernières années. Dans le but de renforcer la communauté Dataverse et d'améliorer l'accessibilité des données pour les chercheurs et le public, Scholars Portal collabore avec la communauté afin de créer une version nationale de Dataverse à laquelle tous les établissements postsecondaires du Canada auront accès. Dans le cadre de l'expansion de ses services, Scholars Portal s'associe à Portage de l'ABRC pour améliorer l'infrastructure de stockage de Dataverse, contribuant ainsi à réduire les coûts de service pour les établissements à

l'échelle nationale. Grâce à ce partenariat, Scholars Portal hébergera et exploitera ce service, la surveillance et la coordination de la communauté nationale de la GDR étant assurées par Portage. Les travaux en cours sont axés sur l'élaboration d'ententes de partenariat avec des consortiums régionaux de bibliothèques ainsi que sur la définition de la structure et des politiques nationales de gouvernance.