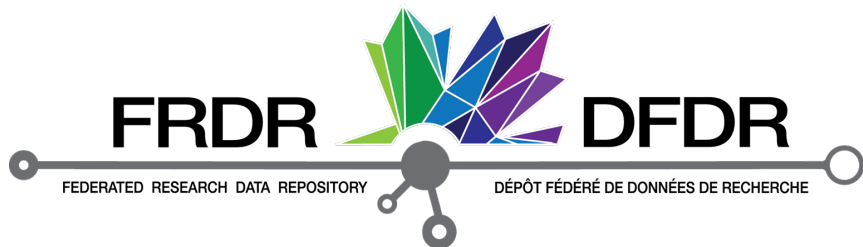


Le Dépôt fédéré de données de recherche Federated Research Data Repository

Services de Curation

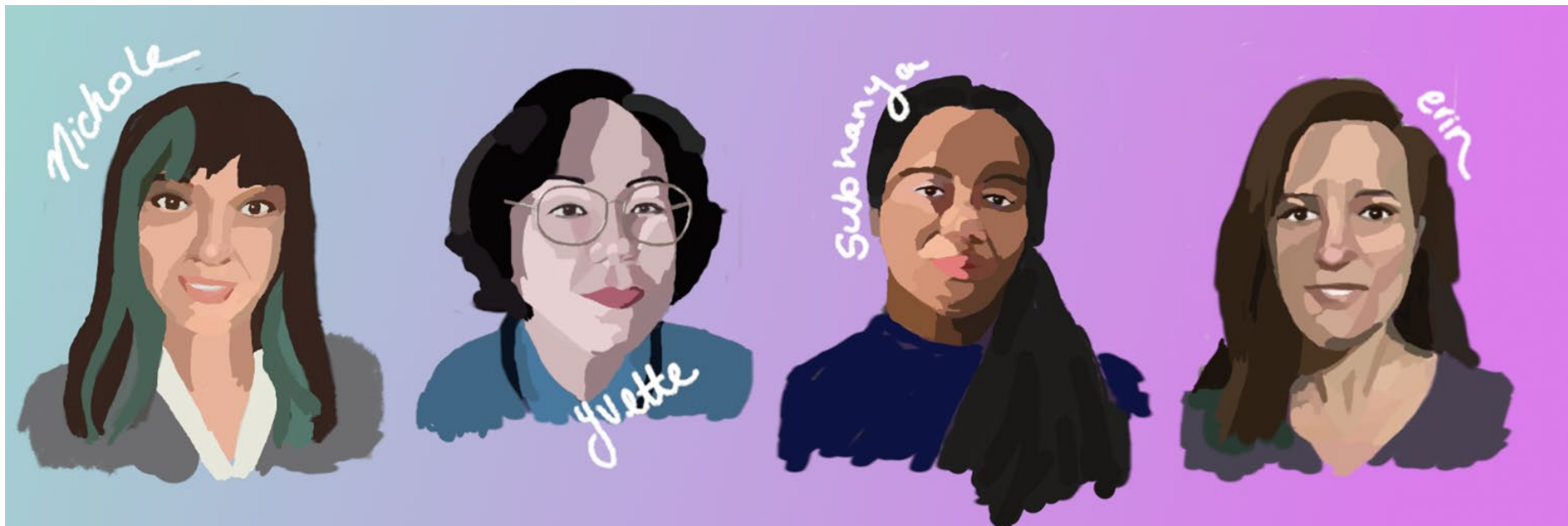


Nous  les données





Rencontrez les curateurs !





Le Dépôt fédéré de données de recherche (DFDR)

- Le DFDR est un service de découverte et un dépôt de données général pour :
 - Les chercheurs canadiens
 - Les données dans toutes les disciplines
 - Des ensembles de données de toutes tailles (bien adaptés aux ensembles de données volumineux)
 - 1 To d'espace de stockage par chercheur
 - Des données brutes ou traitées, provenant d'études antérieures ou actuelles
 - Des données associées ou non à des publications



Considérations de publication dans le DFDR

- Les données seront publiées et devraient donc être relativement inaltérables
- Toutes les données seront éventuellement accessibles au public
 - Les embargos temporaires sont autorisés
 - Les données sensibles doivent être supprimées ou masquées
 - Veuillez prendre en considération les ententes de partage de données que vous devez respecter
- Les données dont l'accès est restreint ne sont pas encore prises en charge, mais des travaux techniques et des considérations politiques sont en cours



Que se passe -t-il après le dépôt ?

- Un courriel de confirmation avec l'identifiant DOI (pas encore enregistré) est émis
- Le jeu de données n'est plus disponible pour l'édition
- Un membre de notre équipe de curation l'examinera
 - Le processus prend environ 2 à 5 jours ouvrables
- Publication de jeux de données !

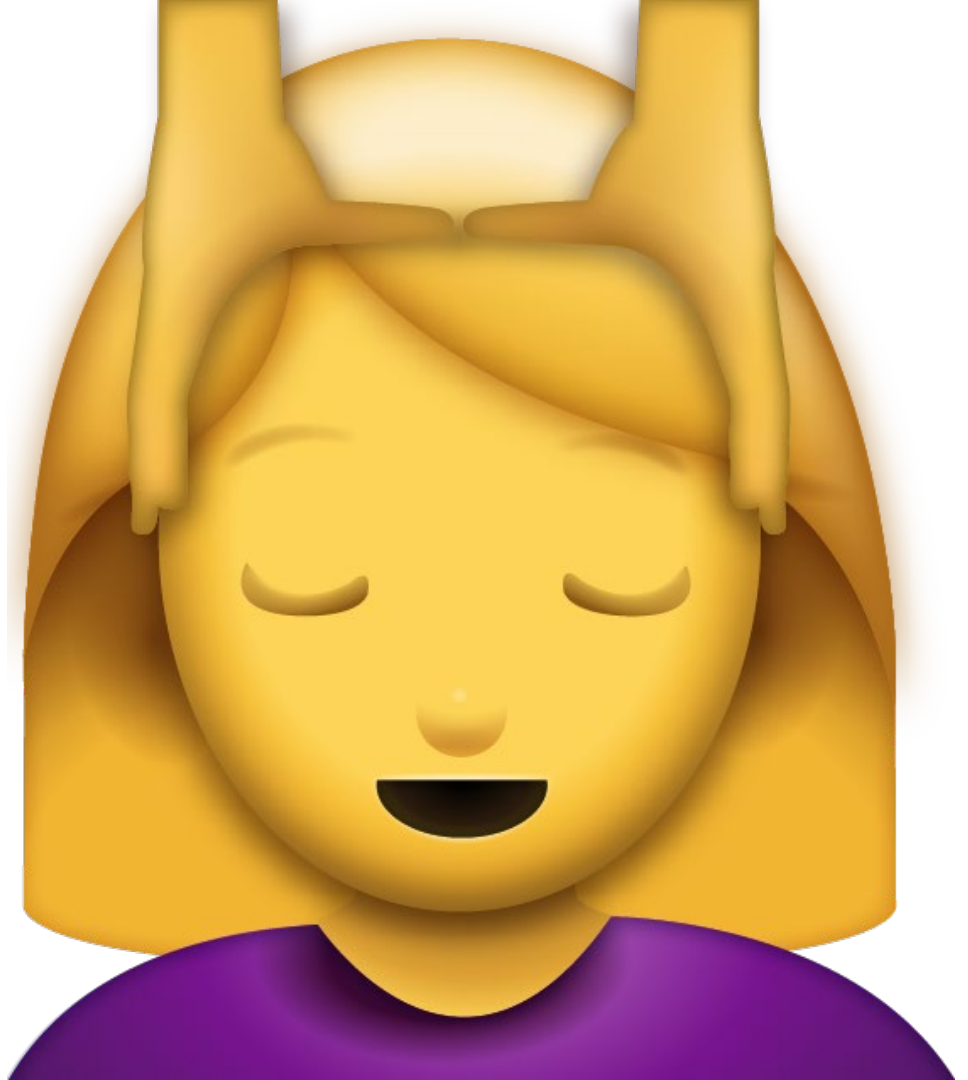
**En quoi consiste la
curation de données ?**





Commencer avec les
données.

Les rendre plus
conformes aux
principes FAIR



Faciles à trouver, accessibles, interopérables et réutilisables

La conformité aux principes FAIR est continue



Hydrometeorological observations at three boreal forest sites (aspen, jack pine, and black spruce) located in central Saskatchewan, Canada

Contact Dataset Administrator

Full metadata record		
DC Field	Value	Language
dc.contributor.author	Ahmed, Hafiz Faizan	
datacite.creator.affiliation	University of Saskatchewan	en_US
datacite.creator.nameIdentifier	https://orcid.org/0000-0002-7002-082X	en_US
dc.contributor.author	Helgason, Warren	
datacite.creator.affiliation	University of Saskatchewan	en_US
datacite.creator.nameIdentifier		en_US
dc.contributor.author	Barr, Alan G.	
datacite.creator.affiliation	Environment and Climate Change Canada	en_US
datacite.creator.nameIdentifier		en_US
dc.contributor.author	Black, T. Andrew	
datacite.creator.affiliation	University of British Columbia	en_US
datacite.creator.nameIdentifier		en_US
dc.date.accessioned	2020-10-29T00:30:41Z	
dc.date.available	2020-10-29T00:30:41Z	



Les données sont-elles **faciles à trouver** pour les humains et les ordinateurs?

F **Identifiant** persistant
Métadonnées riches
Indexées dans une ressource interrogeable

Est-ce que les utilisateurs peuvent **avoir accès aux données** requises une fois qu'ils les ont trouvées?

A Les formats de fichiers sont **ouverts**
Les fichiers **s'ouvrent correctement**

Les données peuvent-elles généralement être **intégrées** à d'autres données?

I **Un langage formel, accessible et partagé** pour la représentation des connaissances
Les références à d'autres (méta)données

Les données sont-elles **reproductibles** ?

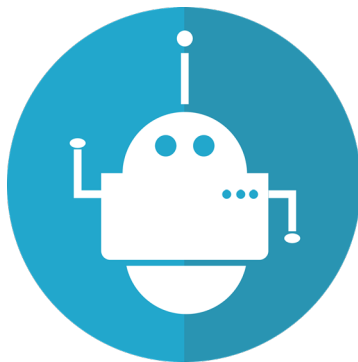
R Un contexte **adéquat**
Des attributs **précis** et pertinents
Une licence d'utilisation des données claire et accessible
Provenance détaillée

paraphrasé de Bishop, B. W., & Hank, C. (2018). Measuring FAIR principles to inform fitness for use.

Trouvable et utilisable par les systèmes informatiques et les humains

Les systèmes informatiques sont doués pour trouver des données qui ont de riches métadonnées, des vocabulaires contrôlés, des identifiants persistants et qui sont indexés dans un dépôt interrogeable.

Les systèmes informatiques savent utiliser des données en formats ouverts et bien structurés (par exemple, des données ordonnées)



Les humains sont doués pour découvrir des données qui ont des métadonnées descriptives, y compris des champs qui peuvent ne pas être indexés. Les fichiers README sont essentiels!

Les humains savent utiliser des données qui ont des variables clairement définies, des instructions, des conditions de licence, des citations, des commentaires, etc.



La curation tout au long du processus de recherche



Planifier → Créer → Traiter → Analyser → Diffuser → Préserver → Réutiliser

Réutiliser

Plan de gestion des données

Recherche

Curation

Préservation

Partage





La curation dans le cadre de la gestion des données de recherche

- Assure la fiabilité, l'exactitude et la validité des données
- Fait la promotion de l'organisation
- Améliore la reproductibilité de la recherche
- Assure la transparence et la responsabilité

La valeur ajoutée de la curation

- Renforce toute la valeur ajoutée de la GDR, plus largement !
 - Améliore la transparence
 - Augmente la probabilité de reproduction de la recherche
 - Et augmente la probabilité de citation correcte
- Plus de citations, plus de partage, plus de connaissances
 - Jusqu'à 25% de plus de citations lorsque les données sont partagées parallèlement à un article !
- Économies de coûts



Curation au DFDR



Consulter les fichiers

Un peu plus en profondeur

Recommander des changements

Améliorer les métadonnées

Transposer les formats de fichiers

Évaluer la conformité aux principes FAIR

Documenter les changements



Icône de FreePik

En pratique


- Arrangement et description
- Révision du code
- Services de recherche
- Documentation
- Évaluation de la conformité aux principes FAIR
- Transposer les formats de fichiers
- Inventaire des dossiers
- Métadonnées

Pour une liste plus exhaustive d'activités et de définitions de curation, consultez Johnston et coll. (2016) *Definitions of Data Curation Activities used by the Data Curation Network*. <http://hdl.handle.net/11299/188638> [en anglais].



Consulter & Un peu plus en profondeur

Fichiers dans le jeu de données - Taille totale - 893.38 KB

 01_wow_art.tiff	149.14 KB
 02_mysterious_art.tiff	123.54 KB
 03_landscape_art.tiff	620.66 KB
 data.csv	0.02 KB
 README.txt	0.01 KB

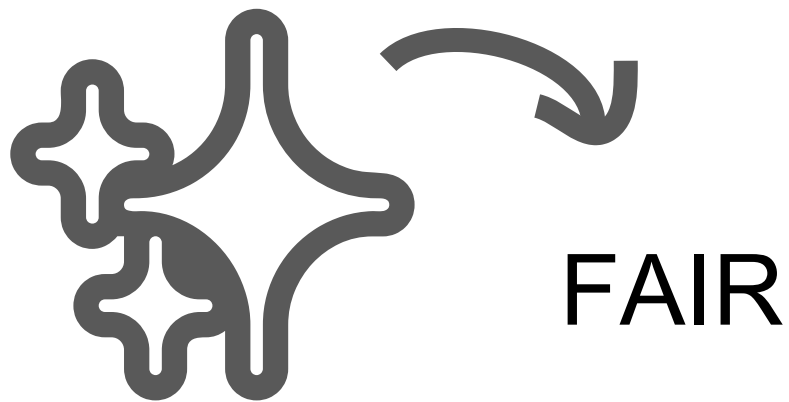


Recommander & Améliorer

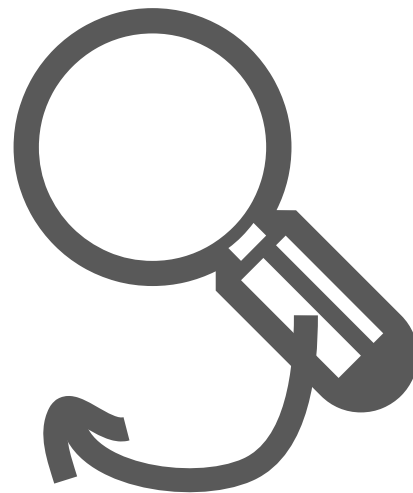
Élément Dublin Core	Valeur	Langue
dc.contributor.author	Rancourt, Y.G.	
datacite.creator.affiliation	Portage	en_US
datacite.creator.nameIdentifier	https://orcid.org/0000-0002-9895-8668	en_US
dc.description	A selection of Paint 3D works, created for this demo.	en_US
dc.description.provenance	Submitted by Y.G. Rancourt	en
dc.rights	Creative Commons Attribution 4.0 International (CC BY 4.0)	en_US
dc.rights.uri	https://creativecommons.org/licenses/by/4.0/	en_US
dc.subject	test	en_US
dc.subject	data	en_US
dc.subject	art	en_US
dc.subject	Paint 3D	en_US
dc.title	TEST Paint 3D Masterpieces	en_US



Transposer & Évaluer



FAIR





Documenter les changements et publier les jeux de données



Photo by [Pineapple Supply Co.](#) on [Unsplash](#)



Conseils et astuces



Documentation

Ce fichier LISEZMOI.txt a été généré le AAAA-MM-JJ par NOM

REMARQUE : Le texte d'aide est précédé des signes **##**. Le supprimer avant d'enregistrer ce fichier!

Il s'agit d'un modèle, n'hésitez pas à le modifier. Veuillez supprimer toute section ou tout élément qui ne s'applique pas à votre ensemble de données. Veuillez ajouter les renseignements supplémentaires qui ne se trouvent pas actuellement dans le modèle et qui sont nécessaires pour interpréter ou réutiliser correctement vos données.

RENSEIGNEMENTS GÉNÉRAUX

1. Titre de l'ensemble de données :

2. Renseignements sur l'auteur

A. Coordonnées du chercheur principal

Nom :

Établissement :

Adresse électronique :

B. Coordonnées de l'associé ou du cochercheur

Nom :

Établissement :

Adresse électronique :

3. Date de la collecte des données (date unique, intervalle, date approximative) :

Format suggéré : AAAA-MM-JJ

4. Emplacement géographique de la collecte des données :

latitude, longitude ou ville/région, province, pays, selon le cas

5. Renseignements sur les sources de financement qui ont appuyé la collecte des données :

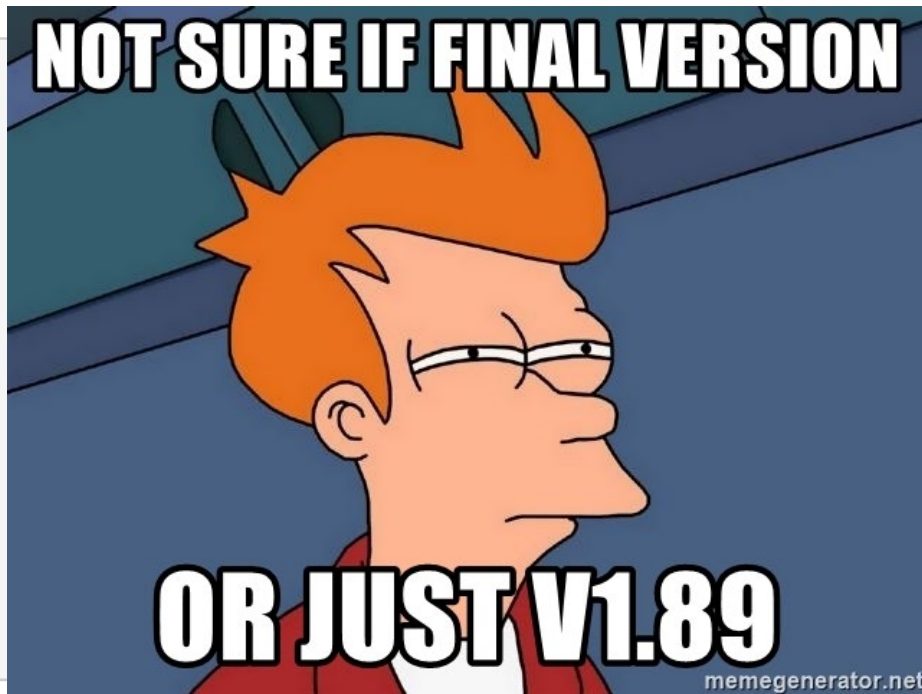


Métadonnées

Element	Notes	Ex.
Titre	Be concise and descriptive. Choose something appropriate for use in a data citation.	Dépenses des bibliothèques membres de l'ABRC pour les abonnements aux ressources savantes sous licence du RCDR pour l'année 2016 - 2017
Auteur	Dataset creators. Be prepared to provide author affiliations and ORCIDs.	Khair, Shahira A. Association des bibliothèques de recherche du Canada.
Description	It may help to think of the dataset description as you would an article abstract. This is a high-level summary that addresses the nature and scope of the dataset.	Ensemble de données et tableau sommaire des dépenses d'abonnement de 28 bibliothèques membres de l'Association des bibliothèques de recherche du Canada (ABRC) pour les ressources savantes sous licence du Réseau canadien de documentation pour la recherche (RCDR) pour l'exercice financier 2016 - 2017.
Mots-clés	Keywords that describe the dataset.	coûts d'abonnement; bibliotheques; ressources savantes
Identifiants connexes	Link to associated publications, code, protocols, methods, and other information that gives context to the data.	https://doi.org/10.5194/tc-14-1919-2359
Droits	Terms of use. The terms you can assign might be limited if your data are derived from external sources.	Creative Commons Attribution 4.0 International (CC BY 4.0)
Bailleurs de fonds	Granting agency or funder(s) that supported your research.	Conseil de recherches en sciences naturelles et en génie du Canada (NSERC)

Nomenclature, format et structure de fichier

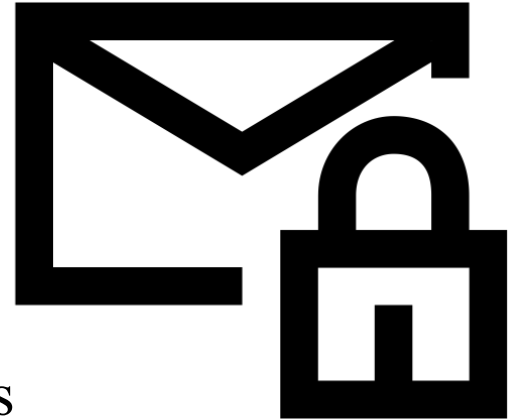
```
├─ Documentation
|   └─ site_information.csv
|   └─ site_1.shp
|   └─ site_2.shp
├─ Data
|   └─ year_01
|       └─ site_1.csv
|       └─ site_2.csv
|   └─ year_02
|       └─ site_1.csv
|       └─ site_2.csv
└─ README.txt
```





Susceptibilités potentielles

- Données d'organisations externes
- Informations sensibles sur les lieux
- Données provenant d'êtres humains



Created by Chameleon Design
from Noun Project

Merci! Questions?

Équipe de curation du DFDR :

Erin Clary, Nichole DeMichelis, Yvette Rancourt, Subhanya Sivajothy

support@frdr-dfdr.ca



compute
canada | **calcul**
canada



Innovation, Science and
Economic Development Canada

Innovation, Sciences et
Développement économique Canada

