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 (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**

1. Plan
2. Create
3. Process
4. Analyze
5. Disseminate
6. Preserve
7. Reuse

---

**Defining Research Data**

- Primary sources supporting research, scholarship or artistic endeavors
- 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


---

**Canadian and International RDM Resources**

- UK Data Archive [www.data-archive.org.uk](http://www.data-archive.org.uk)
- Digital Curation Centre [www.dcc.ac.uk](http://www.dcc.ac.uk)
- Australian National Data Service [www.datags.org.au](http://www.datags.org.au)
- Portage Network [www.portagenetwork.ca](http://www.portagenetwork.ca)
- Research Data Canada [www.resdata.ca](http://www.resdata.ca)

---

**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 (CTRI, 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:**

---

The text is from [Portage Network’s RDM Primer](https://portagenetwork.ca/wp-content/uploads/2018/03/Portage-RDM-Primer-Contact-EN.pdf).
GESTION DES DONNÉES DE RECHERCHE
INFORMATIONS DE BASE

Le modèle du cycle de vie des données de recherche a été créé par le Conseil du leadership sur l’infrastructure de recherche du Canada. Pour plus d’informations, consultez le site Web : https://portagenetwork.ca/

Gestion des données de recherche (GDR)

- Le 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 essentielles à la recherche responsable et peuvent aider les chercheurs à éconомiser du temps et du temps d’investissement dans leurs projets.
- Les spécifications de la GDR servent de base à la mise en œuvre des politiques et des processus pour promouvoir et protéger la reproductibilité des travaux scientifiques.
- L’ensemble des milieux de la recherche peuvent tirer le maximum de valeur des données de recherche largement utilisées, partagées, réutilisées et exploitées différemment.

Le cycle de vie des données de recherche

Planification  Création  Traitement  Analyse  Diffusion  Conservation  Réutilisation

Comment définir les données de recherche

- Des sources primaires qui sont le résultat 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 ou de données de transactions 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.

Politiques canadiennes sur la GDR

- Les trois organismes subventionnaires du gouvernement fédéral (CPS, CRNP 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 au fonds public. 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 3-4-1 rule: Store 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 cloud service (Dropbox, Google Drive, etc.).

3. Describe your data
   3.2. Human Friendly: Describe your variables, so your colleagues will understand what you mean. Data without good metadata is useless. Give your variables clear names.
   3.3. Do not leave cells blank – use numeric values classed out of range to define missing (e.g. 999999 or 0). Describe these values 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 columns).

5. Archive and preserve your data
   5.1. Submit final data files to a repository (a persistent identifier (e.g. handles or DOI)).
   5.2. Provide good metadata for your study so others could find it (use your discipline’s metadata standard, e.g. Business Core, DM, etc.).

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 es, indépendamment de leurs compétences ou de leur niveau d’expertise.

1. Décrire vos données
   1.1. N’oubliez pas vos données avec une version unique.
   1.2. Proposez vos données originales en variant vos fichiers avec un emplacement de données ou en définissant leurs en lecture seuls.
   1.3. Reliez-vous à des données originales (c’est à dire les fichiers numérotés et nommés qui sont au-dessous).

2. Sauvegarder vos données
   2.1. Utilisez le système de 3-4-1, enregistrez trois copies de vos données sur deux supports de stockage différentiels et conservez-les sur un de vos sites.
   2.2. Ne sauvegardez pas ou ne stockez pas des données sensibles dans un espace commercial (Dropbox, Google Drive, etc.).

3. Traiter vos données
   3.1. Pour qu’elles soient facilement utilisées par les utilisateurs, déterminez votre ensemble de données selon une norme de standardisation facilitant la découverte.
   3.2. Store units (e.g. kg or cm) as metadata (in their own columns).
   3.3. Ne laissez pas les cellules vides – utilisez des valeurs manquantes clairement et de manière limitée pour décrire les situations manquantes (e.g. NON en non-propriétaires (e.g. en.8080). Stockez les données dans notre standardisation de données.
   3.4. Conservez vos données dans des formats ouvertes et non-propriétaires.
   3.5. Nommiez bien vos fichiers en utilisant des standardisations de base dans le nom des fichiers.
Online Training modules
  – in progress
    • RDM 101
    • DMP Assistant 101
  – In the future
    • Training repositories 101
    • …
Other training aids

Events

Upcoming Events
There are no upcoming events.

Past Events

Portage RDM Training Assistance Request Form

External Training Resources
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.

Powered by SimpleSurvey
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.

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

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.
DMP Expert Group

**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 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.

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

New to DMP Assistant? Sign up today.
Create a new plan

Please select from the following drop-downs so we can determine what guidelines to apply.

If you aren't responding to specific requirements from a funder or an institution, the Portage Data Stewardship Template is based on internationally accepted standards maintained by a group of research data management experts from research institutions.

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

You may leave blank or select a different organization to your own. If you leave blank, the default Portage DMP template will be used.

Not applicable/not listed.
## My plan (Portage Template)

### 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.

<table>
<thead>
<tr>
<th>Plan Details</th>
<th>Portage Data Management Questions</th>
<th>Share</th>
<th>Export</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Collection</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Documentation and Metadata</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Storage and Backup</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Preservation</td>
<td>(2 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Sharing and Reuse</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Responsibilities and Resources</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ethics and Legal Compliance</td>
<td>(3 questions, 0 answered)</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
DMP Portage Template

Local guidance
DMP Assistant – Sharing

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

<table>
<thead>
<tr>
<th>Email address</th>
<th>Permissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carol Perry</td>
<td>Owner</td>
</tr>
</tbody>
</table>

Add collaborator

Email

Permissions:
- Read only

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
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
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
Thank you!

Contact Information

Jane Fry (Carleton University)
  jane.fry@carleton.ca

Carol Perry (University of Guelph)
  carolp@uoguelph.ca