

# **Sustaining our data through data management: it's easier than you think!**

IASSIST 2018  
Jane Fry and Carol Perry  
Thursday May 25, 2018

# Outline

- What is Portage
- Latest training resources
- Other training aids
- Data Management Plan Assistant
- Over to you!

# What is Portage

- A brief history of Portage

# Latest training resources



## Latest training resources (cont'd)

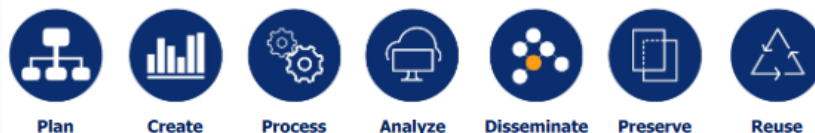
- Portage Expert Training Group
- First deliverable
  - White Paper: [\*RDM Training Landscape in Canada\*](#)
- Our [\*Training Principles\*](#)
  - All training materials put together by Portage are bilingual
- Another point
  - The materials are also freely available to anyone

# RESEARCH DATA MANAGEMENT PRIMER

## Research Data Management (RDM)

- RDM refers to the processes applied throughout the lifecycle of a research project to guide the collection, documentation, storage, sharing, and preservation of research data.
- RDM practices are integral to conducting responsible research and can help researchers save resources by ensuring their data is complete, understandable, and secure.
- RDM practices also follow institutional and funding agency guidelines that protect their investments.
- The broader research community can derive maximum value from research data that can be accessed, shared, reused and repurposed.

## The Research Data Lifecycle



\* Life cycle model developed by the Leadership Council for Digital Research Infrastructure. For more information visit <http://digitalleadership.ca>

## Defining Research Data

- Primary sources supporting research, scholarship or artistic endeavours
- Can be used as evidence to validate findings and results
- May take the form of experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data
- All other digital and non-digital content have the potential to become research data

\*Research data. (n.d.) In CASRAI's Dictionary. Retrieved from [dictionary.casrai.org/Research\\_data](http://dictionary.casrai.org/Research_data)

# RESEARCH DATA MANAGEMENT PRIMER

## Canadian and International RDM Resources



## Canadian RDM Policies

- Canada's "Action Plan on Open Government" outlines the development and adoption of policies, guidelines and tools to support the effective stewardship of scientific data.
- Canada's Tri-Council granting agencies (CIHR, NSERC, SSHRC) have adopted a "Statement of Principles on Digital Data Management", which identifies expectations and responsibilities for the management of data produced with public funding. Completing data management plans can help to meet these expectations and responsibilities.

Contact at your institution:



# GESTION DES DONNÉES DE RECHERCHE

## INFORMATIONS DE BASE

### Gestion des données de recherche (GDR)

- La GDR est l'ensemble des 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.
- Les pratiques de la GDR sont directement liées à la recherche responsable et peuvent aider les chercheurs à économiser des ressources en s'assurant que leurs données sont complètes, compréhensibles et sécurisées.
- Également, les pratiques de la GDR suivent les politiques des institutions et des organismes subventionnaires qui cherchent à protéger leurs investissements.
- L'ensemble des milieux de la recherche peuvent tirer le maximum de valeur des données de recherche lorsqu'elles sont accessibles, partagées, réutilisées et exploitées différemment.

### Le cycle de vie des données de recherche



Le modèle du cycle de vie des données de recherche a été créé par le Conseil du leadership sur l'infrastructure de recherche numérique. Pour plus d'informations, consultez leur site Web : <http://leadershipnumerique.ca>.

### Comment définir les données de recherche

- Des sources primaires qui soutiennent des projets de recherche, des études académiques ou des travaux artistiques.
- Elles peuvent être utilisées comme preuve pour valider des résultats.
- Elles peuvent prendre la forme de données expérimentales, de données d'observation, de données opérationnelles, de données de tiers, de données du secteur public, de données de suivi, de données traitées ou de données réutilisées.
- Tout autre contenu numérique et non numérique a le potentiel de devenir des données de recherche.

\*Research data. (n.d.) Dans le CASRAI Dictionary. Repéré à [dictionary.casrai.org/Research\\_data](http://dictionary.casrai.org/Research_data)

# GESTION DES DONNÉES DE RECHERCHE

## INFORMATION DE BASE

### Ressources canadiennes et internationales en GDR



### Politiques canadiennes sur la GDR

- Le Plan d'action du Canada pour un gouvernement ouvert décrit le développement et l'adoption des politiques, des lignes directrices et des outils pour soutenir la gestion efficace des données scientifiques.
- Les trois organismes subventionnaires du gouvernement fédéral (IRSC, CRSNG et CRSH) ont adopté une Déclaration de principes sur la gestion des données numériques qui décrit les attentes et les responsabilités en matière de gestion des données produites grâce aux fonds publics. L'élaboration de plans de gestion de données peut aider à répondre à ces attentes et responsabilités.

# "Good Enough" Research Data Management

(a brief guide for busy people)

This brief guide presents a set of good data management practices that researchers can adopt, regardless of their data management skills and levels of expertise.

1

## Save your raw data in original format

- 1.1 Don't overwrite your original data with a cleaned version.
- 1.2 Protect your original data by locking them or making them read-only.
- 1.3 Refer to this original data if things go wrong (as they often do).

2

## Backup your data

- 2.1 Use the 3-2-1 rule: Save three copies of your data, on two different storage mediums, and one copy off site.
- 2.2 Do not backup or store sensitive data on a commercial cloud (Dropbox, Google Drive, etc.).

3

## Describe your data

- 3.1 **Machine Friendly:** Describe your dataset with a metadata standard for discovery.
- 3.2 **Human Friendly:** Describe your variables, so your colleagues will understand what you meant. Data without good metadata is useless. Give your variables clear names.
- 3.3 Do not leave cells blank - use numeric values clearly out of range to define missing (e.g. '99999') or not applicable (e.g. '88888') data, and describe these in your data dictionary.
- 3.4 Convert your data to open, non-proprietary formats.
- 3.5 Name your files well with basic metadata in file names.

4

## Process your data

- 4.1 Make each column a variable.
- 4.2 Make each row an observation.
- 4.3 Store units (e.g. kg or cm) as metadata (in their own column).
- 4.4 Document each step processing your data in a README file.

5

## Archive and preserve your data

- 5.1 Submit final data files to a repository assigning a persistent identifier (e.g. handles or DOIs).
- 5.2 Provide good metadata for your study so others could find it (use your discipline's metadata standard, e.g. Darwin Core, DDI, etc.).

portage

This guide was reproduced by the Portage Training Expert Group with permission from its original creator, Eugene Bersky, University of British Columbia. It can be modified and re-used freely under the CC-BY license.



# De bonnes pratiques en gestion des données de recherche

(Un guide sommaire pour gens occupés)

Ce petit guide présente un ensemble de bonnes pratiques que les chercheurs peuvent adopter, et ce, indépendamment de leurs compétences ou de leur niveau d'expertise.

1

## Enregistrer vos données brutes dans leur format original

- 1.1 N'écrivez pas vos données avec une version corrigée.
- 1.2 Protégez vos données originales en verrouillant vos fichiers (avec un mot de passe) ou en définissant un accès en lecture seule.
- 1.3 Référez-vous à ces données originales lorsque les choses tournent mal (comme c'est souvent le cas).

2

## Sauvegarder vos données

- 2.1 Utilisez la règle du 3-2-1: enregistrez trois copies de vos données sur deux supports de stockage différents et conservez-en une dans un lieu hors site.
- 2.2 Ne sauvegardez pas ou ne stockez pas des données sensibles dans un nuage commercial (Dropbox, Google Drive, etc.).

3

## Décrire vos données

- 3.1 **Pour qu'elles soient facilement lisibles par les ordinateurs :** Décrivez votre ensemble de données selon une norme de métadonnées facilitant la découverte.
- 3.2 **D'une manière conviviale pour les humains :** Décrivez vos variables de façon à ce qu'elles soient facilement compréhensibles pour vos collègues. Des données sans bonnes métadonnées sont inutilisables. Donnez des noms clairs à vos variables.
- 3.3 Ne laissez pas les cellules vides - utilisez des valeurs numériques clairement hors limites pour définir les données manquantes (p. ex., '99999') ou non applicables (p. ex., '88888'), et décrivez-les dans votre dictionnaire de données.
- 3.4 Convertissez vos données dans des formats ouverts et non-propriétaires.
- 3.5 Nommez bien vos fichiers en utilisant des métadonnées de base dans le nom des fichiers.

4

## Traiter vos données

- 4.1 Faites de chaque colonne une variable.
- 4.2 Faites de chaque rangée une observation.
- 4.3 Stockez les unités (tels que kg ou cm) en tant que métadonnées (avec leur propre colonne).
- 4.4 Documentez chaque étape du traitement de vos données dans un fichier « LISEZ-MOI ».

5

## Archiver et préserver vos données

- 5.1 Soumettez vos fichiers de données finaux dans un dépôt de données en leur attribuant un identificateur permanent (DOIs ou autres Handles).
- 5.2 Fournissez de bonnes métadonnées à votre étude de manière à ce que les autres puissent la repérer plus facilement (utilisez les normes de métadonnées de votre discipline telles que Darwin Core, DDI, etc.).

portage

Ce guide a été reproduit par le Groupe d'experts sur la formation en GDR de Portage avec la permission de son créateur original, Eugene Bersky, University of British Columbia. Il peut être modifié et réutilisé gratuitement sous la licence CC-BY.





## Latest training resources (cont'd)

- Online Training modules
  - in progress
    - RDM 101
    - DMP Assistant 101
  - In the future
    - Training repositories 101
    - ...

# Other training aids

Portage RDM Training Assistance Request Form

External Training Resources

## Events

---

### Upcoming Events

There are no upcoming events.

### Past Events

## Portage RDM Training Assistance Request Form

Welcome to the Portage Research Data Management (RDM) Training Assistance Request Form.

This form is intended for those organizing RDM training on Canadian academic campuses, and who would like assistance developing training and guidance materials in all areas of RDM. In responding to requests for training assistance, the Portage Training Expert Group (TEG) will draw upon the wealth of knowledge and expertise of the Portage network to assist applicants in developing educational materials and programs.

The wider academic community is also invited to submit RDM training suggestions through this form. In response, TEG will share these requests with local contacts and coordinate RDM training opportunities that respond to the request.

We look forward to helping support the development of RDM knowledge and capacity within your organization. This form will take around 5 minutes to complete. We ask that you provide as much information as possible as this will help us to evaluate the nature of your request and the resources being requested.

8%

Next page

Save and continue later

Powered by [SimpleSurvey](#)

## Formulaire de demande d'une aide en matière de formation en GDR de Portage

Bienvenue sur le formulaire de demande d'aide à la formation sur la gestion des données de recherche (GDR) du réseau Portage.

Ce formulaire est mis à la disposition des personnes qui organisent la formation en GDR sur les campus des universités canadiennes et qui aimeraient avoir de l'aide pour élaborer du matériel de formation pouvant toucher à tous les domaines de la GDR. En réponse aux demandes d'aide à la formation, le groupe d'experts sur la formation puisera parmi la vaste quantité de connaissances et d'expertise du réseau Portage pour aider les demandeurs d'aide à élaborer des programmes et des documents pédagogiques.

Le milieu universitaire dans son ensemble est invité à soumettre des propositions de formation en GDR au moyen de ce formulaire. Les membres du groupe d'experts sur la formation partageront ces demandes avec des contacts locaux et organiseront des formations en GDR qui répondent à la demande.

Nous sommes heureux à l'idée de contribuer au développement des connaissances et des capacités en GDR au sein de votre organisation. Ce formulaire prendra environ cinq minutes à remplir. Nous vous demandons de fournir autant d'informations que possible. Ceci nous aidera à mieux évaluer la nature de votre demande ainsi que les ressources qui sont requises.

8%

Page suivante

Sauvegarder et continuer plus tard

Propulsé par [SimpleSondage](#)

# External Training Resource Library

The Portage Training Expert Group has identified and compiled a diverse range of external resources useful for learning and training in the domain of research data management (RDM), which we invite you to browse and explore.

Please complete the following [Resource Submission Form](#) to recommend additional RDM training resources for this resource library.

## 23 (research data) Things

(Australian National Data Service)

A self-directed learning program developed by the Australian National Data Service for audiences from all communities of practice who want ... [More >>](#)

## 23 Things: Libraries for Research Data

(Research Data Alliance)

An overview of practical, free, online resources and tools to incorporate research data management into the practice of librarianship. Proceed ... [More >>](#)

## Advice and Training

(UK Data Service)

A repository of advice and training materials created by the UK Data Service.

### Filters:

#### Order

Sort by

#### Producer

Any

#### Intended Audience

Any

#### Mode of Delivery

Any

#### Learning Process

Any

#### Related Discipline

Any

#### Cost

Any

#### Language

Any

#### Region of Origin

Any



## Terms of reference

- Provide data management planning resources (tools, templates, guidance, training) that support best practices in research data stewardship.....
- Improve local and national research service coordination
- Facilitate sharing of research administrative information
- Reduce administrative burden on researchers

# DMP Assistant - L'Assistant PGD



[Version française](#)

[Home](#)

[About](#)

[Help](#)

[Terms of Use](#)

**DMP Assistant** is a bilingual tool for preparing data management plans (DMPs). The tool follows best practices in data stewardship and walks researchers step-by-step through key questions about data management.

- Step 1** Sign up with DMP Assistant
- Step 2** Sign in and select a template under Organizations. The Portage template is the default.
- Step 3** Answer the questions that are relevant to your work. Guidance and examples are provided.
- Step 4** Revisit the tool throughout your research to review or revise your answers.

## Sign in




If you have an existing account with DMP Assistant or previous version of DMP Builder.

## Sign up



New to DMP Assistant? Sign up today.

# Customize



[Version française](#)

Signed in as Carol Perry

[My plans](#)[Create plan](#)[About](#)[Help](#)

[Terms of Use](#)

## Create a new plan

Please select from the following drop-downs so we can determine what questions to ask you.

If you aren't responding to specific requirements from a funder or an institution, you can select "Not applicable/not listed".

The Portage Data Stewardship Template is based on internationally accepted standards and is maintained by a group of research data management experts from research data management centers.

**To see institutional questions and/or guidance, select your organization.**

Royal Roads University

Ryerson University

Scholars Portal

University of Calgary

University of Guelph

University of Lethbridge

Université de Moncton

University of New Brunswick

Portage

You may leave blank or select a different organization to your own. If you leave blank, default Portage DMP template will be used [Not applicable/not listed](#).

Create plan

# My plan (Portage Template)

0/20 questions answered

## Tips

Not all questions will apply to all research projects. Researchers are encouraged to answer the questions relevant to their work.

Researchers should revisit the tool throughout their research to review or complete their responses.

Plan details

Portage Data Management Questions

Share

Export

Data Collection (3 questions, 0 answered)



Documentation and Metadata (3 questions, 0 answered)



Storage and Backup (3 questions, 0 answered)



Preservation (2 questions, 0 answered)



Sharing and Reuse (3 questions, 0 answered)



Responsibilities and Resources (3 questions, 0 answered)



Ethics and Legal Compliance (3 questions, 0 answered)



Export



Plan details
Portage Data Management Questions
Share
Export

Data Collection (3 questions, 0 answered )

This section addresses data collection issues such as data types, file formats, naming conventions, and data organisation – factors that will improve the usability of your data and contribute to the success of your project.

What types of data will you collect, create, link to, acquire and/or record?

B
I
A
A
List
List
Link
Table

Save
Not answered yet

Local guidance

Guidance
Share note

Portage Guidance
Examples: numeric, images, audio, video, text, tabular data, modeling data, spatial data, instrumentation data.

University of Guelph guidance
The University of Guelph Library's [Data Resource Centre](#) provides data collection and creation support including support for the design and creation of web surveys and access to data resources including numeric and geospatial databases.
The Library also hosts the [University of Guelph Branch Research Data Centre \(BRDC\)](#) where approved academic staff and students can securely access detailed microdata and administrative datasets from Statistics Canada.
For more information, contact us at [lib.research@uoguelph.ca](mailto:lib.research@uoguelph.ca).

# DMP Assistant – Sharing

Plan details

Portage Data Management Questions

Share

Export

You can give other people access to your plan here. There are three permission levels.

Users with "read only" access can only read the plan.

Editors can contribute to the plan.

Co-owners can also contribute to the plan, but additionally can edit the plan details and control access to the plan.

Add each collaborator in turn by entering their email address below, choosing a permission level and clicking "Add collaborator".

Those you invite will receive an email notification that they have access to this plan, inviting them to register with DMP Assistant if they don't already have an account. A notification is also issued when a user's permission level is changed.

## Collaborators

Email address	Permissions	
Carol Perry	Owner	

## Add collaborator

Permissions:

 ☒

Add collaborator

# DMP Assistant – What's next

Moving this summer to DMP Assistant 2.0

- new Roadmap version

## Major features

- Option to publish DMPs in the “Share” tab
  - (subject to 50% of questions answered)
- Public repository for templates and DMPs
- Improved user statistics for administrators

✓ Notice: Signed out successfully.

# Welcome.

DMPRoadmap has been developed by the **Curation Center** to help you write data management plans.

## Getting started:

- [Digital Curation Centre](#)
- [UC3: University of California Curation Center](#)
- [UK funder requirements for Data Management Plans](#)
- [US funder requirements for Data Management Plans](#)
- [DCC Checklist for a Data Management Plan](#)

[Sign in](#)[Create account](#)

\* Email

\* Password

[Forgot password?](#)

☐ Remember email

Sign in



# DMP Expert Group work

- Two new working groups
  - Exemplar working group
    - Develop/share exemplar DMPs focused on different disciplines
  - Policy working group
    - Review policies & best practices globally
    - Prepare guidance documents for major stakeholders

## Over to you!

- What are we missing?
- What training resources do you want to see?
- How can we help you?

## In sum ...

- We hope you have learned
  - What Portage is
  - All about our latest training resources
  - The other training aids available for you
  - All about the Portage Data Management Plan Assistant
- We also hope that you will contact us if there are other ideas you have



Source: <https://hiveminer.com/Tags/park/portageing>





Source: <https://hiveminer.com/Tags/park,portageing>





Source: <https://hiveminer.com/Tags/park/portageing>

# Thank you!

## Contact Information

Jane Fry (Carleton University)

[jane.fry@carleton.ca](mailto:jane.fry@carleton.ca)

Carol Perry (University of Guelph)

[carolp@uoguelph.ca](mailto:carolp@uoguelph.ca)