

# Guide pour la curation dans Dataverse

1 octobre 2021

*Le Groupe de travail sur le guide de curation de Dataverse*

Préparé par le Groupe d'experts sur la curation du Réseau Portage  
pour la Nouvelle organisation d'infrastructure de recherche  
numérique

**NDRIO**  
New Digital  
Research Infrastructure  
Organization

**NOIRN**  
Nouvelle organisation  
d'infrastructure de  
recherche numérique

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

[engagedri.ca](http://engagedri.ca)

[portagenetwork.ca](http://portagenetwork.ca)

## Table de matières

Introduction .....	1
Scénarios possibles de services de curation .....	2
Niveaux de curation (soit la manière d'utiliser la liste de contrôle de CURATION).....	2
Guides complémentaires .....	3
Le Guide de CURATION .....	4
Consulter .....	5
Niveau 1.....	5
Un peu plus en profondeur .....	6
Niveau 2.....	6
Niveau 3.....	9
Recommander.....	14
Niveau 1.....	14
Améliorer .....	15
Niveau 2.....	15
Transposer .....	17
Niveau 3.....	17
Inclure.....	19
Niveau 2.....	19
Niveau 3.....	20
Optimiser.....	21
Niveau 3.....	21
Noter.....	22
Niveau 2.....	22
Références .....	23
Annexe 1 – Exemples d'ensembles de données avec curation effectuée dans Dataverse.....	26
Niveau 1.....	26
Exemples généraux.....	26
Annexe 2 – Modèles pour la correspondance .....	28
Annexe 3 – Ressources de curation supplémentaires.....	30
Annexe 4 – CURATION dans Dataverse : Guide éclair.....	31

## Introduction

La curation des données – la gestion active des données de recherche telles qu’elles sont créées, traitées, utilisées, archivées, partagées et réutilisées – est un élément essentiel de l'amalgame d'infrastructures, de processus, de schémas et d'expertise en curation qui composent les bonnes pratiques de gestion des données de recherche (GDR).<sup>1</sup> Un processus de curation des données bien exécuté peut améliorer des données en décrivant avec expertise leur contenu, en créant une structure cohérente, en fournissant une documentation significative, en permettant l'automatisation grâce au code et à la syntaxe ainsi qu’en établissant des liens avec d'autres données et d'autres résultats.

Ce guide offre des instructions en étapes pour la curation des nouveaux ensembles de données déposés dans Dataverse. La curation des données fait référence à la gestion active des données de recherche telles qu’elles sont créées, maintenues, utilisées, archivées, partagées et réutilisées.<sup>2</sup> Le guide est construit autour de l’acronyme CURATION pour aider les curateurs, surtout ceux qui commencent dans ce domaine, à se souvenir des étapes du processus de curation. Ce cadre est adapté [des étapes CURATED](#)<sup>3</sup> du Data Curation Network en fonction d’un contexte bilingue et vise à décrire et à fournir des conseils sur les meilleures pratiques de conservation dans Dataverse. La curation des données n’est pas toujours un processus linéaire ; le type de données avec lesquelles vous travaillez, les politiques ou pratiques de votre établissement, votre degré d’aisance avec la curation et le temps que le chercheur peut consacrer au processus peuvent vous obliger à sauter des étapes ou à les exécuter dans un ordre différent. Vous devrez peut-être aussi revenir en arrière et répéter certaines étapes. Compte tenu des priorités concomitantes et du nombre de personnes affectées au service de curation, le niveau de curation que votre établissement est en mesure d’offrir peut également déterminer le nombre d’étapes de curation que vous pouvez effectuer.

Il n’existe pas de modèle unique pour la curation des données. Le niveau et la qualité de la curation dépendra de certains facteurs, notamment des politiques institutionnelles et des objectifs stratégiques, de l’expertise et des capacités et de l’équilibre entre la capacité des ressources et la demande. C’est pourquoi le guide a été élaboré dans un souci de flexibilité. Il peut être utilisé par des curateurs débutants ou expérimentés au sein d’établissements universitaires de toutes tailles, et il peut être adapté par les établissements pour soutenir les politiques et procédures locales. Vous trouverez ci-dessous trois scénarios courants de services de curation des données, suivis des trois niveaux de curation possibles sur lesquelles nous nous sommes basés pour élaborer ce guide.

---

<sup>1</sup> La gestion des données de recherche (GDR) désigne les processus appliqués tout au long du cycle de vie d'un projet de recherche pour guider la collecte, la documentation, le stockage, le partage et la préservation des données de recherche (gouvernement du Canada 2018).

<sup>2</sup> [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>3</sup> <https://datacurationnetwork.org/outputs/workflows/> and <https://doi.org/10.2218/ijdc.v13i1.616> [en anglais]

## Scénarios possibles de services de curation

<b>Curation non médiatisée</b>	Le service de gestion des données de recherche (GDR) n'intervient pas. Le chercheur crée son propre espace dataverse et son propre ensemble de données, soumet ses données et les publie.
<b>Curation semi-médiatisée</b>	L'équipe de gestion des données de recherche crée un espace dataverse ou commence le dépôt d'un ensemble de données et attribue un rôle au chercheur. Le chercheur dépose ses données dans son espace dataverse (ou ensemble de données). Selon la politique locale, soit l'ensemble de données est envoyé pour examen par les administrateurs de l'espace Dataverse de l'établissement, soit le déposant demande à ce que l'ensemble de données soit examiné par l'équipe de gestion des données avant ou après sa publication.
<b>Curation médiatisée</b>	Les données sont soumises par le chercheur au service de gestion des données de recherche. Le service de gestion des données de recherche crée l'espace dataverse, ou seulement l'ensemble de données ; les données sont ensuite préparées par la bibliothèque et publiées après avoir été approuvées par le chercheur.

Tableau 1 : Scénarios possibles de services de curation

## Niveaux de curation (soit la manière d'utiliser la liste de contrôle de CURATION)

La liste de contrôle de CURATION est divisée en 3 niveaux : niveau 1, niveau 2 et niveau 3. Le niveau 1 est l'information de base obligatoire pour publier un ensemble de données dans Dataverse. Selon le niveau de service que votre établissement est en mesure d'offrir,<sup>4</sup> vous pourrez compléter certains éléments des niveaux deux et trois. Voici une description de chaque niveau :

<b>Niveau 1</b>	Ce niveau de curation prépare les ensembles de données à être publiés avec succès dans Dataverse. Il s'agit du niveau de curation minimum nécessaire pour publier avec succès dans dataverse et pour rendre l'ensemble de données facile à repérer.
<b>Niveau 2</b>	Ce niveau de curation améliore la découvrabilité de l'ensemble de données et en favorise la réutilisation. Par exemple, les champs de métadonnées sont complétés et l'ensemble de données est suffisamment documenté pour en permettre la réutilisation.
<b>Niveau 3</b>	Ce niveau de curation prépare les ensembles de données pour la reproductibilité et pour la préservation. Par exemple, la documentation est complète, les scripts et les codes-sources sont analysés et, lorsque possible, les fichiers en format propriétaire sont convertis en format non protégé et adapté pour la préservation long terme.

Tableau 2 : Niveaux de curation

<sup>4</sup> Pour plus d'informations sur les niveaux de curation et la manière dont ils ont été mis en pratique dans deux établissements, veuillez consulter Lafferty-Hess, Sophia, Julie Rudder, Moira Downey, Susan Ivey, Jennifer Darragh, and Rebekah Kati. Conceptualizing Data Curation Activities Within Two Academic Libraries. *Journal of Librarianship and Scholarly Communication* 8, no. 1 (19 juillet 2020) : eP2347. <https://doi.org/10.7710/2162-3309.2347>. [En anglais].

## Guides complémentaires

Le présent Guide est complémentaire à d'autres ressources de curation de données, y compris les directives propres à Dataverse. Les guides ci-dessous approfondissent certaines caractéristiques de la plateforme Dataverse ainsi que les considérations relatives à certains formats de fichiers et proposent des bonnes pratiques pour les métadonnées dans Dataverse. Ils seront particulièrement utiles pour les étapes de curation de niveaux 2 et 3, ou comme référence lorsque vous travaillez avec des formats de données moins communs.

### [Guide des pratiques exemplaires sur les métadonnées de Dataverse Nord](#)<sup>5</sup>

Ce guide bilingue fournit des définitions pour chaque champ de métadonnées dans Dataverse, ainsi que des exemples et des conseils pour s'assurer que les valeurs appropriées sont incluses et formatées correctement. Il distingue les champs de métadonnées obligatoires, recommandés et facultatifs et peut être utilisé pour permettre une description plus riche, au-delà de la portée du présent Guide.

### [Guide sur le dataverse de Scholars Portal](#)<sup>6</sup>

Ce guide bilingue pour le dépôt de données présente les étapes à suivre pour déposer un ensemble de données dans le dépôt de données Dataverse de Scholars Portal, la plus grande instance de Dataverse au Canada. Il explique comment utiliser Dataverse, entre autres comment créer et modifier les ensembles de données. Pour des conseils avancés, consultez le [Harvard Dataverse Project User Guide](#)<sup>7</sup> [en anglais].

### [Curation Workflow du Data Curation Network](#)<sup>8</sup>

Ce guide présente les étapes et propose des listes à cocher qui décrivent le processus de curation de données d'une façon normalisée, quel que soit le type de données ou la plateforme de dépôt. Le cadre « CURATED » a été adapté en « CURATION ».

### [Data Curation Network Data Primers](#)<sup>9</sup>

Élaborés par la communauté qui gravite autour du Data Curation Network, ces documents regroupent des ressources diverses qui aident la curation de divers formes de données. Ces documents fournissent des conseils supplémentaires sur les formats de fichier et sur les exigences logicielles ainsi que des informations qui ne sont pas mentionnées dans les autres guides.

---

<sup>5</sup> <http://hdl.handle.net/2429/73610>

<sup>6</sup> <https://learn.scholarsportal.info/fr/guides/dataverse/>

<sup>7</sup> <https://dataverse.scholarsportal.info/guides/en/latest/user/> [en anglais]

<sup>8</sup> <https://datacurationnetwork.org/outputs/workflows/> [en anglais]

<sup>9</sup> Disponible sur GitHub <https://github.com/DataCurationNetwork/data-primers> et en format PDF <https://hdl.handle.net/11299/202810> [en anglais]

## Le Guide de CURATION

<b>C</b>	<a href="#">Consulter</a> S'assurer que toutes composantes des données et des métadonnées nécessaires à la bonne marche de la publication de l'ensemble de données sont présentes.
<b>U</b>	<a href="#">Un peu plus en profondeur</a> Faire en sorte que l'ensemble de données soit bien décrit et que les futurs utilisateurs aient une idée précise de la nature des données et de la manière dont elles peuvent être utiles.
<b>R</b>	<a href="#">Recommander</a> Demander des informations supplémentaires à la personne qui dépose et suggérer des modifications des métadonnées ou des fichiers qui faciliteront le repérage et la réutilisation des données dans la perspective des principes FAIR.
<b>A</b>	<a href="#">Améliorer</a> Améliorer la soumission pour faciliter le repérage et la réutilisation de l'ensemble de données.
<b>T</b>	<a href="#">Transposer</a> S'assurer que l'ensemble de données emploie des formats de fichier aussi ouverts et aussi courants que possible.
<b>I</b>	<a href="#">Inclure</a> Faciliter la réutilisation, la bonne citation et la reconnaissance de la source des données en incluant les identifiants pérennes pertinents et les informations nécessaires sur les licences.
<b>O</b>	<a href="#">Optimiser</a> Évaluer la conformité globale de l'ensemble de données aux principes FAIR et prendre les mesures nécessaires pour optimiser la facilité de repérage, l'accessibilité, l'interopérabilité et la réutilisation de l'ensemble de données.
<b>N</b>	<a href="#">Noter</a> Consigner par écrit les détails de la démarche du travail de curation.

## Consulter

Le but de l'étape **Consulter**<sup>10</sup> est de faire en sorte que toutes les données et métadonnées **requis** par le système pour publier avec succès la soumission soient présentes et fonctionnent correctement. Vous devez également essayer de recenser toute caractéristique nécessitant une attention particulière (par exemple, les données sensibles et les données d'entités externes).

### Niveau 1

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'ensemble de données a été soumis dans l'espace dataverse approprié.  Pour déterminer si cet ensemble de données a été déposé dans l'espace dataverse approprié, il faut tenir compte des éléments suivants : <ul style="list-style-type: none"><li>▪ Le déposant (ou le groupe de recherche du déposant) a-t-il déjà créé et déposé dans un espace dataverse ?</li><li>▪ L'ensemble de données nécessite-t-il son propre espace Dataverse ou existe-t-il un espace dataverse associé auquel il est lié ?</li><li>▪ L'ensemble de données est-il conforme aux politiques et aux normes de soumission de cet espace dataverse particulier (par exemple, un répertoire géré par la bibliothèque) ?</li></ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le chercheur a confirmé que l'ensemble de données ne pose pas de problème de licence ou de propriété intellectuelle.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le chercheur a confirmé que l'ensemble de données est exempt d'informations identificatoires/sensibles. (C'est-à-dire les informations devant être protégées contre un accès ou une divulgation non autorisée.)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	La documentation de soutien est incluse. <ul style="list-style-type: none"><li>▪ Exemples : guide codification, dictionnaire de données, méthodologie, fichier README, etc.</li></ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Tous les fichiers mentionnés dans la documentation sont inclus dans l'ensemble de données.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les champs de métadonnées requis sont exacts. <ul style="list-style-type: none"><li>▪ Utilisez le <a href="http://hdl.handle.net/2429/73610">Guide des pratiques exemplaires sur les métadonnées de Dataverse Nord</a><sup>11</sup> pour évaluer l'exhaustivité des champs de métadonnées requis (titre, nom d'auteur, courriel de la personne-ressource, description, sujet).</li></ul>

<sup>10</sup> Notez que les deux premières étapes — Consulter et Un peu plus en profondeur — se recoupent. Les tâches de l'étape Consulter sont toutes des tâches de niveau 1 et doivent être accomplies pour publier un ensemble de données dans Dataverse. Les tâches de l'étape Un peu plus en profondeur sont de niveau 2 et 3 et peuvent être effectuées en fonction du niveau de service de votre établissement.

<sup>11</sup> <http://hdl.handle.net/2429/73610>

## Un peu plus en profondeur

Le but de l'étape **Un peu plus en profondeur**<sup>12</sup> est de s'assurer que l'ensemble de données est bien décrit et que les utilisateurs finaux auront une idée claire de ce que sont les données et de la manière dont elles peuvent être utilisées. Examiner les métadonnées et la documentation pour en vérifier la rigueur et la clarté, créer ou recommander une documentation supplémentaire au besoin et vérifier les problèmes d'utilisation tels que les données manquantes, les échecs d'exécution de code, les titres ambigus et les problèmes de présentation des données. Vous pouvez également vérifier les risques de divulgation de renseignements personnels, les violations des droits de propriété intellectuelle, etc. selon les politiques de votre établissement et du niveau de service de curation fourni dans votre espace.

Pour un aperçu complet des mesures que vous pourriez prendre pour comprendre les données et évaluer leur exhaustivité et leur facilité d'utilisation, voir étape 3.0 dans [Curating Research Data Volume Two: A Handbook of Current Practice](#)<sup>13</sup> [en anglais]. Les documents [Curation Primers](#)<sup>14</sup> du Data Curation Network [en anglais] constituent également une excellente source d'informations et de conseils sur les meilleures pratiques en matière de curation de divers types de fichiers. Ils comportent des informations utiles sur les outils que vous pouvez utiliser pour examiner les fichiers et sur les types d'informations qui seront nécessaires pour garantir la facilité d'utilisation des données au fil du temps.

## Niveau 2

### La documentation de soutien est rigoureuse, précise et complète

Si l'ensemble de données comporte un fichier README, un guide de codage, des manuels d'utilisation ou toute autre documentation, vérifiez l'exactitude et l'exhaustivité de ces documents. La documentation doit offrir des informations contextuelles sur l'ensemble de données afin de faciliter son utilisation. Si la documentation est inadéquate, travaillez avec le chercheur pour améliorer la documentation existante ou fournissez-lui des modèles et des conseils pour lui permettre de créer de la documentation. Les guides suivants [Guide to writing « readme » style metadata](#) de l'Université Cornell<sup>15</sup> et [Créer un fichier README pour vos ensembles de données](#) de l'Université de Colombie-Britannique<sup>16</sup> sont des ressources utiles pour les curateurs et les chercheurs.

---

<sup>12</sup> Notez que les deux premières étapes — Consulter et Un peu plus en profondeur — se recoupent. Les tâches de l'étape Consulter sont toutes des tâches de niveau 1 et doivent être accomplies pour publier un ensemble de données dans Dataverse. Les tâches de l'étape Un peu plus en profondeur sont de niveau 2 et 3 et peuvent être effectuées en fonction du niveau de service de votre établissement.

<sup>13</sup> <https://hdl.handle.net/11299/185335> [en anglais]

<sup>14</sup> Disponible sur GitHub <https://github.com/DataCurationNetwork/data-primers> et en format PDF <https://hdl.handle.net/11299/202810> [en anglais]

<sup>15</sup> <https://data.research.cornell.edu/content/readme> [en anglais]

<sup>16</sup> <https://doi.org/10.5281/zenodo.4058961>

Oui	Non	Certains problèmes	S.O.	La documentation inclut :
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Des informations contextuelles sur les données (la manière dont les données ont été collectées ou générées, le but de la recherche).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les conventions de nommage et la structure des fichiers, si nécessaire.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Une description de la manière dont les données ont été modifiées ou traitées.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Des informations sur la confidentialité et sur les restrictions concernant l'utilisation secondaire des données.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les noms des étiquettes et des variables, des valeurs autorisées et des unités de mesure, et des codes et des classifications, le cas échéant.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'environnement informatique nécessaire pour exécuter les codes inclus (système d'exploitation, bibliothèques logicielles et toute autre dépendance).

### Les fichiers s'ouvrent et les contenus s'affichent correctement

Téléchargez l'ensemble de données et décompressez les fichiers archivés ou compressés. Examinez le contenu et réglez tous les problèmes liés aux fichiers en format propriétaire.

**Conseil :** Pour les types de fichiers peu connus, la documentation fournit-elle des indications sur le logiciel utilisé pour générer le fichier ou sur la manière de le visualiser ? Essayez d'ouvrir le fichier dans un éditeur de texte, car même les fichiers binaires peuvent avoir des titres en texte clair qui fournissent des informations sur l'instrument ou le logiciel qui a généré le fichier.

S'il n'est pas possible d'ouvrir tous les fichiers, on recommande de vérifier un sous-ensemble qui contient :

- au moins un de chaque type de fichier dans l'ensemble de données soumis ;
- des fichiers de script, du code et tout ce que vous soupçonnez d'être protégé par une licence ;
- tous les fichiers qui peuvent comporter des informations sensibles.

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les fichiers s'ouvrent correctement et le contenu de dossiers compressés est extrait sans problème. Un logiciel inconnu, des versions incompatibles ou l'absence d'accès au logiciel peuvent être des raisons pour lesquelles les fichiers ne s'exécutent pas.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si les fichiers ne sont pas accessibles, le chercheur a-t-il fourni une version non propriétaire des fichiers ?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si une version non propriétaire des fichiers ne peuvent pas être fournis, le chercheur a-t-il fourni un fichier README qui décrit les données et le logiciel nécessaire pour les utiliser ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le contenu des fichiers est conforme à la structure et à l'encodage.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le fichier README décrit les fichiers de données et comprend les éléments suivants : <ul style="list-style-type: none"> <li><input type="checkbox"/> Description sommaire</li> <li><input type="checkbox"/> Format de fichier (mentionnés si le format est propriétaire)</li> <li><input type="checkbox"/> Taille de fichier</li> <li><input type="checkbox"/> Chemin et/ou arborescence<sup>17</sup></li> <li><input type="checkbox"/> Une somme de contrôle (<i>checksum</i> en anglais) qui peut être utilisée pour vérifier l'intégrité des données (par exemple, md5, sha256).<sup>18</sup></li> </ul>

### Les fichiers et dossiers sont nommés et structurés correctement

Idéalement, les fichiers sont nommés et organisés d'une manière compréhensible et navigable par un utilisateur final potentiel. La structure des répertoires doit être suffisamment simple pour permettre de trouver les fichiers et les répertoires doivent être nommés de manière à communiquer clairement leur contenu. Notez que les critères suivants (fournis par [l'Université de Victoria](#)<sup>19</sup>) sont généraux et que les conventions peuvent différer selon le type de données et la discipline. Pour plus d'informations, le guide de l'Université d'Ottawa sur le [nommage de fichier et la gestion de versions](#)<sup>20</sup> peut être utile.

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les noms de fichiers sont exempts d'espaces ou de caractères spéciaux et utilisent des traits de soulignement ou des tirets comme délimiteurs (par exemple Datacollection_20201009_v02.csv).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les noms de fichiers et de répertoires sont concis, cohérents et compréhensibles.

<sup>17</sup> Envisager l'utilisation d'outils pour automatiser cette production :

- Windows : utilisez [la fonction d'arborescence « tree »](#) de la ligne de commande
- MAC : utilisez [Homebrew](#) pour installer [le paquet d'arborescence « tree »](#)
- Linux : installez et utilisez [la commande d'arborescence « tree »](#)

<sup>18</sup> Envisager l'utilisation d'outils pour automatiser cette production :

- Windows : utilisez le the programme [HASHMYFILES](#)
- MAC : utilisez la fonction « md5 » ou « shasum -a 256 » de la ligne de commande
- Linux : utilisez la fonction « md5sum » ou « sha256sum » de la ligne de commande

<sup>19</sup> [https://libguides.uvic.ca/ld.php?content\\_id=35154390](https://libguides.uvic.ca/ld.php?content_id=35154390) [en anglais]

<sup>20</sup> <https://biblio.uottawa.ca/fr/services/professeurs/gestion-donnees-recherche/designation-fichiers-disposition-donnees>

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les dates dans les noms de fichiers utilisent un formatage uniforme (par exemple, AAAAMMJJ).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les noms de fichiers utilisent des zéros de tête pour les numéros de version (par exemple v_012).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les fichiers sont regroupés dans une structure de dossiers logique. La structure des dossiers peut dépendre du code, de la syntaxe ou du produit d'un logiciel.

## Niveau 3

### Le code est bien commenté et produit les résultats voulus

Le déposant peut avoir inclus des fichiers de script qui ont été utilisés pour traiter ou analyser les données, du code qui étend un modèle existant, des fichiers exécutables ou d'autres logiciels. Bien qu'il puisse être avantageux de conserver les fichiers de script avec les données, il existe de nombreuses options de dépôt spécialement conçues pour le code et les logiciels qui disposent de systèmes de gestion des versions robustes et permettent un développement continu. Vous pouvez suggérer une autre solution pour le code de publication et d'archivage, bien que vous puissiez néanmoins souhaiter examiner le code ou le logiciel et les métadonnées ou la documentation associées pour votre dépôt.

Oui	Non	Certains problèmes	S.O.	Si l'ensemble de données comprend un logiciel ou un code...
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le code est-il également disponible dans GitLab, GitHub, Bitbucket ou un autre dépôt construit à cet effet ? <ul style="list-style-type: none"> <li>Le cas échéant, vous pouvez demander au déposant de l'archiver dans les archives <a href="https://www.softwareheritage.org">Software Heritage</a><sup>21</sup> ou <a href="https://zenodo.org">Zenodo</a><sup>22</sup> et de le relier à partir de l'enregistrement des métadonnées de l'ensemble de données.</li> <li>Si ce n'est pas le cas et que le chercheur a inclus d'autres éléments en plus des scripts de traitement ou d'analyse de données (par exemple des fichiers .r ou .m ou des fichiers .do), suggérez un dépôt spécialement conçu à cet effet.</li> </ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le code (ou une partie du code) est-il dérivé d'une autre source ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si le code est dérivé d'une autre source, la licence choisie est-elle compatible avec la licence de la source originale ?

<sup>21</sup> <https://www.softwareheritage.org/save-and-reference-research-software/?lang=fr> [en anglais]

<sup>22</sup> <https://guides.github.com/activities/citable-code/>

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	S'il est exécuté, le code source produit-il sans erreur les résultats attendus ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si un fichier exécutable est inclus, le code source est-il également disponible ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Le code est bien commenté :</p> <ul style="list-style-type: none"> <li>▪ Le code contient des informations telles que : auteur, numéro de version, nom de fichier, licence.</li> <li>▪ La fonction ou l'objectif du code est clair (selon le fichier README ou la description intégrée).</li> <li>▪ Le cas échéant, le déposant a inclus des informations sur la manière d'exécuter le code.</li> <li>▪ Les dépendances nécessaires sont énumérées.</li> <li>▪ Les commentaires sont concis et clairs et décrivent l'intention des lignes de code qui suivent, OU</li> <li>▪ Le code est expressif (peut être compris par les humains et les machines).</li> </ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Le fichier README ou le code comprend des informations sur :</p> <ul style="list-style-type: none"> <li>▪ La licence.</li> <li>▪ Le nom et les coordonnées du développeur.</li> <li>▪ La version et la date de la dernière modification ou exécution du code.</li> <li>▪ Toutes les références pour le code (ou une partie du code) qui a été repiquée.</li> <li>▪ Des informations sur les entrées requises et les sorties attendues.</li> <li>▪ Des instructions pour l'exécution du code. S'il existe plusieurs fichiers de script, l'ordre dans lequel ils sont exécutés doit être clair.</li> <li>▪ Toutes les dépendances nécessaires.</li> <li>▪ Des informations sur l'environnement dans lequel le code a été développé ou peut être utilisé.</li> </ul>

## La soumission contient des éléments potentiellement sensibles

**Note :** La plupart des instances de Dataverse peuvent restreindre l'accès aux ensembles de données, mais ne peuvent pas fournir une enclave de données convenable pour les données sensibles. Consultez votre administrateur Dataverse pour confirmer la capacité de votre Dataverse à cet égard.

Les données potentiellement sensibles devraient être signalées par les chercheurs avant le dépôt, mais ils ne les signaleront pas systématiquement. Le titre ou la description de l'ensemble de données ou la présence de formulaires de consentement des participants ou d'accords de participation peuvent vous aider à repérer un ensemble de données qui mérite une attention particulière, mais certaines données sensibles ne sont pas aussi « évidentes » au moment du dépôt et de la curation. Cela comprend, notamment les données collectées auprès des populations autochtones, ou sur leurs terres, leurs ressources et leurs environnements; les connaissances traditionnelles ; les données collectées sur des propriétés privées ; les données indiquant l'emplacement d'espèces vulnérables ou de sites protégés ; les données propriétaires ou collectées avec un partenaire industriel ; les données réutilisées ou fournies par un tiers ; et d'autres données dont la collecte ou la publication est soumise à un accord de partage de données.

Vous pouvez signaler les contenus qui semblent être en violation avec les conditions d'utilisation de votre espace et travailler avec les chercheurs pour limiter les risques de divulgation. Si vous le faites, indiquez clairement que la responsabilité finale du respect des conditions d'utilisation, du consentement des participants ou de tout accord de partage des données incombe à l'équipe de recherche..

Notez que la manière dont vous traitez les données sensibles peut dépendre de considérations ou de politiques locales, ce qui peut même empêcher les curateurs de déterminer la sensibilité des données sans avis extérieur. De nombreux intervenants peuvent apporter leur soutien, notamment votre comité d'éthique de la recherche, votre bureau de recherche, votre bureau des relations avec les autochtones, votre bureau de liaison avec l'industrie et d'autres services associés au soutien de la recherche sur votre campus.

## Ressources de données sensibles :

- Groupe en éthique de la recherche du Gouvernement du Canada : [Lignes Directrices pour verser des données existantes dans des dépôts publics](#)<sup>23</sup>
- Boîte à outils du réseau Portage pour les données sensibles – destiné aux chercheurs
  - Partie 1 : [Glossaire terminologique sur l'utilisation des données sensibles à des fins de recherche](#)<sup>24</sup>
  - Partie 2 : [Matrice de risque lié aux données de recherche avec des êtres humains](#)<sup>25</sup>
  - Partie 3 : [Langage en matière de gestion de données de recherche pour le consentement éclairé](#)<sup>26</sup>
- Curation Primers du Data Curation Network
  - [Curation of Data Collected by Informed Consent](#)<sup>27</sup> [en anglais]

<sup>23</sup> [https://ethics.gc.ca/fra/depositing\\_depots.html](https://ethics.gc.ca/fra/depositing_depots.html)

<sup>24</sup> <https://doi.org/10.5281/zenodo.4088985>

<sup>25</sup> <https://doi.org/10.5281/zenodo.4107118>

<sup>26</sup> <https://doi.org/10.5281/zenodo.4107185>

<sup>27</sup> <https://hdl.handle.net/11299/218838> [en anglais]

- [Human Participants Data Essentials](#)<sup>28</sup> [en anglais]
- [Directives de dépersonnalisation des données](#) du réseau Portage<sup>29</sup>
- Global Biodiversity Information Facility (GBIF) : [Current Best Practices for Generalizing Sensitive Species Occurrence Data](#)<sup>30</sup> [en anglais]

Les approches suivantes peuvent servir à déterminer si un ensemble de données contient ou non des éléments sensibles et des restrictions au partage des données :

- Selon la description, déterminez si des informations sensibles dans l'ensemble de données peuvent être divulguées.
- Vérifiez si les formulaires de consentement contiennent des formulations qui interdisent le partage des données.
- Renseignez-vous auprès du déposant sur la présence d'informations sensibles. **Note** : si le déposant doit remplir un formulaire d'information sur l'ensemble de données avant de soumettre sa demande, cette question peut être incluse.
- Ouvrez des fichiers et recherchez des contenus sensibles (par exemple, des informations précises sur la localisation des espèces et des sites protégés, des contenus offensants, des renseignements personnels directs et indirects, etc.).

Oui	Non	Incertain	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'ensemble de données comporte de l'information sur : <ul style="list-style-type: none"> <li>▪ Des êtres humains et des données sur la santé</li> <li>▪ Des thèmes/sujets autochtones</li> <li>▪ Des mineurs ou des personnes incapables de donner un consentement éclairé</li> <li>▪ Des espèces vulnérables</li> <li>▪ Des données soumises au droit d'auteur ou à toute autre restriction légale sans licence appropriée</li> <li>▪ Du contenu illégal/offensant</li> <li>▪ D'autre contenu qui pourrait être potentiellement sensible ou qui viole manifestement les conditions d'utilisation du dépôt</li> </ul>

<sup>28</sup> <https://hdl.handle.net/11299/216579> [en anglais]

<sup>29</sup> <https://portagenetwork.ca/fr/outils-et-ressources/directives-de-la-gdr-sur-la-covid-19/directives-sur-la-depersonnalisation-des-donnees/>

<sup>30</sup> <https://doi.org/10.15468/doc-5jp4-5g10> [en anglais]

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si la réponse à l'une des questions précédentes est « Oui », est-ce qu'au moins l'un des éléments suivants est inclus ? <ul style="list-style-type: none"> <li>▪ Formulaire de consentement</li> <li>▪ Entente de participation</li> <li>▪ Entente de partage des données</li> <li>▪ Tout autre document autorisant la publication de données pour les rendre disponibles pour un usage secondaire</li> </ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le formulaire de consentement ou l'accord de partage des données comprend-il une formulation qui permet au déposant de partager les données dans un dépôt ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Si les données ne peuvent être partagées que dans des circonstances particulières, le déposant a-t-il choisi les mesures appropriées pour en limiter l'accès ? <b>Note</b> : Les serveurs Dataverse ne sont pas cryptés et ne doivent pas être utilisés pour stocker des données sur des êtres humains identifiables.

### La soumission comporte des données ou du code de sources externes

Les ensembles de données doivent être inspectés pour détecter les données ou les codes provenant de sources externes afin de s'assurer que les déposants ont les droits ou les autorisations nécessaires pour partager les données et que l'attribution à ces sources a été effectuée correctement. Consultez le service ou les experts en droits d'auteur de votre établissement pour déterminer comment les politiques en matière de propriété intellectuelle des partenaires affectent les soumissions dans vos dépôts.

Oui	Non	Incertain	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'ensemble de données contient des informations propriétaires ou à accès restreint. Exemple : données ou codes sous licence commerciale ou propriétaire.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les données ou codes agrégées d'autres sources ont été cités correctement.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'accord de partage des données est indiqué ou inclus et permet la diffusion des informations.

## Recommander

À l'étape **Recommander**, vous pouvez demander des informations supplémentaires au déposant ou suggérer des modifications aux métadonnées et aux fichiers qui permettront de trouver et d'utiliser plus facilement l'ensemble de données conformément aux principes FAIR. Recommander est considéré comme une tâche de curation de niveau 1. À ce niveau, il pourrait ne pas être nécessaire de demander des modifications pour chaque dépôt, mais vous devrez contacter les chercheurs si les étapes minimales requises pour publier un ensemble de données dans Dataverse ne sont pas respectées. Si vous effectuez des tâches aux niveaux 2 et 3, vous devrez probablement contacter les chercheurs plus fréquemment. Soyez prêt à gérer les attentes, car certains déposants seront plus réceptifs aux recommandations ou aux changements que d'autres. Toutefois, même les chercheurs qui ne peuvent pas appliquer vos recommandations à ce dépôt peuvent les prendre en compte et les utiliser pour améliorer le prochain dépôt.

### Niveau 1

- Il peut être utile de signaler les questions relatives aux métadonnées ou aux données lorsque vous travaillez sur les autres étapes du cadre de CURATION.
- Vos demandes doivent être classées par ordre de priorité, et la différence entre ce qui est nécessaire pour publier l'ensemble de données et ce qui serait agréable à avoir doit être claire.
- Si l'ensemble de données est si incomplet qu'il nécessite une révision importante, vous devriez en discuter en personne plutôt que par courriel.
- Si vous avez des lignes directrices pour aider les chercheurs pendant la procédure de dépôt, servez-vous de celles-ci pour amorcer la conversation.
- Pour gagner du temps et être conséquent, utilisez des modèles comme point de départ pour vos messages. Pour vous aider à démarrer, voir [Curating Research Data Volume Two : A Handbook of Current Practice](#),<sup>31</sup> section 3.5 pour un exemple de message et des informations supplémentaires.<sup>32</sup>

---

<sup>31</sup> <https://hdl.handle.net/11299/185335> [en anglais]

<sup>32</sup> Une variante de ce modèle a été traduite en français et elle est disponible à l'Annexe 2

## Améliorer

L'étape **Améliorer** a pour but d'améliorer la soumission afin de faciliter davantage la découverte et l'utilisation. Cette étape comprend toutes les mesures que vous pouvez prendre (en tant que curateur) pour améliorer la documentation et les métadonnées au-delà de ce qui a été fait dans les étapes Consulter et Un peu plus en profondeur. Pour vous guider dans l'enrichissement des métadonnées, utilisez le [Guide des pratiques exemplaires sur les métadonnées de Dataverse Nord](#).<sup>33</sup>

### Niveau 2

Les métadonnées sont rigoureuses, exactes et complètes

**Conseil** : Les termes de classification des mots-clés et des sujets fournissent des informations supplémentaires sur l'ensemble de données en le situant dans un domaine d'étude et en facilitant son indexation, sa classification et sa découverte.

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Utilisation de titres descriptifs.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Définition des termes et des acronymes.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Absence de jargon (pour la lisibilité humaine) et de symboles (pour la lisibilité informatique).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Utilisation de langage particulier et précis.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Ajout de termes dans le champ <b>mot-clé</b> de l'ensemble de données.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Ajout de termes dans le champ <b>classification par sujets</b> de l'ensemble de données.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Chaque terme a été attribué à son propre domaine (c'est-à-dire un terme par case).
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les termes sont fournis à partir d'un vocabulaire contrôlé (c'est-à-dire que les sous-champs <b>vocabulaire</b> et <b>adresse URL du vocabulaire</b> sont complétés).

Des liens vers des publications, des ensembles de données et d'autres ressources connexes sont inclus

<sup>33</sup> <http://hdl.handle.net/2429/73610>

Des choses publiées ailleurs peuvent fournir plus de contexte au nouveau dépôt et aider l'utilisateur final à mieux comprendre l'ensemble de données. Certaines de ces ressources peuvent être mentionnées dans la description de l'ensemble de données ou dans le fichier README, ou elles peuvent avoir été découvertes au cours du processus de révision de l'ensemble de données. Si possible, fournissez des liens directs vers les publications, le code, les modèles, la documentation, les données sources et autres ressources connexes.

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le champ <b>publication connexe</b> et les sous-champs connexes (citation, type d'identifiant, numéro d'identifiant, URL) sont complets.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Tous les ensembles de données connexes sont notés dans le champ <b>ensembles de données connexes</b> .
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les liens vers d'autres ressources publiées, telles que les modèles ou codes associés, la documentation, l'instrument d'enquête, le protocole d'étude, les plans d'analyse, le plan de gestion des données sont ajoutés au champ <b>autres références</b> .
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le champ et les sous-champs <b>renseignements sur la subvention</b> (organisme subventionnaire et numéro de subvention) sont complets.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le champ <b>sources de données</b> est complet.

## Transposer

L'objectif de l'étape **Transposer** est de produire des ensembles de données qui utilisent des formats ouverts et communs si possible. Les formats ouverts aideront à faire en sorte que les données soient accessibles au public le plus large possible, y compris à ceux qui n'ont pas accès à des logiciels spécialisés, et que des mesures de préservation puissent être prises à l'avenir. Il y aura des moments où des formats spécialisés seront nécessaires (par exemple, des types/structures de données complexes qui nécessitent une représentation particulière), mais lorsque d'autres options existent, vous pouvez envisager la transformation. La perte de données ou d'informations intégrées dans la structure des données ou des métadonnées peut être inquiétante. Dans la mesure du possible, demandez au chercheur de fournir un autre format plutôt que de transformer les fichiers vous-même afin qu'il puisse confirmer que le résultat est exact et complet.

La conservation des fichiers originaux aux côtés des fichiers transformés peut être la solution idéale, car les données originales restent disponibles sous la forme utilisée par le chercheur et peuvent être ouvertes et manipulées dans leur logiciel d'origine. Toutefois, en fonction de la taille globale de la soumission, le déposant peut ne pas vouloir publier à la fois les fichiers originaux en format protégé et les fichiers en format ouvert.

Si les fichiers sont transformés, assurez-vous que les noms de fichiers indiquent les fichiers originaux et les fichiers transformés. Si le curateur ou le spécialiste de la préservation est responsable de la transformation, les fichiers originaux doivent être conservés conformément à la politique du dépôt. Si les fichiers ne peuvent pas être transformés, assurez-vous que le fichier README décrit le contenu du fichier ainsi que le logiciel nécessaire pour consulter ou utiliser les fichiers.

## Niveau 3

Les formats de fichiers sont ouverts ou documentés correctement

Oui	Non	Certains problèmes	S.O.	Identification des formats de fichiers et des logiciels utilisés pour créer les fichiers...
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	L'ensemble de données contient des formats de fichiers propriétaires ou des formats qui ne sont pas adaptés à la préservation?  Consultez les <a href="#">formats de fichiers privilégiés Dataverse NO</a> <sup>34</sup> [en anglais] ou <a href="#">Guide concernant les formats recommandés par Bibliothèque et Archives nationales du Québec</a> . <sup>35</sup>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les fichiers propriétaires peuvent-ils être transformés en formats non propriétaires sans perte de données ?

<sup>34</sup> <https://site.uit.no/dataverseno/deposit/prepare/> [en anglais]

<sup>35</sup> [https://www.banq.qc.ca/documents/archives/archivistique\\_gestion/publications\\_proposees/Guide-formats-BAnQ\\_Final.pdf](https://www.banq.qc.ca/documents/archives/archivistique_gestion/publications_proposees/Guide-formats-BAnQ_Final.pdf)

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Si les fichiers <b>ne peuvent pas</b> être transformés, la documentation décrit-elle :</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Le format et le contenu des fichiers ?</li> <li><input type="checkbox"/> Le logiciel ou l'instrument qui a généré les fichiers ?</li> <li><input type="checkbox"/> Le logiciel nécessaire pour visualiser ou utiliser les fichiers ?</li> <li><input type="checkbox"/> Un logiciel gratuit pour visualiser les fichiers, le cas échéant ?</li> </ul>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>La documentation décrit les données, les formules utilisées, les titres de colonnes, les étiquettes de variables, etc. afin que les données soient utilisables.</p>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>La documentation est présentée dans un format recommandé (PDF/A ou .txt)</p>

## Inclure

L'objectif de l'étape **Inclure** est de faciliter la réutilisation des données ainsi que l'attribution et le crédit appropriés en indiquant les identifiants persistants pertinents et les informations de licence appropriées. Une partie de cette opération se fera probablement automatiquement ; par exemple, le système attribuera un identifiant DOI et l'enregistrera lorsque l'ensemble de données sera publié, mais d'autres informations (par exemple, les identifiants ORCID) doivent être ajoutées manuellement.

Une licence décrira les conditions de réutilisation de l'ensemble de données publié. La licence par défaut dans Dataverse est [CC0 pour œuvre dédiée au domaine public](#),<sup>36</sup> mais vous devrez peut-être y apporter quelques changements. Idéalement, la licence choisie pour l'ensemble de données répondra aux besoins du déposant sans être trop restrictive. On doit tenir compte de la préférence du déposant, des exigences de redistribution de toutes les données ou tous les codes obtenus ou dérivés d'une source externe ainsi que des exigences de n'importe quel accord de partage de données que le déposant pourrait avoir conclu. Des outils, tels que [Creative Commons License Chooser](#)<sup>37</sup> et GitHub's [Choose an Open Source License](#)<sup>38</sup> [en anglais], et des ressources telles que [Open Source Initiative](#)<sup>39</sup> [en anglais], [la Free Software Foundation](#)<sup>40</sup> et [l'Open Knowledge Foundation](#)<sup>41</sup> [en anglais] peuvent aider à sélectionner une licence appropriée. La communauté Dataverse a également publié un modèle qui peut être utile pour rédiger une entente de réutilisation des données.<sup>42</sup> Veuillez noter que Creative Commons recommande de ne pas utiliser ses licences pour du code ou des logiciels.<sup>43</sup>

## Niveau 2

### Indiquer les identifiants persistants lorsque possible

Oui	Non	Certains problèmes	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Cet ensemble de données fait-il partie d'une publication (ou d'un fichier supplémentaire d'une publication) ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le cas échéant, y a-t-il un titre ou un lien permanent disponible dans la documentation ou la fiche de métadonnées ? Vérifiez si le lien vous amène à la bonne publication.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Est-ce qu'il y a un identifiant d'auteur pertinent, notamment ORCID ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Est-ce qu'il y a un identifiant d'organisme subventionnaire pertinent ?

<sup>36</sup> <https://creativecommons.org/publicdomain/zero/1.0/deed.fr>

<sup>37</sup> <https://creativecommons.org/choose/?lang=fr>

<sup>38</sup> <https://choosealicense.com/> [en anglais]

<sup>39</sup> <https://opensource.org/licenses> [en anglais]

<sup>40</sup> <https://www.gnu.org/licenses/license-list.fr.html>

<sup>41</sup> <https://opendatacommons.org/> [en anglais]

<sup>42</sup> <https://dataverse.org/best-practices/sample-dua> [en anglais]

<sup>43</sup> <https://creativecommons.org/faq/fr/#puis-je-appliquer-une-licence-creative-commons-à-un-logiciel>

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le DOI des données attribué par le système est enregistré et mène à la page de renvoi de l'ensemble de données (si la curation a lieu avant la publication, vous ne pourrez le confirmer qu'après la publication de la curation).
--------------------------	--------------------------	--------------------------	--------------------------	---

### Niveau 3

#### Réviser la licence et les conditions d'utilisation de l'ensemble de données

Il incombe au chercheur de choisir pour l'ensemble de données qu'il publie une licence appropriée et conforme aux conditions d'utilisation qui ont été appliquées à toutes les ressources utilisées pour générer son ensemble de données. Vous pouvez discuter des points suivants avec le chercheur.

Oui	Non	Incertain	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le déposant a-t-il déjà choisi une licence ? <input type="checkbox"/> Le cas échéant, quelle licence ou quelles conditions d'utilisation ont été choisies ? _____
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Le dépôt contient des données ou des codes obtenus ou dérivés d'une source externe ? <input type="checkbox"/> Le cas échéant, quelle licence ou quelles conditions d'utilisation ont été attribuées à la source originale ? _____
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	La licence choisie par le déposant est-elle moins contraignante que les conditions fixées par une source externe ? <input type="checkbox"/> Le cas échéant, consulter le déposant pour lui proposer une licence plus appropriée. <input type="checkbox"/> Autrement, vous pouvez aller de l'avant avec la licence que le déposant a choisie.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Les données obtenues de sources tierces ou dérivées d'ensembles de données existants, sont-elles correctement attribuées ?
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Encouragez la citation de données pour les ensembles de données externes.

## Optimiser

Lors de l'étape **Optimiser**, [le caractère FAIR](#)<sup>44</sup> d'un ensemble de données est formellement évalué et vous ou le chercheur pouvez prendre des mesures pour vous assurer que les données soient faciles à repérer, accessibles, interopérables et réutilisables. Bien que certains des principes FAIR soient déjà pris en compte dans le cadre de CURATION, cette étape est une évaluation formelle. Vous pouvez formaliser et évaluer la conformité globale aux principes FAIR en utilisant un outil particulier. Vous pouvez effectuer cette tâche à des fins internes, pour améliorer votre propre pratique, ou vous pouvez utiliser les résultats de l'évaluation pour encadrer les demandes que vous envoyez au chercheur à l'étape Recommander.

## Niveau 3

### Évaluer l'ensemble de données et optimiser la conformité aux principes FAIR

Utilisez l'un des outils offerts pour évaluer si l'ensemble de données est conforme aux principes FAIR :

- [SATIFYD](#)<sup>45</sup> (Self-Assessment Tool to Improve the FAIRness of Your Dataset) – outil d'auto-évaluation pour améliorer la conformité de l'ensemble de données aux principes FAIR offert par Data Archiving and Networked Services (DANS) qui évalue la conformité selon vos réponses à 12 questions et offre des conseils pour améliorer le résultat.
- [CSIRO 5-star Data Rating Tool](#)<sup>46</sup> est un système d'auto-évaluation de la conformité aux principes FAIR. Les critères d'évaluation sont adaptés à partir des critères d'évaluation des données d'OzNome,<sup>47</sup> qui fournissent des mesures spécifiques pour obtenir des données fiables et conformes aux principes FAIR.
- [FAIR Self Assessment Tool](#)<sup>48</sup> du Australian Research Data Commons (ARDC) a été conçu pour les bibliothécaires de données et le personnel informatique. Les réponses sont sélectionnées dans un menu déroulant et la conformité globale aux principes FAIR est indiquée en temps réel par une « barre de progression » verte.

---

<sup>44</sup> <https://doi.org/10.1038/sdata.2016.18> [en anglais]

<sup>45</sup> <https://satisfyd.dans.knaw.nl/> [en anglais]

<sup>46</sup> <https://doi.org/10.4225/08/5a12348f8567b> [en anglais]

<sup>47</sup> <https://confluence.csiro.au/display/OZNOME/Data+ratings> [en anglais]

<sup>48</sup> <https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/> [en anglais]

## Noter

L'objectif de l'étape **Noter** est de s'assurer que vous avez fait un compte rendu écrit et précis de votre travail de curation. Certaines actions et informations peuvent être consignées automatiquement dans l'enregistrement de métadonnées (par exemple, le nom du curateur, la date à laquelle l'ensemble de données a été soumis, renvoyé ou approuvé) et d'autres peuvent être documentées dans un journal de curation, soit un document normalisé qui résume les modifications apportées, les personnes qui les ont apportées et les raisons de ces modifications. Le format, le contenu et l'étendue du journal de curation peuvent varier en fonction des exigences de votre dépôt et du degré d'interaction que vous avez avec vos déposants. Bien qu'il s'agisse de la dernière étape du processus de curation, il est utile de créer et de remplir le journal au fur et à mesure des étapes de la curation, puis de le finaliser à la fin du processus. Vous trouverez de plus amples informations sur les journaux de curation dans [Curating Research Data Volume Two: A Handbook of Current Practice](#), dont un gabarit (section 3.2, page 86)<sup>49</sup> [en anglais].

## Niveau 2

### Créer un journal de curation pour documenter les décisions et les actions posées

Oui	Non	S.O.	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Créez un fichier de journal en texte brut pour documenter les informations suivantes :</p> <ul style="list-style-type: none"><li><input type="checkbox"/> Un identifiant DOI ou un identifiant interne de l'objet, ou un autre moyen de lier le fichier de journal à l'ensemble de données.</li><li><input type="checkbox"/> Tout problème potentiel découvert dans le cadre du processus de curation.</li><li><input type="checkbox"/> Les changements de haut niveau généraux recommandés au déposant et les questions pour le déposant.</li><li><input type="checkbox"/> Un résumé général des réponses des déposants aux demandes et aux questions.</li><li><input type="checkbox"/> Modifications conséquentes de l'ensemble de données :<ul style="list-style-type: none"><li><input type="checkbox"/> Les modifications apportées aux fichiers ou à la documentation.</li><li><input type="checkbox"/> Le nom des fichiers ajoutés ou retirés de l'ensemble de données et les motifs du retrait.</li><li><input type="checkbox"/> La liste des métadonnées modifiées, ajoutées ou supprimées et une explication de ces changements.</li></ul></li></ul>

<sup>49</sup> <https://hdl.handle.net/11299/185335> [en anglais]

## Références

Australian Research Data Commons. s.d. FAIR Self Assessment Tool. Consulté le 20 mars 2021.

<https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>. [En anglais].

Bascik, Teresa, Philippe Boisvert, Alexandra Cooper, Martine Gagnon, Mark Goodwin, John Huck, Amber Leahey, Michael Steeleworthy, et Sally Taylor. 2020. Guide des pratiques exemplaires sur les métadonnées de Dataverse Nord : Version 2.0. Vancouver : Bibliothèque de l'Université de la Colombie-Britannique.

<http://hdl.handle.net/2429/73610>.

Bibliothèque et Archives nationales du Québec. 2020. Guide concernant les formats recommandés par BANQ. Consulté le 2 juin 2021.

[https://www.banq.gc.ca/documents/archives/archivistique\\_gestion/publications\\_proposees/Guide-formats-BANQ\\_Final.pdf](https://www.banq.gc.ca/documents/archives/archivistique_gestion/publications_proposees/Guide-formats-BANQ_Final.pdf).

Bibliothèque de l'Université d'Ottawa. Équipe de gestion des données de recherche. s.d. Nommage de fichiers et gestion de versions. Consulté le 9 juin 2021. <https://biblio.uottawa.ca/fr/services/professeurs/gestion-donnees-recherche/designation-fichiers-disposition-donnees>.

Brigham, Doug. 2020. Créer un fichier README pour vos ensembles de données (Guide rapide). Zenodo.

<https://doi.org/10.5281/zenodo.4058961>.

Chapman, Arthur D. 2020. Current Best Practices for Generalizing Sensitive Species Occurrence Data.

Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-5jp4-5g10>. [En anglais].

Cornell University Research Data Management Service Group. s.d. Guide to Writing « Readme »

Style Metadata. Consulté le 20 mars 2021. <https://data.research.cornell.edu/content/readme>. [En anglais].

Creative Commons. s.d.-a. CC0 1.0 universel (CC0 1.0) transfert dans le domaine public. Consulté le 9 juin 2021.

<https://creativecommons.org/publicdomain/zero/1.0/deed.fr>.

Creative Commons. s.d.-b. Choose a license. Consulté le 4 juin 2021.

<https://creativecommons.org/choose/?lang=fr>.

Creative Commons. s.d.-c. Foire Aux Questions : Puis-je appliquer une licence Creative Commons à un logiciel ?

Consulté le 8 septembre 2021. <https://creativecommons.org/faq/fr/#puis-je-appliquer-une-licence-creative-commons-%C3%A0-un-logiciel>.

Darragh, Jen, Alicia Hofelich Mohr, Shanda Hunt, Rachel Woodbrook, Dave Fearon, Jennifer Moore, et Hannah Hadley. 2020. Human Subjects Data Essentials Data Curation Primer. Data Curation Network. Récupéré de l'Université de Minnesota, University Digital Conservancy. <https://hdl.handle.net/11299/216579>. [En anglais].

Data Curation Network. 2019. Data Curation Primers. [Collection]. l'Université de Minnesota, University Digital Conservancy. <https://hdl.handle.net/11299/202810>. [En anglais].

Data Curation Network. s.d. The DCN Curation Workflow. Consulté le 20 mars 2021. <https://datacurationnetwork.org/outputs/workflows/>. [En anglais].

Dataverse Project. s.d. Sample Data Usage Agreement. Consulté le 28 mars 2021. <https://dataverse.org/best-practices/sample-dua>. [En anglais].

DataverseNO. s.d. Prepare Your Data: Preferred File Formats. Consulté le 28 mars 2021. <https://site.uit.no/dataverseno/deposit/prepare/>. [En anglais].

Fankhauser, Eliane, Jerry de Vries, Nina Westzaan, et Vesa Åkerman. 2019. SATIFYD: Self-Assessment Tool to Improve the FAIRness of Your Dataset. Data Archiving and Networked Services. Consulté le 20 mars 2021. <https://satisfyd.dans.knaw.nl/>. [En anglais].

Free Software Foundation. s.d. Licences commentées. Modifié pour la dernière fois 4 juillet 2021. <https://www.gnu.org/licenses/license-list.fr.html>.

GitHub. s.d. Choose an Open Source License. Modifié pour la dernière fois 23 mars 2021. <https://choosealicense.com/>. [En anglais].

GitHub. s.d. Making your code citable: GitHub Guides. Consulté le 7 septembre 2021. <https://guides.github.com/activities/citable-code/>. [En anglais].

Gouvernement du Canada. 2018. Foire aux questions : Politique des trois organismes sur la gestion des données de recherche. Innovation, Sciences et Développement économique Canada. Modifié pour la dernière fois 15 mars 2021. [https://www.ic.gc.ca/eic/site/063.nsf/fra/h\\_97609.html#1d](https://www.ic.gc.ca/eic/site/063.nsf/fra/h_97609.html#1d).

Gouvernement du Canada. 2021. Lignes Directrices pour verser des données existantes dans des dépôts publics. Groupe en éthique de la recherche. Modifié pour la dernière fois 25 mai 2021. [https://ethics.gc.ca/fra/depositing\\_depots.html](https://ethics.gc.ca/fra/depositing_depots.html).

Le groupe d'experts sur la curation du Portage. 2019. Introduction à Curation de données. Consulté le 2 juin 2021. [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).

Le groupe d'experts sur les données sensibles. 2020-a. Boîte à outils pour les données sensibles — destiné aux chercheurs, Partie 1 : Glossaire terminologique sur l'utilisation des données sensibles à des fins de recherche. Zenodo. <https://doi.org/10.5281/zenodo.4088985>.

Le groupe d'experts sur les données sensibles. 2020-b. Boîte à outils pour les données sensibles — destiné aux chercheurs, Partie 2 : Matrice de risque lié aux données de recherche avec des êtres humains. Zenodo. <https://doi.org/10.5281/zenodo.4107118>.

Le groupe d'experts sur les données sensibles. 2020-c. Boîte à outils pour les données sensibles — destiné aux chercheurs, Partie 3 : Langage en matière de gestion de données de recherche pour le consentement éclairé. Zenodo. <https://doi.org/10.5281/zenodo.4107185>.

Le groupe de travail sur la COVID-19 de Portage. Directives sur la dépersonnalisation des données. Consulté le 29 mars 2021. <https://portagenetwork.ca/fr/outils-et-ressources/directives-de-la-gdr-sur-la-covid-19/directives-sur-la-depersonnalisation-des-donnees/>.

Hunt, Shanda, Alicia Hofelich Mohr, et Rachel Woodbrook. 2021. Consent Forms Data Curation Primer. Data Curation Network. Récupéré de l'Université de Minnesota, University Digital Conservancy. <https://hdl.handle.net/11299/218838>. [En anglais].

l'Institute for Quantitative Social Sciences. s.d. Dataverse Project: User Guide. Version 5.3. Modifié pour la dernière fois 3 décembre 2020. <https://dataverse.scholarsportal.info/guides/en/latest/user/>. [En anglais].

Johnston, Lisa R. 2017. *Curating Research Data Volume Two: A Handbook of Current Practice*. Chicago: Association of College and Research Libraries. Récupéré de l'Université de Minnesota, University Digital Conservancy. <https://hdl.handle.net/11299/185335>. [En anglais].

Johnston, Lisa R., Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, Claire Stewart, Mara Blake, Joel Herndon, Timothy M. McGeary, Elizabeth Hull, et Elizabeth Coburn. 2018. Data Curation Network: A Cross-institutional Staffing Model for Curating Research Data. *International Journal of Digital Curation* 13 (1): 125-140. <https://doi.org/10.2218/ijdc.v13i1.616>. [En anglais].

Khair, Shahira. 2020. Deposit Guidelines for UVic Dataverse. Librairie de l'Université de Victoria, Services à les données de recherche. [https://libguides.uvic.ca/ld.php?content\\_id=35154390](https://libguides.uvic.ca/ld.php?content_id=35154390). [En anglais].

Lafferty-Hess, Sophia, Julie Rudder, Moira Downey, Susan Ivey, Jennifer Darragh, et Rebekah Kati. 2020. Conceptualizing Data Curation Activities Within Two Academic Libraries. *Journal of Librarianship and Scholarly Communication* 8 (1): eP2347. <https://doi.org/10.7710/2162-3309.2347>. [En anglais].

l'Open Knowledge Foundation. s.d. Open Data Commons: Legal Tools for Open Data. Consulté le 28 mars 2021. <https://opendatacommons.org/>. [En anglais].

Open Source Initiative. s.d. Licenses et Standards. Consulté le 28 mars 2021. <https://opensource.org/licenses>. [En anglais].

Scholars Portal. s.d. Guide de Dataverse de Scholars Portal. Consulté le 28 mars 2021. <https://learn.scholarsportal.info/fr/guides/dataverse/>.

Software Heritage. s.d. Save and reference research software. Consulté le 8 février 2021. <https://www.softwareheritage.org/save-and-reference-research-software/?lang=fr>. [En anglais].

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, ... Barend Mons. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>. [En anglais].

Yu, Jonathan, et Simon Cox. 2017. 5-Star Data Rating Tool, v4. The Commonwealth Scientific and Industrial Research Organisation. Software Collection. Consulté le 20 mars 2021. <https://doi.org/10.4225/08/5a12348f8567b>. [En anglais].

Yu, Jonathan. s.d. OzNome: Data Ratings. Dernière mise à jour 22 mai 2018. <https://confluence.csiro.au/display/OZNOME/Data+ratings>. [En anglais].

## Annexe 1 – Exemples d’ensembles de données avec curation effectuée dans Dataverse

Cette annexe présente des exemples d'ensembles de données dont la curation et la publication ont été effectuées dans le portail Dataverse de Scholars Portal. La première série d'exemples montre le niveau minimal de curation nécessaire pour déposer des données dans Dataverse. La série d'exemples suivante aborde différents éléments de la curation et fournit des exemples de bonnes pratiques de curation.

### Niveau 1

Les étapes de curation de niveau 1 préparent les ensembles de données pour la publication dans Dataverse. Ce niveau de curation est le minimum requis pour publier avec succès dans Dataverse et rendre l'ensemble de données retrouvable. Les étapes de niveau 1 comprennent la confirmation que les champs de métadonnées obligatoires sont complets et exacts. Il est possible d'avoir recours à des fichiers Readme et à la curation de fichiers.

- Ipsos Canadian Public Affairs Dataverse - <https://dataverse.scholarsportal.info/dataverse/ipsos>
- Ipsos Global @dvisor Dataverse (University of Toronto) - [https://dataverse.scholarsportal.info/dataverse/ipsos\\_UofT](https://dataverse.scholarsportal.info/dataverse/ipsos_UofT)
- Winseck, Dwayne, 2020, « Canadian Media Concentration Research Project Dataset 2019 », <https://doi.org/10.5683/SP2/YSZMOQ>, Scholars Portal Dataverse, V1
- Hamarat, Yaprak, 2019, Données photographiques pour la thèse de doctorat intitulée « L'esthétique de l'engagement écologique : l'impensé des politiques environnementales », <https://doi.org/10.5683/SP2/IS0SWH>, Scholars Portal Dataverse, V1

### Exemples généraux

Voici des exemples qui illustrent une bonne curation des ensembles de données.

- Margaret B. Harrison; Practice and Research in Nursing Group: Wound Care Collaborative; Ian Graham; E. Andrea Nelson; Elizabeth VanDenKerkhof; Karen Lorimer; Connie Harris; Meg Carley; The Canadian Bandaging Trial Group, 2013, « Practice and Research in Nursing (PRN) Wound Studies, 1999-2009 [Canada] », <https://hdl.handle.net/10864/CORX8>, Scholars Portal Dataverse, V6
  - Cet ensemble de données donne un aperçu des études utilisées pour créer l'ensemble de données et une description détaillée de la façon dont les données ont été traitées. L'équipe de curation a travaillé avec le gestionnaire de données pour garantir la confidentialité des participants tout en maximisant la partie de l'ensemble de données ouverte à la réutilisation.
- Agricultural and Forest Meteorology Group. Elora Research Station/Guelph Turfgrass Institute, 2018, « Weather records for the Elora Research Station, Elora, Ontario [Canada]: Meteorological data 2018 », <https://doi.org/10.5683/SP/RQRDSH>, Scholars Portal Dataverse, V3.
  - Cet ensemble de données comprend une description claire résumant les données qui ont été collectées et les anomalies qui ont affecté l'enregistrement. Il contient un ensemble de données tabulaires et un guide de codification bien structurés.

- Paul, Jason; Baltzer, Jennifer L.; Kokelj, Steve V., 2020, « Near-surface permafrost ground ice characteristics and ecological and physical drivers of transient layer ice content in discontinuous permafrost near Yellowknife, NT », <https://doi.org/10.5683/SP2/LX5IJN>, Scholars Portal Dataverse, V1
  - La description de l'ensemble de données résume la collecte, l'emplacement et la méthode des données. Les métadonnées sont enrichies d'informations géospatiales pour améliorer l'indexation et la recherche. Les fichiers individuels comprennent des métadonnées descriptives identifiant leur contenu avant le téléchargement. Toutes les variables sont définies dans les feuilles de calcul. Les fichiers décrivent toutes les variables dans toutes les feuilles de calcul. Le fichier Readme comprend des informations méthodologiques, des coordonnées personnelles et un aperçu du fichier à titre de référence. Le chercheur a choisi une licence CC BY-NC-SA 4.0 et les utilisateurs finaux doivent accepter les conditions d'utilisation avant de pouvoir télécharger les données.
- Moubayed, Anna-Maria, 2018, « All About Eve and Other Stories: Representations of Eve in French Romanesque Sculpture (and more) », <https://doi.org/10.5683/SP/2Y5C1X>, Scholars Portal Dataverse, V1
  - Un exemple d'ensemble de données en sciences humaines. En plus du fichier de données, tous les carnets de recherche sur le terrain sont inclus, ainsi qu'un lien vers une carte interactive et les fichiers nécessaires pour recréer la carte.
- Gagné, Monique; Ward, W. Peter, 2019, « Birth weight and economic growth data sets, Utrecht Hospital, 1880-1940, [2012] », <https://doi.org/10.5683/SP2/WNY6FG>, Scholars Portal Dataverse, V1.
  - Cette publication contient l'ensemble de données original, auquel ont été ajoutés un guide de codification, une description de la disposition des enregistrements et un dictionnaire de données pour faciliter la réutilisation.

## Annexe 2 – Modèles pour la correspondance

Les modèles suivants peuvent servir à la correspondance avec les déposants. Vous devrez peut-être modifier un modèle pour qu'il reflète plus précisément l'ensemble de données nécessitant une curation ; autrement, vous souhaitez peut-être modifier un modèle pour qu'il reflète mieux votre voix. Dans certains cas, l'utilisation du mot « Question » au lieu de « Recommandation » peut être plus appropriée. Veuillez noter que le texte entre parenthèses doit être supprimé ou remplacé par les détails de l'ensemble de données nécessitant une curation avant d'envoyer votre message.

### Demande de changement sans renvoi de soumission au déposant

**Objet** : Soumission récente Dataverse [Indiquez le titre résumé de l'ensemble de données]

Bonjour [NOM],

Merci de votre récente soumission [Indiquez le titre complet de l'ensemble de données] au [Nom du dépôt]. [Vous pouvez ajouter quelque chose de positif ! Par exemple, un complément de documentation, une étude intéressante, etc.]. Lorsque nous recevons un ensemble de données, nous l'examinons pour nous assurer que les ensembles de données que nous hébergeons sont aussi exhaustifs, accessibles et compréhensibles que possible. Ainsi, nous avons examiné votre soumission et nous souhaitons vous faire part des recommandations suivantes :

Recommandation #1

Recommandation #2

Recommandation #3

Recommandation #4

Si vous souhaitez apporter ces modifications, veuillez [XXX] [Par exemple, nous envoyer par courriel le fichier README révisé et nous l'utiliserons pour remplacer celui dans votre soumission ; OU, confirmer les modifications et nous mettrons à jour les métadonnées et publierons votre soumission].

N'hésitez pas à communiquer avec nous si vous avez des questions ou des préoccupations.

Cordialement,  
[Indiquez votre nom]

### Demande de changement avec renvoi de soumission au déposant pour examen

**Objet** : Soumission récente Dataverse [Indiquez le titre résumé de l'ensemble de données]

Bonjour [NOM],

Merci de votre récente soumission [Indiquez le titre complet de l'ensemble de données] au [Nom du dépôt]. [Vous pouvez ajouter quelque chose de positif ! Par exemple, un complément de documentation, une étude intéressante, etc.]. Lorsque nous recevons un ensemble de données, nous l'examinons pour nous assurer que les ensembles de données que nous hébergeons sont aussi exhaustifs, accessibles et compréhensibles que

possible. Ainsi, nous avons examiné votre soumission et nous souhaitons vous faire part des recommandations suivantes :

Recommandation #1

Recommandation #2

Recommandation #3

Recommandation #4

Nous vous avons renvoyé votre demande afin que vous puissiez modifier vos fichiers. Pour y accéder, veuillez vous connecter à [XXXX]. Lorsque vos modifications sont terminées, veuillez vérifier que tout semble exact et soumettre à nouveau votre ensemble de données.

Si vous avez des questions concernant l'une de nos recommandations ou si vous éprouvez des difficultés à accéder à votre soumission ou à la modifier, veuillez nous en informer.

Cordialement,

[Indiquez votre nom]

## Annexe 3 – Ressources de curation supplémentaires

Le groupe d'experts sur la curation (GEC) de Portage NOIRN gère le [Guide de survie pour la curation de données](#), un espace où les professionnels de la gestion et de la curation des données de recherche peuvent partager des ressources liées à la pratique de la curation des données. Il a pour but d'aider à renforcer la capacité de curation des données et de soutenir le milieu de la curation des données au Canada et ailleurs.

Ressources supplémentaires - concernant la curation en général et la curation dans les Dataverse en particulier - ont été indexées dans le Guide de survie pour la curation des données.

## Annexe 4 – CURATION dans Dataverse : Guide éclair

- **Niveau 1:** Ce niveau de curation prépare les ensembles de données pour la publication dans Dataverse et permet de les rendre minimalement repérables.
- **Niveau 2:** Ce niveau de curation améliore la possibilité de découvrir les ensembles de données et contribue à permettre leur réutilisation.
- **Niveau 3:** Ce niveau de curation prépare les ensembles de données pour la reproductibilité et pour la préservation.

Une lettre	La définition	Les principales tâches	Niveau 1	Niveau 2	Niveau 3
<b>C</b>	<p><b>Consulter</b> S’assurer que toutes les composantes des données et des métadonnées nécessaires à la bonne marche de la publication de l’ensemble de données sont présentes.</p>	<input type="checkbox"/> L’ensemble de données a été soumis dans l’espace dataverse approprié	X		
		<input type="checkbox"/> Le chercheur a confirmé que l’ensemble de données ne pose pas de problème de licence ou de propriété intellectuelle	X		
		<input type="checkbox"/> Le chercheur a confirmé que l’ensemble de données est exempt d’informations identificatoires/sensibles	X		
		<input type="checkbox"/> La documentation de soutien est incluse	X		
		<input type="checkbox"/> Tous les fichiers mentionnés dans la documentation sont inclus dans l’ensemble de données	X		
		<input type="checkbox"/> Les champs de métadonnées requis sont exacts	X		
<b>U</b>	<p><b>Un peu plus en profondeur</b> Faire en sorte que l’ensemble de données soit bien décrit et que les futurs utilisateurs aient une idée précise de la</p>	<input type="checkbox"/> La documentation est rigoureuse, précise et complète		X	
		<input type="checkbox"/> Les fichiers s’ouvrent et les contenus s’affichent correctement		X	
		<input type="checkbox"/> Les fichiers et dossiers sont nommés et structurés correctement		X	

	nature des données et de la manière dont elles peuvent être utiles.	<input type="checkbox"/> Le code est commenté et produit les résultats souhaités			X
		<input type="checkbox"/> La soumission contient des éléments potentiellement sensibles			X
		<input type="checkbox"/> La soumission comporte des données ou du code de sources externes			X
<b>R</b>	<u>Recommander</u> Demander des informations supplémentaires à la personne qui dépose et suggérer des modifications des métadonnées ou des fichiers qui faciliteront le repérage et la réutilisation des données dans la perspective des principes FAIR.	<input type="checkbox"/> Comprendre le nombre de recommandations à transmettre à la personne qui dépose et la portée de ces demandes -- concentrez-vous sur trois à cinq changements	X		
		<input type="checkbox"/> Demander des informations supplémentaires à la personne qui dépose	X		
<b>A</b>	<u>Améliorer</u> Améliorer la soumission pour faciliter le repérage et la réutilisation de l'ensemble de données.	<input type="checkbox"/> Les métadonnées sont rigoureuses, exactes et complètes		X	
		<input type="checkbox"/> Des liens vers des publications, des ensembles de données et d'autres ressources connexes sont inclus		X	
<b>T</b>	<u>Transposer</u> S'assurer que l'ensemble de données emploie des formats de fichier aussi ouverts et aussi courants que possible.	<input type="checkbox"/> Les formats de fichiers sont ouverts ou correctement documentés			X
<b>I</b>	<u>Inclure</u> Faciliter la réutilisation, la bonne citation et la reconnaissance de la	<input type="checkbox"/> Indiquer les identifiants persistants lorsque possible		X	

	<p>source des données en incluant les identifiants pérennes pertinents et les informations nécessaires sur les licences.</p>	<input type="checkbox"/> Réviser la licence et les conditions d'utilisation de l'ensemble de données			X
O	<p><b>Optimiser</b> Évaluer la conformité globale de l'ensemble de données aux principes FAIR et prendre les mesures nécessaires pour optimiser la facilité de repérage, l'accessibilité, l'interopérabilité et la réutilisation de l'ensemble de données.</p>	<input type="checkbox"/> Évaluer l'ensemble de données et optimiser la conformité aux principes FAIR			X
N	<p><b>Noter</b> Consigner par écrit les détails de la démarche du travail de curation.</p>	<input type="checkbox"/> Créer un journal de curation pour documenter les décisions et les actions posées		X	