Data Management Plans

James Doiron
RDM Services Coordinator, UofA Libraries
Academic Director, UofA Research Data Centre
Co-Chair, Portage DMP Expert Group

Shifting Horizons II: Realities of RDM Services
University of Ottawa
February 20, 2020
What is Research Data Management?

RDM = managing data throughout all phases of the research lifecycle...through active phases and beyond

In the Spring of 2018, the Tri-Agencies released a draft of the “Tri-Agency Research Data Management Policy”

Release of a final policy is expected in Spring 2020

The draft policy includes suggested requirements related to three primary areas:

1. Institutional Strategy
2. DMPs
3. Deposit
Institutional Strategy:

“Each institution administering tri-agency funds is required to create an institutional research data management strategy”

Data Management Plans:

“...specific funding opportunities may require DMPs to be submitted to the appropriate agency at time of application

Data Deposit:

“Grant recipients are required to deposit into a recognized digital repository all digital research data, metadata and code that directly support the research conclusions in journal publications, pre-prints, and other research outputs that arise from agency-supported research
What is a data management plan (DMP)?

A DMP:

- Is a formal document which clearly articulates the strategies and tools you will implement to effectively manage your data.

- Speaks to the management of data both during the active phases of your research and after the completion of the research project.

The objective of a DMP is to address issues related to data management prior to starting your research project!
A DMP provides information across key research lifecycle categories:
Why are DMPs important?

A data management plan is important to the research process as it can help you to:

- **set out consistent strategies** prior to starting your research for how data will be managed throughout its entire lifecycle
- **identify the strengths & weaknesses** in your current practices and make decisions on how to integrate effective data management practices into your process
- **prepare data for future reuse, preservation and sharing**
- **reduce the overall cost of research** by increasing project efficiencies
When should I start creating a DMP?

A DMP should be developed as early in the research process as possible!

Every research project should **begin** with the creation of a DMP, as it will be used to guide the research process.
DMP General guidelines

Begin by providing a description of your research project, its focus, and purpose.

Avoid extensive use of discipline specific jargon - your DMP should be easily understood by anyone!

Provide clarification for any acronyms used.

Do not leave sections or questions blank.

Provide rationale for decisions made - help others understand why you have made a decision.

Your DMP is a living document - update it as needed!
DMP Sections: Data Collection

Include descriptions of how you will collect data, including from where and in what format(s).

Describe any software and/or platforms that will be used for data collection.

Provide an estimate of the amount of data you will collect (e.g., MBs/GBs/TBs).

Clearly explain how you will both store and transfer data.

Explain how you will organize your data, including details relating both to file naming and versioning.
DMP Sections: Documentation & Metadata

Choose a metadata standard suited to your discipline and/or chosen data repository, or provide rationale for creating your own.

Describe what information will be needed for others to understand or reuse your data.

Describe how you will consistently capture documentation throughout the project.
DMP Sections: Storage & Backup

Provide an estimate of storage space needed during the active phases of your research - remember to take into account file versioning, backups, and data growth!

State a data backup schedule, automatic being most ideal.

If needed, follow the 3-2-1 backup rule:
3 copies of your data, on 2 different storage media, with 1 located offsite.

Describe how collaborators or research team will be able to access, modify, contribute, and work with your data.
DMP Sections: Preservation

Consider the value of your data and decide which, if any, should be preserved.

Consider optimal file formats (preferably non-proprietary) for supporting long-term preservation.
DMP Sections: Sharing & Reuse

Consult with colleagues or librarians to choose an appropriate data repository or search re3data.org to find one.

Choose a repository that assigns permanent identifiers to datasets (e.g., DOI) to enhance discoverability, accessibility, and citability.

If applicable, describe how you will ensure file integrity, anonymization and de-identification.

Explain what uses can be made of your data through licenses like Creative Commons.

Consider the appropriate sharing of your data, including any funding or confidentiality requirements.
DMP Sections: Responsibilities & Resources

Identify data stewardship roles and responsibilities of project members and other organizations during and after the project.

Estimate and describe any required resources and costs for data management and long-term access to your data.
Describe if there are any legal, ethical, and intellectual property issues when managing and sharing your data.

Explain how you will comply with any applicable privacy legislation and laws, including funding and institutional requirements.

Describe how you will ensure your data are securely managed after the project is completed.
THE PORTAGE NETWORK is dedicated to the shared stewardship of research data in Canada through:

- Developing a national research data culture
- Fostering a community of practice for research data
- Building national research data services and infrastructure

Launched in 2015 by the Canadian Association of Research Libraries, Portage works within the library community to coordinate expertise, services, and technology in research data management, seeking to collaborate with other research data management stakeholders.

Research data culture represents widely shared values and principles for digital data management.

A community of practice for research data consists of stakeholders working collaboratively to ensure data are accessible to address complex research issues.

Research data management activities span sectors, domains, and jurisdictions, necessitating nationally coordinated research data management services and infrastructure.

> Read more
Portage Network

Portage is a national RDM network launched by CARL in 2015 which coalesces initiatives to build capacity and coordinate RDM activities in Canada

1. **Networks of Expertise:**
   - Pan-Canadian RDM expertise
   - Provides resources, tools, and experts in the area of RDM

1. **Infrastructure Platforms**
   - Works to assemble and deliver essential RDM infrastructure and service components

*Portage provides bi-lingual expertise and resources*
Based on a national network of experts

Expert Groups
- Data Management Planning
- Curation
- Data Discovery
- Preservation
- RDM Training
- Research Intelligence
- Data Repositories

125+ Experts

Working Groups
- Dataverse North
- Responsible Research Data Management Practices for Sensitive Data
- FRDR Service Model
- Institutional RDM Strategies

59 Organizations
Based on a national network of experts

Expert Groups

- Data Management Planning
- Curation
- Data Discovery
- Preservation
- RDM Training
- Research Intelligence
- Data Repositories

125+ Experts

59 Organizations

Four working groups:

1. Exemplars
2. Policy
3. Training
4. DMP Repository

Working Groups

- Dataverse North
- Responsible Research Data Management
- FRDR Service Model
- Institutional RDM Strategies
DMP Assistant Key Features

- Supports cross institutional research collaboration
- Freely available to end users
- Web-based, bilingual tool
- Endorsed by the Tri-Agencies
- Guidance & Examples
- Can have multiple DMPs
- Shared access - co-owner, editor, or read-only
- Exportable (e.g., pdf, text, MS Word)
- Living document - modifiable when needed

DMP Assistant

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.
**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, sign in with your credentials.

New to DMP Assistant? Sign up today.

Please note that we are currently working on single sign-in authentication. For now, please create a new DMP Assistant account. You will have the option to link your DMP Assistant account to your campus ID when that feature becomes available.
Upon signing in, researchers can either access an existing DMP.....

.....or create a new one
Each DMP has study level information associated with it.
Researchers can choose to answer questions within any given section at any time.

**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>University of Alberta Data Management Questions</th>
<th>Share</th>
<th>Export</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td>(6 questions, 6 answered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Documentation and Metadata</strong></td>
<td>(3 questions, 3 answered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Storage and Backup</strong></td>
<td>(2 questions, 2 answered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preservation</strong></td>
<td>(3 questions, 3 answered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Sharing and Reuse</strong></td>
<td>(4 questions, 4 answered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Responsibilities and Resources</strong></td>
<td>(4 questions, 3 answered)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**What types of data will you collect, create, acquire and/or record?**

**UAAlberta Guidance**

Examples: Images, audio, video, text, tabular data, modeling data, spatial data, instrumentation data

**What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?**

**UAAlberta Guidance**

Proprietary file formats which require specialized software or hardware to use are not recommended, but may be necessary for certain data collection or analysis methods. Using open file formats or industry-standard formats (e.g., those which are widely used by a given community) is preferred whenever possible.
From here you can download your plan in various formats. This may be useful if you need to submit your plan as part of a grant application.

Select what format you wish to use and click to 'Export'.

Format

pdf  Export
Short tutorial video
https://libcasts.library.dal.ca/Portage/DMP_Assistant/

Portage’s DMP Assistant

Creating a data management plan
Questions & Discussion

James Doiron
RDM Services Coordinator, UofA Libraries
Academic Director, UofA Research Data Centre
Co-Chair, Portage DMP Expert Group

jdoiron@ualberta.ca