

Introduction à la GDR : ce que vous devez savoir



Série de webinaires CAUL-CBUA sur la gestion des données de recherche (GDR)

14 octobre 2021

Nadia Zurek

Bibliothécaire, Université du Québec à Trois-Rivières

Membre du Groupe de travail sur la GDR du BCI



Objectifs de la séance

Objectif principal:

- Se familiariser avec les concepts clés de la gestion des données de recherche

Sous-objectifs:

- Comprendre l'importance de la GDR
- Connaitre le contexte canadien de la GDR
- Intégrer quelques bonnes pratiques en matière de GDR
- Identifier les organismes et les ressources en soutien à la GDR

Plan de la séance

- **Qu'est-ce que la GDR?**
- **L'importance de la GDR**
- **Les principes FAIR**
- **Le contexte canadien : la politique des trois organismes**
- **Les données ouvertes**
- **La GDR en détail**
 1. Plan de gestion des données (PGD)
 2. Bonnes pratiques
 3. Dépôts de données
- **Ressources nationales et locales**

Mise en contexte

Qu'est-ce que la GDR?



Que sont les données de recherche ?

- **De différents types et formats :**
 - Données d'observation, de tierces parties, factuelles, personnelles, etc.
 - Textes, chiffres, images, films, schémas, programmes informatiques, etc.
- **Essentielles à la recherche**
 - Preuve pour confirmer ou infirmer l'hypothèse

L'importance de la GDR

Pour les chercheurs

- Meilleure organisation du travail
- Conservation sécuritaire des données
- Protéger de la perte ou du vol des données
- Visibilité et impact
- Nouvelles collaborations
- Réutilisation des données
- Intégrité, transparence

Pour la science

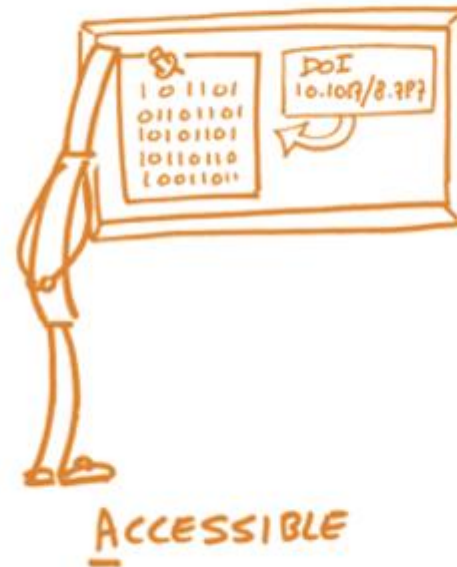
- Avancement accéléré des connaissances
- Reproductibilité de la recherche
- Limite la répétition des travaux
- Validation des résultats de la recherche

Les principes FAIR

FAIR DATA PRINCIPLES



Facilement repérable



Accessible



Interopérable



Réutilisable

Source : Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Brinken, H., & Heller, L. (2018). *Open Science Training Handbook*. <https://doi.org/10.5281/zenodo.1212496>

Les principes FAIR

Ressource en français sur les principes FAIR : [DoraNum](#)

Facilement repérable (findable):

- Identifiants uniques
- Métadonnées
- Dépôts de données

Accessible:

- Protocoles standards, libres et ouverts
- Authentification pour les données sensibles

Interopérable :

- Vocabulaire contrôlé
- Formats non propriétaires

Réutilisable :

- Métadonnées riches et exactes
- Contextualisation (ex: indication de la provenance)
- Licences

Le contexte canadien

Le contexte canadien

Les 3 organismes subventionnaires canadiens

CRSH  SSHRC

Sciences humaines et sociales



Sciences naturelles et génie



Sciences de la santé

Le contexte canadien

Organismes subventionnaires canadiens (CRSNG, CRSH, IRSC)

2016 : [Déclaration de principes](#) sur la gestion des données numériques

2018 : [Consultation publique](#)

mars 2021: [Politique des trois organismes sur la gestion des données de recherche](#) :

- ✓ Une **stratégie institutionnelle** de gestion des données (échéance : Mars 2023).
- ✓ Un **plan de gestion des données** lors du dépôt d'une demande de subvention (Dès le printemps 2022).
- ✓ Le **dépôt des données** et des métadonnées dans un dépôt numérique reconnu (Progressivement, après l'analyse des stratégies institutionnelles).

Ce n'est pas
une politique
de données
ouvertes!

Les données ouvertes

Le principe « aussi ouvert que possible, aussi fermé que nécessaire »

Directive du parlement européen et du conseil (2021)

Attention au partage des données « sensibles » :

- Enjeux éthiques (données personnelles ou identificatoires)
- Enjeux juridiques
- Enjeux commerciaux

Les données ouvertes



Est-ce que mes données sont partageables?

[Matrice de risque](#) lié aux données de recherche avec les êtres humains (Portage)

[Arbre décisionnel](#) pour le partage des données avec participants humains (Portage)

[Outil d'aide à la décision](#) sur le partage des données (FRQ)

Pour en savoir plus sur les données personnelles et leur confidentialité:

https://ethics.gc.ca/fra/depositing_deposits.html

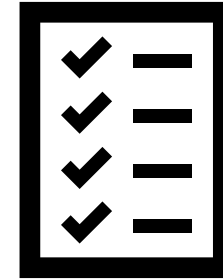
https://publications.gc.ca/collections/collection_2015/ger-pre/RR4-2-2014-fra.pdf (Chapitre 5)

<https://biblio.uottawa.ca/fr/services/professeurs/gestion-donnees-recherche/considerations-juridiques-ethiques>



La GDR en détail

1. Le plan de gestion des données (PGD)

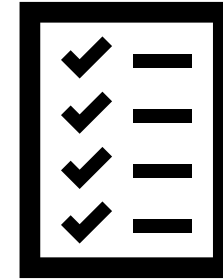


« [Le PGD] est un **document synthétique** qui aide à **organiser** et **anticiper** toutes les étapes du **cycle de vie de la donnée**. Il explique pour chaque jeu de données comment seront gérées les données d'un projet, depuis leur création ou collecte jusqu'à leur partage et leur archivage. »

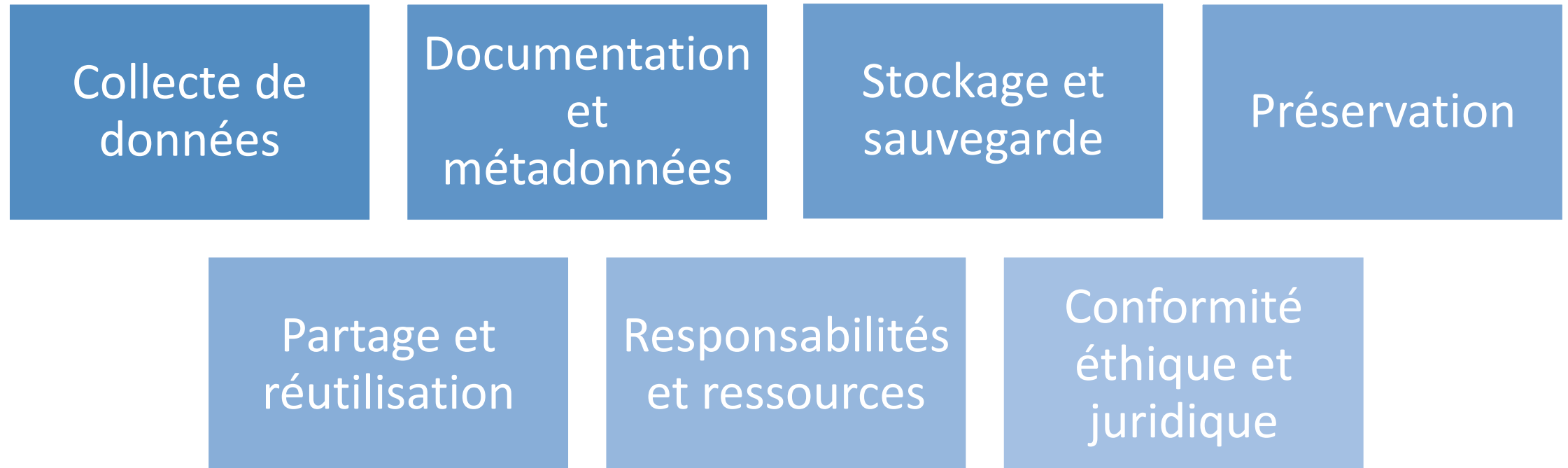
Source : DoRANum. (2018). *Fiche Synthétique Plan de gestion des données DMP*. Repéré à <https://dorum.fr/plan-gestion-donnees-dmp/fiche-synthetique/>

- Le plan est évolutif et doit être mis à jour

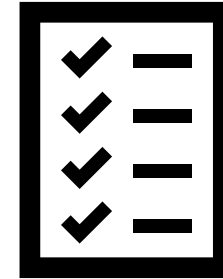
1. Le plan de gestion des données (PGD)



Le PGD contient environ une vingtaine de questions portant sur ces aspects:



1. Le plan de gestion des données (PGD)



Pour en savoir plus sur les PGD, inscrivez-vous [au webinaire](#) « À l'aide! J'ai besoin d'un plan de gestion de données » (28 octobre)!

2. Les bonnes pratiques en GDR



PLANIFIER

- Rédiger un PGD
- Obtenir le consentement des participants quant à la diffusion éventuelle des données avant leur collecte
- Identifier le bon processus pour protéger les données potentiellement sensibles (ex: anonymisation ou accès contrôlé)

2. Les bonnes pratiques en GDR



STOCKER ET SAUVEGARDER LES DONNÉES

- Choisir la plateforme de stockage de données : sécurité, coût, lieu des serveurs, récupération de fichiers, versionnage, etc.
- Suivre la règle du 3-2-1 :
 - 3 copies sauvegardées
 - sur 2 supports de stockage distincts,
 - dont 1 situé hors campus.
- Définir une fréquence de sauvegarde

Pour en savoir plus sur la plateforme [ACENET](#), inscrivez-vous [au webinaire](#) « Exploiter l'infrastructure de recherche numérique de pointe pour favoriser la gestion des données de recherche » (25 novembre)!

2. Les bonnes pratiques en GDR



ORGANISER LES DONNÉES

- Hiérarchie des dossiers
- Nommage des fichiers ([guide](#) de l'UQÀM)
- Documentation aidant à contextualiser les données:
 - Fichier LISEZMOI.txt (modèle en [français](#) et en [anglais](#))
 - [Dictionnaire de données](#) →
 - Métadonnées riches

Nom	Définition	Type	Longueur	Règle	Exemple
CodeP	Code du participant	AN	6		635958
CivilP	État civil du participant	A	20	5 types d'état civil	uni civilement
Datenaiss	Date de naissance du participant	D	10	aaaa-mm-jj	1980-01-01
NbEnf	Nombre d'enfants du participant	N	20		2

- **Uniformité, constance et logique**

2. Les bonnes pratiques en GDR



FORMAT DES FICHIERS

- Privilégier des formats de fichiers ouverts
- Dans certaines disciplines, les formats propriétaires sont acceptés
- Consulter ce [tableau](#) de UK Data Service sur les formats recommandés
Ex: .txt; .csv; .xml; .tif; .mp4.
- Objectifs:
 - ✓ Préservation à long terme
 - ✓ Éviter les erreurs de conversion
 - ✓ Permettre un meilleur partage des données

Exercice

Enquête sur la moutarde

questionnaire_2021.pdf

enquetesurlamoutarde_lesconsignes.txt

Enquête_moutarde_données_2021_1.xlsx

Enquête_moutarde_données_2021_2.xlsx

VideoATremblay.mp4

VerbatimATremblay.doc

VideoDSmith.mp4

VerbatimDSmith.doc

VideoJ.Jacques.mp4

VerbatimJ.Jacques.doc

VideoR.St-Pierre.mp4

VerbatimR.St-Pierre.doc

Enquete sur la moutarde

Documentation

Enquete_moutarde_questionnaire_2021.pdf

Enquete_moutarde_consignes.txt

Enquete_moutarde_Dictionnaire_Donnees_2021.pdf

Lisez-moi_Enquete_moutarde_2021_2021-09-27.txt

Donnees_sondage : données tabulaires

Enquete_moutarde_donnees_2021_v01.csv

Enquete_moutarde_donnees_2021_v02.csv

Verbatim : verbatim des entrevues

Enquete_moutarde_Verbatim_P1_2021.txt

Enquete_moutarde_Verbatim_P2_2021.txt

Enquete_moutarde_Verbatim_P3_2021.txt

Enquete_moutarde_Verbatim_P4_2021.txt

Video (si autorisé): vidéos des entrevues

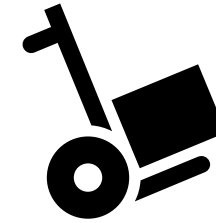
Enquete_moutarde_Video_P1_2021.mp4

Enquete_moutarde_Video_P2_2021.mp4

Enquete_moutarde_Video_P3_2021.mp4

Enquete_moutarde_Video_P4_2021.mp4

3. Les dépôts de données



Types de dépôts :

Spécialisé
(ex: [Pangea](#), [GenBank](#), [PsychData](#),...)

PANGAEA. GenBank
PsychData

Généraliste
(ex: DFDR – [vidéo](#), [Zenodo](#), [OSF](#),...)

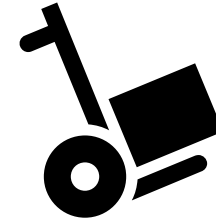
DFDR DÉPÔT FÉDÉRÉ DE DONNÉES DE RECHERCHE
zenodo
OSF

Institutionnel
(s'il est disponible, ex: Dataverse)

The Dataverse[®] Project

Trouver un dépôt : re3data.org

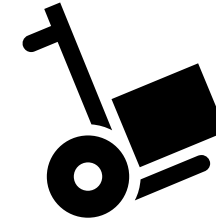
3. Les dépôts de données



Choisir un dépôt – Les priorités:

1. Exigences de l'organisme subventionnaire ou de l'éditeur
2. Dépôt spécialisé
3. Dépôt institutionnel ou généraliste

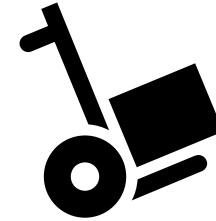
3. Les dépôts de données



Choisir un dépôt – Critères à évaluer :

- Réputation
- Pays de stockage des données
- Coûts et volume des données
- Langue pour la description
- Identifiant unique (doi)
- Administrateurs des jeux de données (niveau de contrôle)
- Types de licences et de formats
- Gestion des accès restreints

3. Les dépôts de données



Pour en savoir plus sur les dépôts de données, inscrivez-vous [au webinaire](#) « Où déposer vos données – Dataverse? DFDR? Autre? » (10 novembre)!

Besoin d'aide?

Ressources nationales

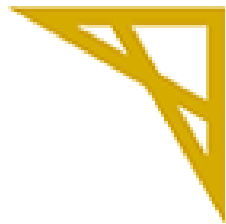
[Assistant PGD](#): un outil d'aide à la préparation d'un PGD en ligne

Réseau Portage - [Ressources](#):

- Guides
- Webinaires et vidéos
- Dictionnaires et glossaires



Nouveau nom :



**Alliance de recherche
numérique** du Canada

Ressources locales

Pages web ou guides sur la GDR dans les universités

Bibliothécaire disciplinaire

Bibliothécaire en soutien à la recherche ou même spécialiste en GDR

Questions précises? Consultez les pages web de votre institution

- Enjeux éthiques (ex: demande d'approbation éthique): Comité éthique de la recherche avec les êtres humains
- Propriété intellectuelle : Bureau du droit d'auteur/de la propriété intellectuelle
- Logiciels et serveurs : Service de technologie informatique, Centre de services technologiques



Merci!

Questions?



Sources

Parlement européen, Conseil de l'Union européenne. (2021). Directive du parlement européen et du conseil. *Journal officiel de l'Union européenne*. Repéré à <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32021L1024&from=EN>

Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Brinken, H., & Heller, L. (2018). *Open Science Training Handbook*. <https://doi.org/10.5281/zenodo.1212496>

URFIST Méditerranée. (2021). Les principes FAIR. <https://doi.org/10.13143/Z7S6-ED26>

Lewis & Ruth Sherman Centre for Digital Scholarship et McMaster University. (2021). *Best Practices for Managing Data in your Research*. Repéré à <https://scds.github.io/intro-rdm/intro.html>.

Dennie, D., Papin, D. et Du Ruisseau, M. (2020, 7 mai). *Webinaire GT-GDR BCI : Introduction au choix d'un dépôt de données de recherche* [vidéo]. YouTube. <https://www.youtube.com/watch?v=9Q2d45sZ4VE&t=2s>

Bibliothèques de l'UQÀM. (s. d.). *Recommandations pour le nommage de vos fichiers électroniques*. Repéré à http://guides.bibliotheques.uqam.ca/docs/gestion_donnees_recherche/recommandations_nommage_fichiers.pdf