





Support your research with data management planning!

James Doiron RDM Services Coordinator, University of Alberta Library Academic Director, UofA Research Data Centre Co-chair, Portage DMP Expert Group

March 23, 2021

What is Research Data Management?

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



Tri-Agency RDM Policy

In March 2021, the Tri-Agency released their <u>Research Data Management Policy</u>

The policy includes general requirements and implementation dates across three main pillars:



Tri-Agency RDM Policy - "Take away points"

Institutional Strategy:

Each postsecondary institution eligible to administer Tri-Agency funds is required to create and release an institutional RDM strategy by March 2023

Data Management Plans:

For certain funding opportunities, data management plans (DMPs) will be required to be submitted at the time of application, as outlined in the call for proposals; in these cases, the DMPs will be considered in the adjudication process.

NOTE: <u>By spring 2022,</u> the agencies will identify the initial set of funding opportunities subject to the DMP requirement. The agencies will pilot the DMP requirement in targeted funding opportunities before this date.

Data Deposit:

Grant recipients are required to deposit into a digital repository all digital research data, metadata and code that directly support the research conclusions in journal publications and pre-prints that arise from agency-supported research.

NOTE: After reviewing the institutional strategies and in line with the readiness of the Canadian research community, the agencies will phase in the deposit requirement.

FAIR https://www.go-fair.org/fair-principles/

FAIR is a set of guiding principles focused towards making data:

*Key Reading:

- Findable Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.
- Accessible Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.
- **Interoperable** Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- Reusable Data and collections have clear usage licenses and provide accurate information on provenance.



Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*3:160018 doi: 10.1038/ sdata.2016.18 (2016).

FAIR Principles: Key Resource

go-fair.org (https://www.go-fair.org/fair-principles/)

- Detailed information across the FAIR principles
- Implementation Networks
- News
- Event
- Resources! —



Home > Resources

F/IR

> GO FAIR Materials

- > Materials for INs
- > Materials for Countries
- Materials from GO FAIR meetings
- Media & Communications
 Material
- GO FAIR Workshop Series
 Metadata for Machines
 Workshops
 - Germany GOes FAIR
 Workshops
 - > Pillar-Specific Workshops
 - > Manifesto Writing Worksho
- > FAQ
- > RDM Starter Kit
 > More on FAIR
- > Glossary

This page is dedicated to resources that you might find useful in your FAIR endeavours. We have collected our GO FAIR materials as well as interesting papers & publications, tools and more for your information and use:

FAIR Principles Implementation Networks News Events Resources About GO FAIR Q

GO FAIR Materials

Resources

- GO FAIR Workshop Series
- FAQ
- Starter Kit for Research Data Management
- More on FAIR
- Glossary

2

OCAP

The First Nation Principles of OCAP are a set of standards that establish how First Nations data should be collected, protected, used or shared:

- Ownership: refers to the relationship of First Nations to their cultural knowledge, data & information - a community/group collectively owns information in the same way that an individual owns his/her personal information
- Control: affirms that First Nations communities have rights in seeking control over all aspects of research
 from start to finish - that impact them. This extends to control of resources and review processes and management of information.
- Access: First Nations must have access to information and data about themselves and their communities regardless of where it is held, and have the right to manage and make decisions regarding access to their collective information.
- **Possession:** Refers to the physical control of data the mechanism by which ownership can be asserted and protected.



OCAP Principles: Key Resource

First Nations Information Governance Centre (https://www.FNIGC.ca/)

- Fundamentals of OCAP online training program
- **FNIGC** data online
- First Nations Data Centre (data by request)
- First Nations surveys (i.e., regional • health, early childhood, education, labour, oral health)
- FNIGC online library





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:

- <u>set out consistent strategies</u> 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
- <u>reduce the overall cost of research</u> 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

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

Clearly explain how you will both store and transfer data Provide an estimate of the amount of data you will collect (e.g., MBs/GBs/TBs).

Explain how you will organize your data, including details relating both to file naming and versioning

DMP Sections: Documentation & Metadata

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

Describe how you will consistently capture documentation throughout the project.

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

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!

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

State a data backup schedule, automatic being most ideal. Describe how collaborators or research team will be able to access, modify, contribute, and work with your data.

DMP Sections: Preservation

Not all data that you create necessarily *needs* to be preserved consider such things as the value of your data, funding requirements, etc., and decide which, if any, should be preserved. Consult with experts in the Libraries as needed!

> Consider optimal file formats (preferably nonproprietary) for supporting long-term preservation. Optimally preserved data are easily accessible and use by anyone, without requiring proprietary software to do so.

DMP Sections: Sharing & Reuse

Consider the appropriate sharing of your data, including any funding or confidentiality requirements. Explain what uses can be made of your data through licenses like <u>Creative Commons</u>.

Consult with colleagues or librarians to choose an appropriate data repository or search <u>re3data.org</u> 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 deidentification.

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.

DMP Sections: Ethics & Legal Compliance

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 including the secure management of sensitive data and in accordance with any ethical obligations

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. <u>Networks of Expertise</u>:
- 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 & Metadata
- Preservation
- RDM Training
- Research Intelligence
- Data Repositories
- Sensitive Data
- Dataverse North WG





portagenetwork.ca









DMP guidance resources

Creating an effective DMP (English/French)



CREATE AN EFFECTIVE DATA MANAGEMENT PLAN

This brief guide presents a general framework for creating an effective data management plan (DMP) to help you plan and organize your research and to meet research funder requirements.

To prepare your DMP, visit the Portage DMP Assistant tool.

GENERAL GUIDELINES

- Begin by providing a description of your research project, its focus, and purpose.
- Avoid the 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!

DATA COLLECTION

- Include descriptions of how you will collect data, including from where and in what format(s).
- · Provide an estimate of the amount of data you will collect (e.g., MBs/GBs/TBs).
- Describe any software and/or platforms that will be used for data collection.
- 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.

DOCUMENTATION AND METADATA

- Describe what information will be needed for others to understand or reuse your data.
- Describe how you will consistently capture documentation throughout the project.
- Choose a metadata standard suited to your discipline and/or chosen data repository or provide rationale for creating your own.



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) Portage Network | portage@carl-abrc.ca | portagenetwork.ca



STORAGE AND 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.
- If needed, follow the 3-2-1 backup rule: 3 copies of your data, on 2 different storage media, with 1 located offsite.
- State a data backup schedule, automatic being most ideal.
- Describe how collaborators or research team will be able to access, modify, contribute, and work with your data.

PRESERVATION

- Not all data that you create necessarily needs to be preserved consider such things as the value of your data, funding requirements, etc., and decide which, if any, should be preserved. Consult with experts in your Library as needed.
- Consider optimal file formats (preferably non-proprietary) for supporting long-term preservation.

SHARING AND REUSE

- Consider the appropriate sharing of your data, including any funding, ethical and/or confidentiality requirements.
- Explain what uses can be made of your data through licenses like Creative Commons.
 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 deidentification.

RESPONSIBILITIES AND 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 longterm access to your data.



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) Portage Network | portage@carl-abrc.ca | portagenetwork.ca

DMP exemplars

Discipline and/or methodology focus	Title & Link					
	Belgians and French in the Prairies					
Digital Humanities	Soundscape Study					
	Women's Print History Project (1750-1830)					
	Historical Canadian Census Data					
Mixed Methods (surveys & qualitative interviews)	Mixed Methods Fictional Exemplar					
	Ecohydrology Research Group					
Natural Sciences	Computational Reproducibility in High-Performance Computing					
Social Sciences	People, Places, Policies & Prospects: Affordable Rental Housing for Those in Greatest Need					
	Usage of Academic Profile Websites					

DMP exemplars



Data Management Plan Exemplar #3: Mixed Methods

Fictional Exemplar

Data Collection

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

We will be collecting surveys which will then be exported into tabular format.

We will also be conducting both semi-structured interviews and focus groups that will produce both digital audio and text (transcriptions) based 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?

Our file formats will exist both in non-proprietary and proprietary formats. The non- proprietary formats will ensure that these data are able to be used by anyone wishing to do so once they are deposited and made openly available.

Surveys will exist in .csv (non-proprietary), MS Excel, & SPSS (both proprietary) formats. For more information regarding SPSS see: SPSS Wikipedia https://en.wikipedia.org/wiki/SPSS

Interviews & focus groups data will exist in .mp3 (non-proprietary), MS Word & NVivo (both proprietary) formats. For more information regarding NVivo see: NVivo Wikipedia https://en.wikipedia.org/wiki/NVivo

Any survey data deposited for sharing and long-term access will be in .csv format so that anyone can use them without requiring proprietary software.

The final de-identified versions of the interviews and focus groups transcripts will be exported into a basic non-proprietary text format for deposit, long-term preservation and access.

If data are collected using laptops or mobile devices, please explain how you will securely store and transfer the data.

Laptops are not being used for any data collection, though encrypted digital voice recorders (DVRs) will be used to collect both interviews and transcripts. Interviews and focus group digital audio files will not be stored on the DVRs, only collected and then securely transferred to the project's cloud based virtual research environment space via a secure FTP (File Transfer Protocol).



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) Portage Network | portage@carl-abrc.ca | portagenetwork.ca

Discipline and/or methodology

Digital Humanities

Mixed Methods

(surveys & qualitative interview)

Natural Sciences

Social Sciences

2

MODÈLES DE PGD

Type de recherche	Titre					
Sciences humaines numériq	<u>Plan de gestion des données pour le projet</u> <u>Belges et Français dans les Prairies</u>					
	Étude du paysage sonore du silence (Modèle)					
	Sciences Humaines Numériques					
	Sciences humaines numériques et données secondaires					
Méthodes mixtes	Modèle fictif sur « Méthodes mixtes »					

MODÈLES DE PGD

Type de recherche	Titre
Sciences naturelles	Plan de gestion des données du groupe de recherche en écohydrologie (Modèle)
	<u>Plan de gestion des données pour</u> <u>Reproductibilité computationnelle dans</u> <u>l'informatique haute performance (Modèle)</u>
Sciences sociales	Plan de gestion de données pour Personnes, places, politiques et perspectives (Modèle)
	Plan de gestion des données pour Utilisation des sites web de profils universitaires (Modèle)

DMP Assistant Key Features



DMP Assistant



Home Public DMPs DMP Templates Help

🛃 Language 🗸

✓ Notice: Signed out successfully.

Welcome to DMP Assistant.

DMP Assistant has been developed by the Portage Network 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 Assistant equivalent in France
- France funder requirements for Data Management Plans

Sign in	Create account
* Email	
* Passw	ord
Forgot pa	assword? nber email
Sign in	



Upon signing in, you can either access an existing DMP.....

(ASSISTANT) My Dash

LIBRARY

UNIVERSITY OF ALBERTA

My Dashboard Create plats 🖉 Reference - Help

Notice: Signed in successfully.

My Dashboard

The table below lists the plans that you have created, and that have been shared with you by others. You can edit, share, download, make a copy, or remove these plans at any time.

Q	Search								
Project Title 🗧 🗧	Template	¢	Edited	•	Role	Test	Visibility	Shared	
Portage DMP Exemplar - "Mixed Methods"	University of Alberta Template		03-03- 2021		Owner	0	Private	Yes	Actions-
Test	Portage Template		03-03- 2021		Owner	0	Private	No	Actions-
Test for migration to 2.0	University of Alberta Template		02-03- 2021		Owner	0	Private	No	Actions-
My plan (Portage Template)	Portage Template		05-11- 2020		Owner	0	Private	No	Actions-
My plan (University of Alberta Template)	University of Alberta Template		01-10- 2019		Owner	0	Private	No	Actions-
My plan (Portage Template)	Portage Template		13-09- 2019		Owner	0	Private	No	Actions-
My plan (Portage Template)	Portage Template		16-04- 2019		Owner	0	Private	No	Actions-
My plan (University of Alberta Template)	University of Alberta Template		29-08- 2018		Owner	0	Private	No	Actions-
My plan (University of Alberta Template)	University of Alberta Template		17-11- 2017		Owner	0	Private	No	Actions-
My plan (Portage Template)	Portage Template		28-09- 2017		Owner	0	Private	No	Actions-

View all

Create plan

1 2 Next Last



Each DMP has study level information associated with it



LIBRARY

-

SEEE

My Dashboard Create plans 🗐 R

UNIVERSITY OF ALBERTA

🖉 Reference 🗸 🛛 Help

Portage DMP Exemplar - "Mixed Methods'



ID

8675309

Principal Investigator

You can choose to answer questions within any given section at any time...

My Dashboard Create p	plans 🗾 Reference 🗸 Help	
UNIVERSITY OF ALBERT	ΓΑ	
University of All	Portage DMP Exemplar - "Mixed Methods"	
	Project Details Plan overview Write Plan Share Download	
	expand all collapse all 22/22 answered	
	Data Collection (6 / 6)	+
	Documentation and Metadata (3 / 3)	+
	Storage and Backup (2 / 2)	+
	Preservation (3 / 3)	+
	Data Sharing and Reuse (4 / 4)	+
	Responsibilities and Resources (4 / 4)	+

Dedicated space as well as guidance customized guidance is provided for each question...

Documentation and Metadata (3 / 3)

What documentation will be needed for the data to be read and interpreted correctly in the future? This includes study-level documentation, data-level description, and any other contextual information required to make the data usable by other researchers.

B I \∃ - \∃ - ∅ \⊞ -

Survey data will be collected within <u>REDCap</u>, a secure electronic data capture and management software hosted by the Women & Children's Health Research Institute (WCHRI) at the University of Alberta. REDCap features include the ability to develop and export a data dictionary which will outline all codes and variables within the survey. Key documentation related variables will be automatically populated within the survey data including time and date stamps and other key information to support data management and analytic activities.

All qualitative interviews will include summary information including: data collector, location of interview, and the date that the interview was conducted. Additionally, qualitative interviews will have accompanying field notes contaning key contextual information and metadata.

File naming documentation will be developed and implemented. Components of file names will include as needed: file version (raw, edit, master, analytic), date (ie., dd/mm/yyyy), and any applicable key contextual information, (e.g., geographical location, interviewer initials or code).

Guidance		Comments			
UAlberta					
Your docun level inforn		may include study- out:			
 when any r condi contection collection 	it was cre elevant st tions of us extual deta tion meth mentation	udy documents			
A complete may includ	-	on of the data files			
 naming and labelling conventions explanations of codes and variables 					

• any information or files require to reproduce derived data.

More information about data documentation is available at the UK Data Archive

Save

For any given question you are able to have 'share notes' with your collaborators

Portage DMP Exemplar - "Mixed Methods"

Project Details	Plan overview	Write Plan	Share	Download								
expand all colla	pse all				22/22 answered							
Data Collectio	n (6 / 6)											_
												_
	f data will you co - ≟ - ⊘ ⊞ -		acquire a	ind/or recor	d?		Gu	iidance		Comr	ments (1)	_
We will be collecting surveys which will then be exported into tabular format. We will also be conducting both semi-structured interviews and focus groups that will produce both digital audio and text (transcriptions) based data.					surv form mite	nat like a	e expor .csv file n@hotma s to sha	ted int ?? ail.com Edit are wi	o tabular <mark>2 years ago</mark> Remove th			
											Å	

You can easily add collaborators to your DMP, with different permissions settings...

(ASSISTANT) UNIVERSITY OF ALBERTA LIBRARY Portage DMP Exemplar - "Mixed Methods" **Project Details** share Download Plan overview Write Plan Set plan visibility Public or organisational visibility is intended for finished plans. You must answer at least 0% of the questions to enable these options. Note: test plans are set to private visibility by default. Private: visible to me, specified collaborators and administrators at my organisation Organisation: anyone at my organisation can view Public: anyone can view Manage collaborators Invite specific people to read, edit, or administer your plan. Invitees will receive an email notification that they have access to this plan. Email address Permissions mitchelldoiron@hotmail.com Owner jdoiron@ualberta.ca Co-owner 🛟 Remove Invite collaborators * Email * Permissions Co-owner Editor Read only Submit

A range of options are available both for formatting and downloading your DMP...



Moving forward....



• DMP Assistant

- →DMP Assistant 2.0 Bilingual webinars soon to come see Portage events page
- → New discipline and methodology specific templates
- → More DMP exemplars to come and based on the new templates
- → More features and support for machine actionable DMPs

DMP Assistant Steering Committee

→consists of a wide range of stakeholders (e.g., Portage, host institution, researchers, Tri-Agencies, First Nations Information Governance Centre)

→Helps to guide the ongoing development and sustainability of DMP Assistant

→Helps to define and prioritize work flow and roadmap activities

DMP Expert Group

- →New DMPEG members
- →DMP Assistant 2.0 tutorial video working group
- →DMP Assistant 2.0 customization guide working group

Portage Webinar – Support Your Research with DMP Assistant 2.0!

Date: March 30, 2021

Time: 1 pm ET

Location: Online

Registration



Data management plans, or DMPs, are one of the foundations of good research data management. Hosted by the University of Alberta Library and supported by the Portage Network, the DMP Assistant is a national, open, bilingual data management planning (DMP) tool to help researchers better manage their data throughout the lifespan of a project. The tool develops a DMP by prompting researchers to answer a number of key data management questions, supported by best-practice guidance and examples. Building on the preceding DMP-focused webinar, this session will be of interest to researchers, graduate students, librarians, and research support stakeholders. Participants will take an in-depth look at the newly launched DMP Assistant 2.0, including all of its enhanced key features for both end-users and institutional administrators, as well as a brief look at the future of the platform.

Speaker:

Robyn Nicholson, Data Management Planning Coordinator, Portage Network

This webinar will be presented in English with simultaneous interpretation in French and will be recorded. Engagement during the webinar in both official languages is welcomed. Previous webinar recordings are available on the CARL YouTube Channel. Links to the recordings and slides can also be found on Portage Training Resources.

Register here!







Questions & Discussion

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

jdoiron@ualberta.ca