GETTING STARTED WITH DATA MANAGEMENT PLANNING

May 2019
GOALS FOR TODAY

1. **RDM 101:**
   1. How to avoid a data management nightmare (a short video)
   2. RDM 101 modules (what is data management, RDM policy, steps for planning, etc)

2. How to create an RDM plan using the DMP assistant
PART 1. RDM 101

What is research data management and why should I do it?
A SCENARIO

How to avoid a data management nightmare

NYU Health Sciences Library
Karen Hanson  |  Kevin Read  |  Alisa Surkis
PORTAGE TRAINING
MODULES:

RDM 101

https://portagenetwork.ca/training-resources/portage-training-resources/rdm-101/

• **Module 1:** Background and Learning Objectives

• **Module 2:** What is Research Data Management (RDM?)

• **Module 3:** Canadian Policy Review (we will not cover this in its entirety today)

• **Module 4:** Steps Towards Good Research Data Management.
PART 2. USING THE DMP ASSISTANT

To make a research data management plan
WHERE TO START: DATA MANAGEMENT PLANS

• We are going to focus on the planning aspect of Data Management.
• This is an on-going, evaluative task

• Elements of Data Management Plan:
  1. Data Collection
  2. Documentation & Metadata (“data about data”)
  3. Storage & backup
  4. Preservation
  5. Sharing & Reuse
  6. Responsibilities & Resources
  7. Ethics & Legal compliance
The Portage Network:
https://portagenetwork.ca/  |  https://reseauportage.ca

Tool:
DMP Assistant (Assistant PGD in French): an all-purpose tool for preparing data management plans (DMPs).

- takes researchers step-by-step through a number of key questions about data management.
- Guidance and examples are provided.
- 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.
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.

**Sign up**

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.
Since I already have an account, we are brought directly to the “My Plans” section upon logging in. You can see some of the sample plans I’ve created. This section lists the name, template used, who owns the plan, if it has been shared, last edit date. You can also edit existing plans, share, export and delete them.

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th>Shared?</th>
<th>Last edited</th>
<th>Select an action</th>
</tr>
</thead>
<tbody>
<tr>
<td>My plan (Portage Template)</td>
<td>Me</td>
<td>No</td>
<td>17-02-2017</td>
<td>Edit Share Export Delete</td>
</tr>
<tr>
<td>Test plan for research data management needs assessment at Dalhousie University.</td>
<td>Me</td>
<td>No</td>
<td>24-04-2017</td>
<td>Edit Share Export Delete</td>
</tr>
<tr>
<td>My plan (University of British Columbia Generic Template)</td>
<td>Me</td>
<td>No</td>
<td>24-04-2017</td>
<td>Edit Share Export Delete</td>
</tr>
<tr>
<td>My plan (Portage Template)</td>
<td>Me</td>
<td>No</td>
<td>24-05-2017</td>
<td>Edit Share Export Delete</td>
</tr>
</tbody>
</table>
I’m going to choose Dalhousie’s branded version with the standard portage template. Institutions can customize template questions and add guidance. Researchers need to only answer questions that are applicable to them.
These fields can all be edited - you probably don’t want to call your plan “My plan”. You can also save each section as you go, and come back later.
These are the questions that you can fill out for your DMP. You only need to answer questions that are relevant to your project. Guidance is provided, and your institution can also add additional guidance.

<table>
<thead>
<tr>
<th>Sections</th>
<th>Questions</th>
</tr>
</thead>
</table>
| Data Collection        | - What types of data will you collect, create, link to, acquire and/or record?  
- 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?  
- What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized? |
| Documentation and Metadata | - What documentation will be needed for the data to be read and interpreted correctly in the future?  
- How will you make sure that documentation is created or captured consistently throughout your project?  
- If you are using a metadata standard and/or tools to document and describe your data, please list here. |
| Storage and Backup     | - What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it?  
- How and where will your data be stored and backed up during your research project?  
- How will the research team and other collaborators access, modify, and contribute data throughout the project? |
| Preservation           | - Where will you deposit your data for long-term preservation and access at the end of your research project?  
- Indicate how you will ensure your data is preservation ready. Consider preservation-friendly file formats, ensuring file integrity, anonymization and de-identification, inclusion of supporting documentation. |
| Sharing and Reuse      | - What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).  
- Have you considered what type of end-user license to include with your data?  
- What steps will be taken to help the research community know that your data exists? |
| Responsibilities and Resources | - Identify who will be responsible for managing this project’s data during and after the project and the major data management tasks for which they will be responsible.  
- How will responsibilities for managing data activities be handled if substantive changes happen in the personnel overseeing the project’s data, including a change of Principal Investigator?  
- What resources will you require to implement your data management plan? What do you estimate the overall cost for data management to be? |
| Ethics and Legal Compliance | - If your research project includes sensitive data, how will you ensure that it is securely managed and accessible only to approved members of the project?  
- If applicable, what strategies will you undertake to address secondary uses of sensitive data?  
- How will you manage legal, ethical, and intellectual property issues? |
Once you fill out the basic plan details, you are brought to the Question section. Guidance to the questions is provided on the right. You can also add links, tables and format the text. On the right, you also have the option to provide notes for any collaborators. You can also save each question as you answer it and return to the plan later.
You can add collaborators and add permissions. There are three permission levels. Collaborators will have to register for the DMP Assistant if they don’t already have an account.
Test plan for research data management needs assessment at Dalhousie University.

This page gives you an overview of your plan. It tells what your plan is based on and gives an overview of the questions that you will be asked.

Plan name: Test plan for research data management needs assessment at Dalhousie University.
ID: e437393
Grant number: n/a
Principal Investigator/Researcher: Erin MacPherson
Principal Investigator/Researcher ID: n/a
Plan data contact: -
Description: This project is to assess the data management needs of the research community at Dalhousie University. The data will identify the volume and type of data collected by researchers. The results of this study will help us to determine the types of data management services that the Dalhousie Libraries can offer researchers.

This plan is based on:
Institution: Dalhousie University

If you need to make changes, you can always go back and edit plan details or the questions anytime.
Click on the “Settings” option for more great features before you export your document. You can select which sections to export (including plan details and questions). For some formats, you can also choose your font type, size and margins.
RDM Needs Assessment Data Management Plan

Project Name  RDM needs assessment at Dalhousie University (test plan).
Project Identifier  FundID: testRDM010101
Grant Title  n/a
Principal Investigator / Researcher  Erin MacPherson
Project Data Contact  Erin.MacPherson@dal.ca, 902-893-3296
Description  This project is to assess the data management needs of the research community at Dalhousie University. The data will identify the volume and type of data collected by researchers. The results of this study will help us to determine the types of data management services that the Dalhousie Libraries can offer researchers.
Institution  Dalhousie University

Data Collection
What types of data will you collect, create, link to, acquire and/or record?
There will be lots of 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?
Data will be in many different formats!

What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?
My lab has created a file naming and version control system that I have attached to this document.
DMP ASSISTANT EXERCISE

- Use the case study provided (Civil Engineering).
- 30 minutes to complete exercise.

**Instructions:**

1. Create an account with the DMP Assistant: [https://portagenetwork.ca/](https://portagenetwork.ca/). Login.
2. Erin or Sarah will create the basic plan and “share it” with you.
3. Work in pairs. We will assign sections to each pair or group. Work through the DMP Assistant questions.
4. Share what you came up with, what you found difficult.
CASE STUDY SUMMARY

Civil Engineering professor studying traffic patterns:

- Generates “profiles” of movement through intersections
- Goal is to improve traffic flow
- Works with the State Department of Transportation who use this information to make decisions on spending and inform public policy.

Data Collection:

- Raw data is collected using a form that is proprietary to the sensor
- Raw data exported to .csv files, added to MySQL databases (multiple) holding 3-4 months of data. Extracted to MS Excel
- Other components include video, database of road conditions, weather conditions, etc.

Data handling

- Several graduate students working in the lab
- No standard naming convention or metadata/explanations included
- Field names are unclear

Data storage

- Professor sends copies to department in zipped format over FTP servers
- Would like to share the data if allowed
- Stored on lab’s centralized server, backed up with whatever campus IT uses.
- Would also like to archive the data long-term.

Reference:
QUESTIONS

• What do you like about the DMP Assistant?
• What would make the DMP Assistant more useful? Is anything missing?
• What could a librarian (a liaison librarian or RDM librarian) do to help a researcher who may be in a situation like this? What type of training or events might be helpful?
• How would you like to see this tool promoted in your department/faculty?
Any Questions?