

Data Management Planning with DMP Assistant

MacDATA Workshop: February 26, 2016

V. Jadon, J. Brodeur
McMaster University Library



**MAPS
DATA
GIS**

Overview

1. Introduction to Research Data Management (RDM) [20 mins]
 - a. An RDM primer - The Canadian context
 - b. RDM activities
 - c. Introduction to data management planning
2. Exercise: Planning for research data management [40 mins]
3. Discussion and wrap-up [20 mins]

Workshop Objectives

By the end of this workshop, attendees should be able to:

1. Describe research data management (RDM) and summarize basic RDM activities,
2. Create a data management plan (DMP) for a current or future research project; and,
3. Identify areas where RDM support is needed, and locate appropriate resources.

If you shared your data another researcher or collaborator, would they be able to:

- a. Interpret and understand it?
- b. Use it in new analyses?

Would someone (including you) be able to find, interpret and use your data 20 years from now?

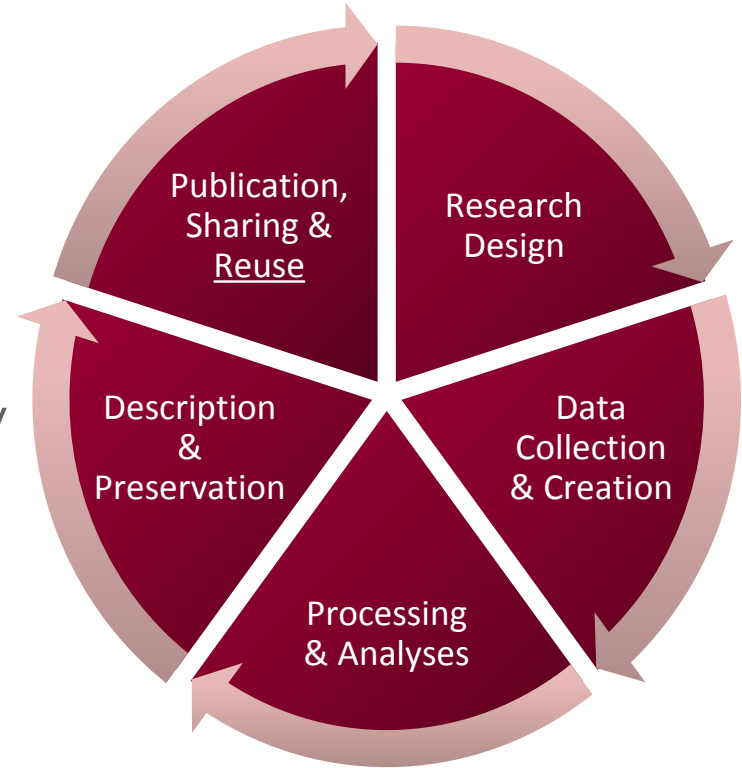
Research Data Management Primer: RDM in the Canadian Context

Research Data Management is...

... the **active organization** & maintenance of data

... the **application of best practices** to ensure data security, accessibility, usability, and integrity

... a **set of activities** resulting in self-describing data sets that can be discovered and reused.



Applying RDM best practices will benefit...

Researchers and their collaborators

- ✧ Improves research efficiency and productivity
- ✧ Provides extra credit for research work
- ✧ Increases research impact
- ✧ (May) help to meet funding requirements

Research Communities

- ✧ Accelerates discovery
- ✧ Enables validation and verification

Funders, governments and the public

- ✧ Improves return on investment
- ✧ Increases research transparency
- ✧ Data as a public good

Canadian Government & Funding Agencies

- ✧ Canada's Action Plan on Open Government (2014-2016)
- ✧ Tri-Agency Open Access Policy on Publications (Feb-2015)
- ✧ Tri-Agency Statement of Principles on Digital Data Management (Jul-2015)
- ✧ Comprehensive Brief on Research Data Management Policies (Aug-2015)

Tri-Agency Statement of Principles on Digital Data Management

Expectations:

- **Data management planning**
 - Adherence to standards
 - Collection and storage
 - Metadata
 - Preservation, Retention and Sharing
 - Timeliness
 - Efficient and cost-effective

Publishers & Scientific Organizations

- ✧ Data sharing policies
- ✧ Recommended data repositories
- ✧ Publisher-supported data repositories



Support Organizations and Communities

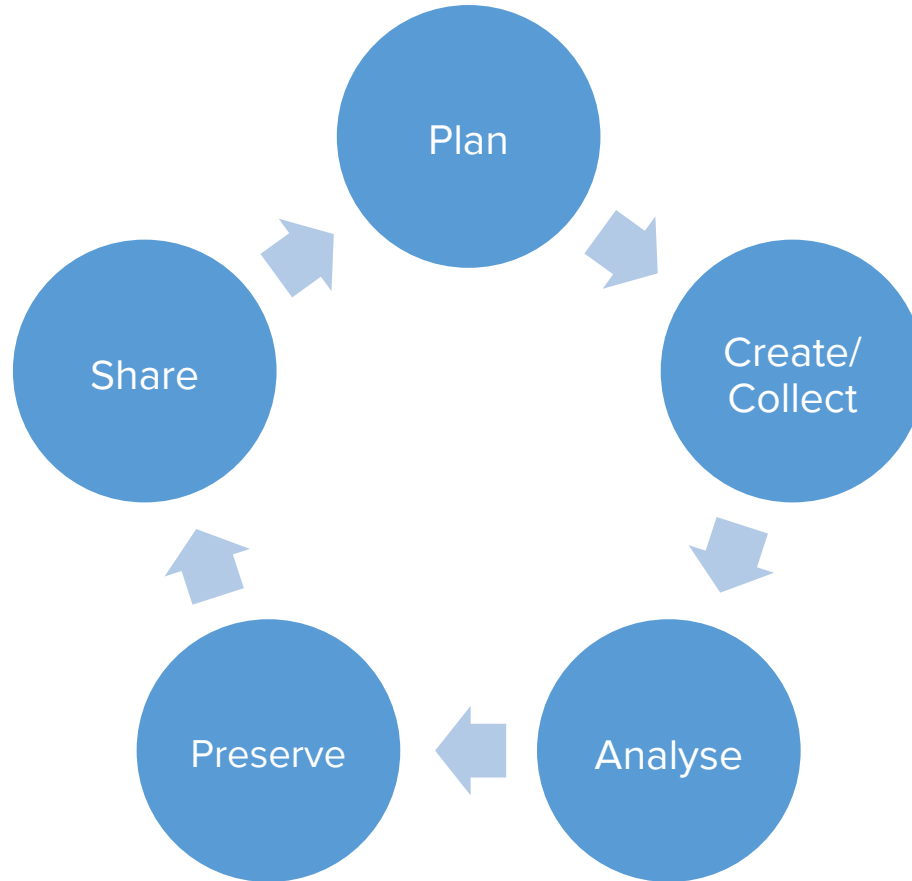
- ✧ Standards of practice
- ✧ Training, expertise and information
- ✧ Tools and resources



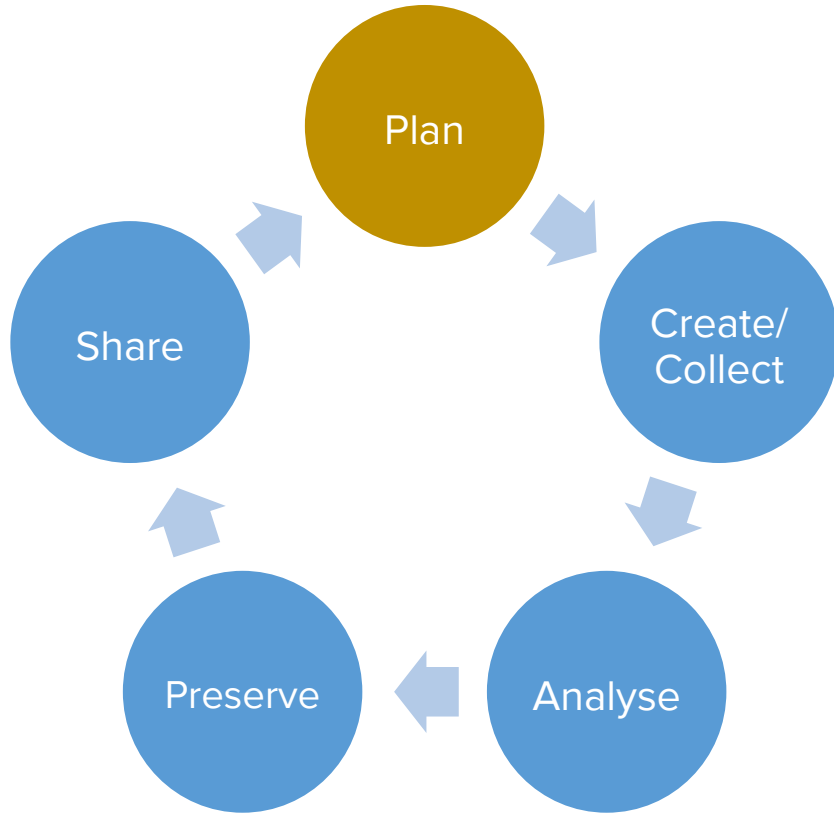
RDM
@McMaster

RDM in the Data Life Cycle: Common Challenges

RDM in the Data Lifecycle



Planning

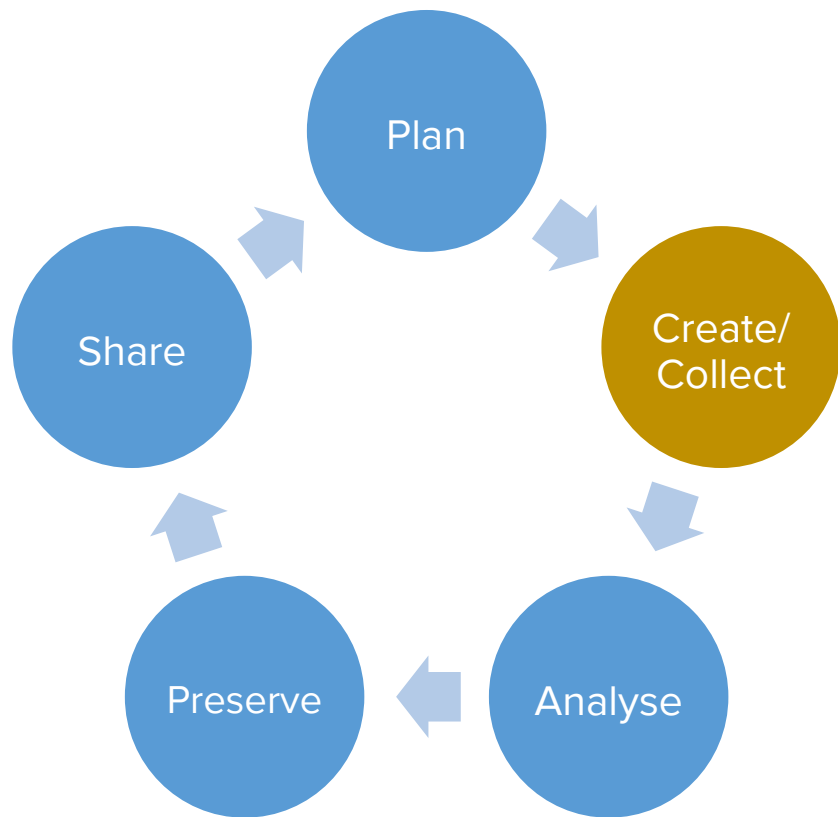


What are the stipulations in institutional, funder or publisher data policies to be followed?

What resources do you require to manage your data?

Who is responsible for data management and long-term stewardship?

Creating and Collecting

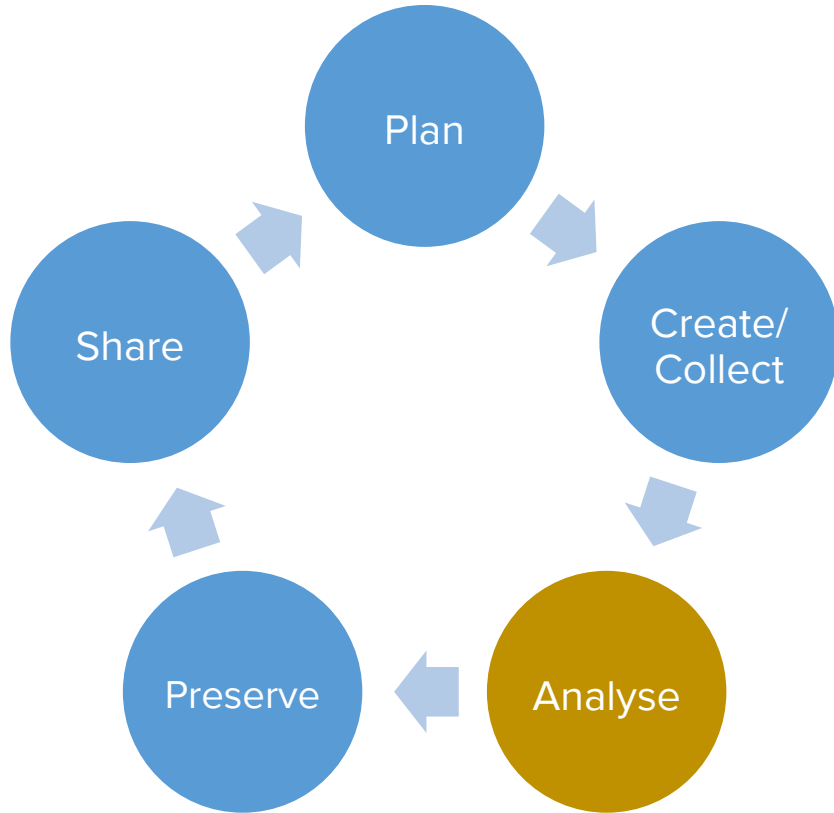


How much data will you collect or create?
Where will you store it and back it up?

Is your created or collected data in a
suitable format for sharing and long-term
preservation?

What documentation and metadata
should accompany the data?

Analyzing

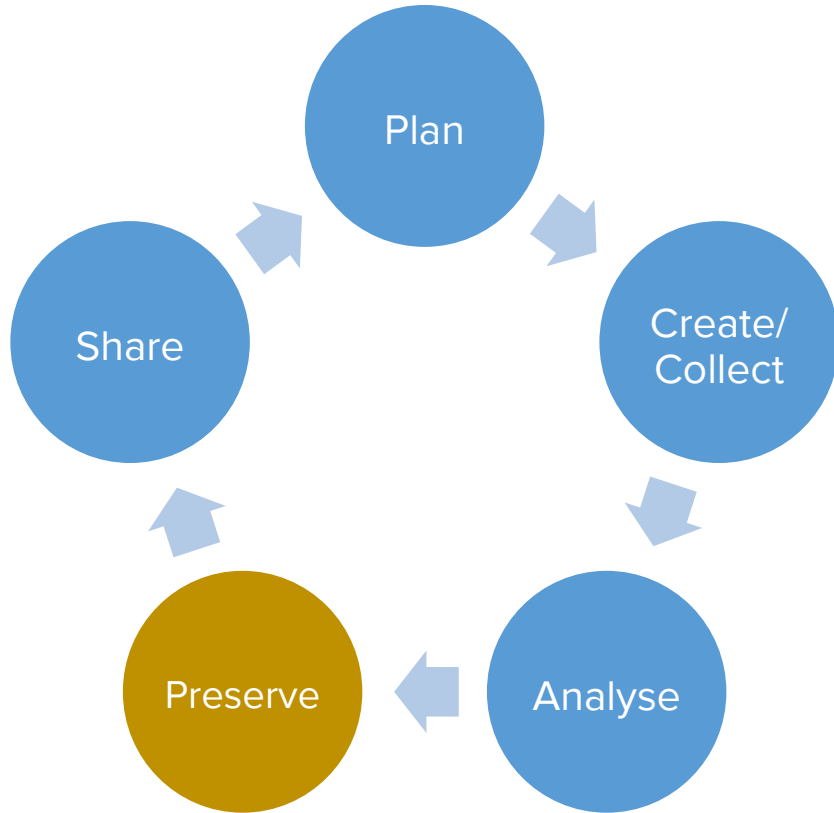


How will you manage any ethical or privacy issues before analysing the data?

How will you securely store (potentially large and cleaned) data pre- and post-analysis?

Who will have access to this data for analysis?

Preserving

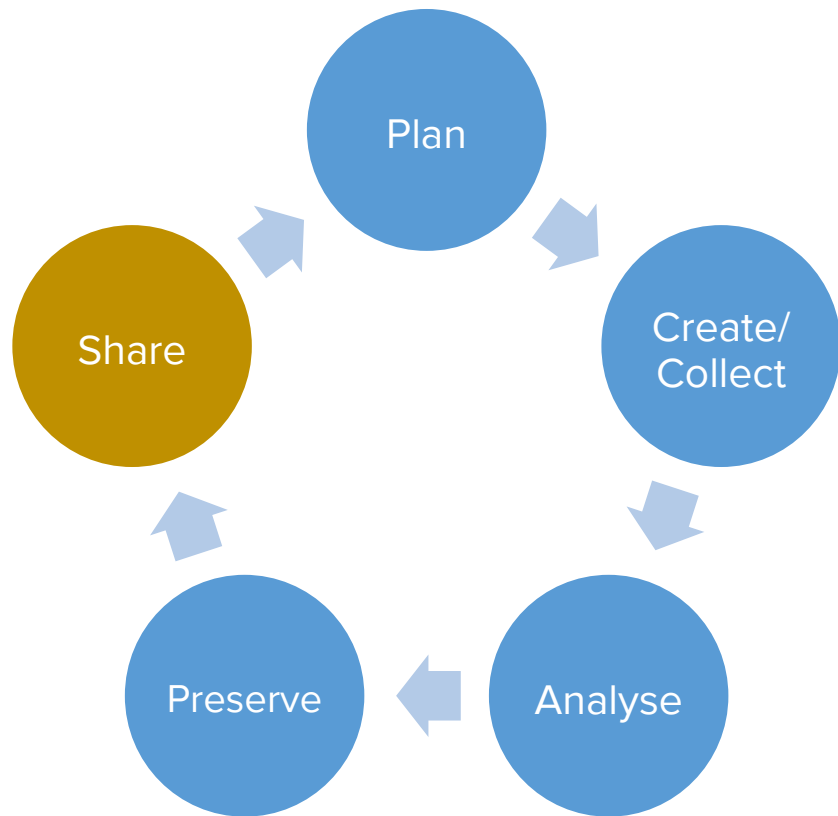


What data should be retained and preserved?

Where will you preserve your data? Who will have access to this preserved data?

For how long you are going to preserve your research data?

Sharing



What data will you make publicly available for research and reuse?

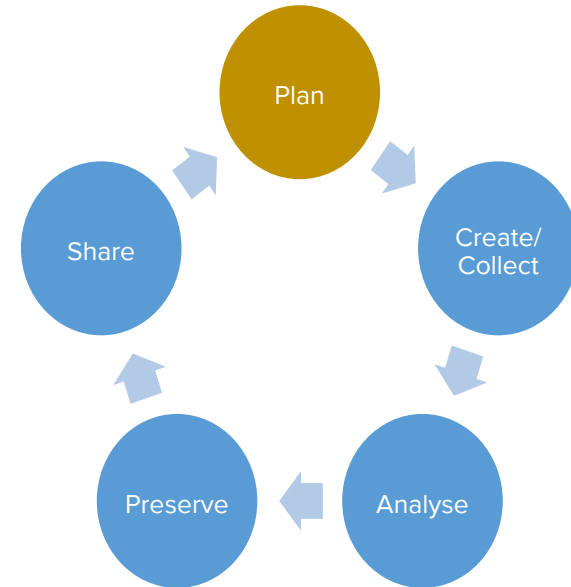
What resources are in place for sharing your data among multiple sites?

How will you manage restrictions e.g. license, privacy issues etc. associated with your data before sharing?

Research Data Management Planning

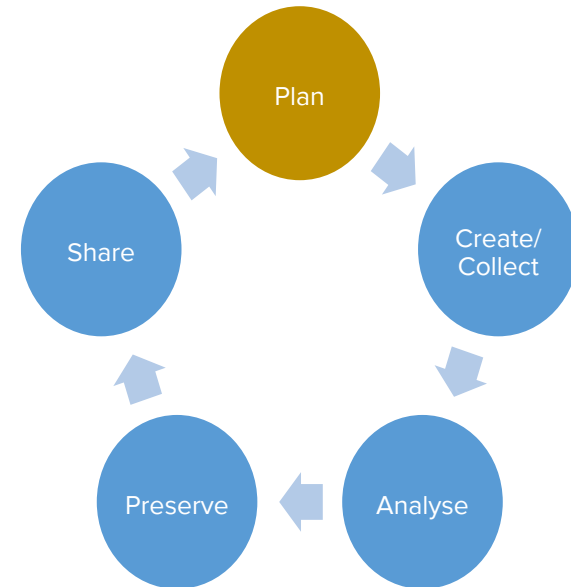
A Research Data Management Plan (DMP) *should:*

- ✧ Describe how you will manage data through all stages of your research
- ✧ Communicate a strategy for creating share-worthy and share-ready data products



An effective DMP *will:*

- ✧ Be completed at the time of study design
- ✧ Ensure compliance with policies / obligations
- ✧ Document and organize research activities
- ✧ Help identify support requirements
- ✧ *(Likely)* evolve with your study...



DMPs: The International Context

UK:

- ✧ AHRC, BBSRC, CRUK, ESRC, MRC, NERC, STFC, Wellcome Trust

US:

- ✧ NIH*, NSF,

EU

- ✧ Portugal, EC-Horizon 2020

Shearer, Kathleen. 2015. Comprehensive Brief on Research Data Management Policies. Science.gc.ca, April 2015. Available at: <http://www.science.gc.ca/default.asp?lang=En&n=1E116DB8-1>

DMPs: The Canadian Context

From the *Statement of Principles* (2015):

“Data management planning is necessary at all stages of the research project lifecycle, from design and inception to completion.”



SSHRC  CRSH



DMP ASSISTANT PGD

DATA MANAGEMENT PLANS
PLANS DE GESTION DES DONNÉES

A web-based, *bilingual* data management planning tool made available to all researchers in Canada through the Portage network

A guide for best practices in **data stewardship**:

Available at: <https://portagenetwork.ca> | <https://reseauportage.ca>

Activity - Develop a DMP (~40 minutes)

Using the **CARL Portage DMP Assistant**:



- ✧ Create an account
- ✧ Develop a DMP for a current, upcoming or prospective research project
- ✧ Browse through the DMP questions
- ✧ Complete the answerable questions in rough form (or internally)
- ✧ Note any confusing, problematic or difficult questions

Navigate to the DMP Assistant:

<https://portagenetwork.ca/>

Sign in with your email / password



Your account was successfully confirmed. Please sign in.

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 in

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

[Forgot your password?](#)

☐ Remember me

Campus ID single sign-in authentication coming soon!

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

Click “Create plan”



Set organization to “McMaster University”

Select the “Portage” template



Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren't responding to specific requirements from a funder or an institution, you can choose the **Portage Data Stewardship Template**. The Portage Data Stewardship Template is based on internationally accepted standards and best practices. It has been prepared and is maintained by a group of research data management experts from research libraries across Canada.

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, default Portage DMP template will be used. [Not applicable/not listed.](#)

Choose a template

There are a number of possible templates you could use. Please choose one.

Tick to select any other sources of guidance you wish to see.

☒ McMaster

<https://portagenetwork.ca/>

Enter relevant personal,
organizational and study information



Expand boxes to complete questions
for each section



Provide answers in the left-hand box
Use right-hand box for guidance/comments



The image displays three sequential screenshots of the Portage Template web application, illustrating the workflow from initial setup to detailed data collection questions.

Top Screenshot: My plan (Portage Template)
This view shows the initial setup page. It includes tabs for "Plan details", "Portage Data Management Questions", "Share", and "Export". A "Cancel" button is in the top right. The main area prompts the user to "Please fill in the basic project details below and click 'Save' to save". Fields include Plan name (pre-filled with "My plan (Portage Template)"), ID, Grant number, Principal Investigator/Researcher (pre-filled with "RDM @McMaster"), Principal Investigator/Researcher ID, Plan data contact, and a Description text area. "Save" and "Cancel" buttons are at the bottom right.

Middle Screenshot: My plan (Portage Template)
This view shows the "Portage Data Management Questions" section. A "Tips" box at the top states: "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." Below this is a list of sections with expandable arrows: "Data Collection (3 questions, 0 answered)", "Documentation and Metadata (1 question, 0 answered)", "Storage and Backup (1 question, 0 answered)", "Preservation (2 questions, 0 answered)", "Sharing and Reuse (1 question, 0 answered)", and "Responsibilities and Resources (3 questions, 0 answered)".

Bottom Screenshot: Detailed Data Collection Question
This view shows a specific question under the "Data Collection" section. The question is: "What types of data will you collect, create, link to, acquire and/or record?". Below the question is a rich text editor with a toolbar (bold, italic, link, etc.). To the right of the editor is a "Guidance" box titled "Portage Guidance" which provides examples: "Examples: numeric, images, audio, video, text, tabular data, modeling data, spatial data, instrumentation data." A "Save" button is at the bottom left.

Reflections and Discussion

What are your reflections on:

- ✧ The overall experience?
- ✧ The interface and its content?

What sections / questions were particularly challenging? Howso?

What kind of support or resources would help you complete this plan?

What comes next?

- ✧ Institutional customization of DMP Assistant
- ✧ Amalgamation of DMP tools DMPOnline/Assistant & DMPTool
- ✧ Integration with REBs?
- ✧ Funding agency guidelines / policies / mandates

DMP guidance

Canada

- ✧ Updated Portage Network website (*Coming soon*): <https://portagenetwork.ca>
- ✧ RDM@McMaster: <http://library.mcmaster.ca/rdm/planning/dmp>
- ✧ UBC Libraries: <http://researchdata.library.ubc.ca/plan/>

International

- ✧ CDL **DMP**Tool [US]: https://dmptool.org/community_resources
- ✧ DCC [UK]: <http://www.dcc.ac.uk/resources/data-management-plans>

DMP support



Portage Network

Coming soon

<https://portagenetwork.ca>

RDM
@McMaster

RDM@McMaster

<http://library.mcmaster.ca/rdm>

rdmgmt@mcmaster.ca

Thank You.

RDM
@McMaster

Check out <http://library.mcmaster.ca/rdm> for more information

Contact us at rdmgmt@mcmaster.ca