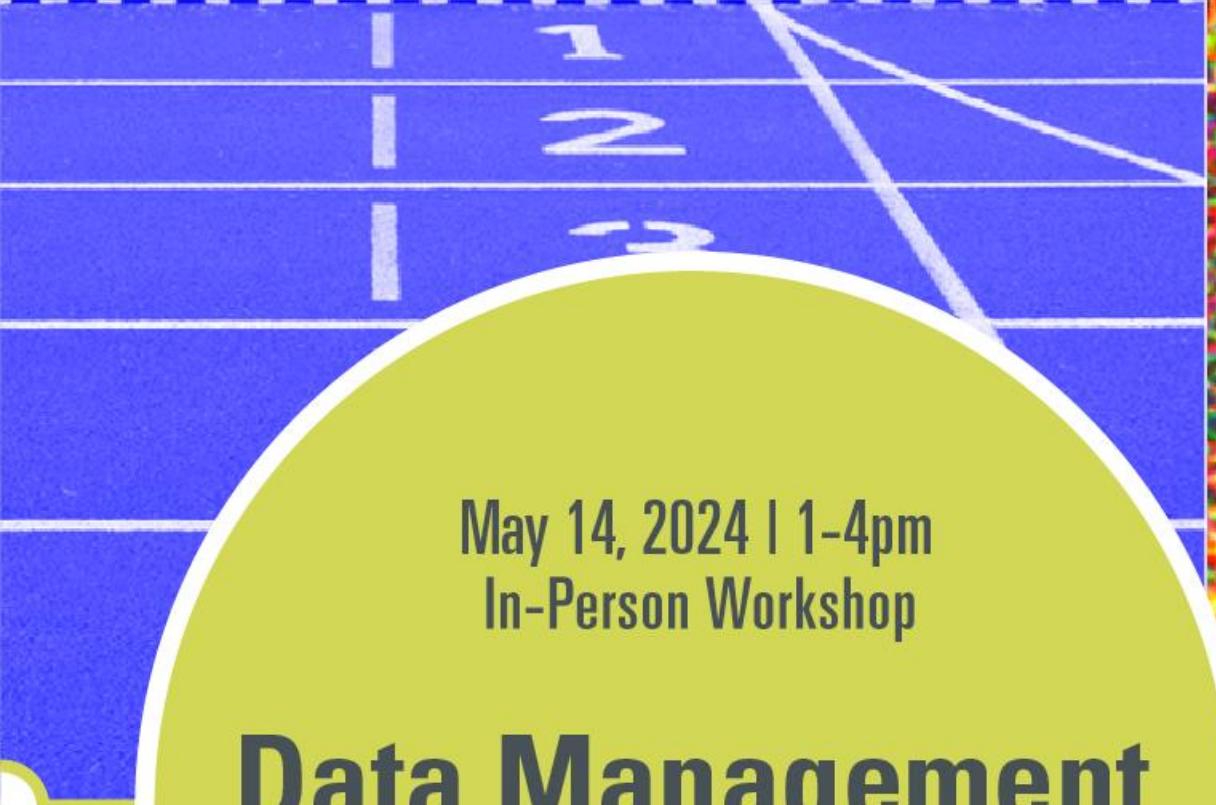
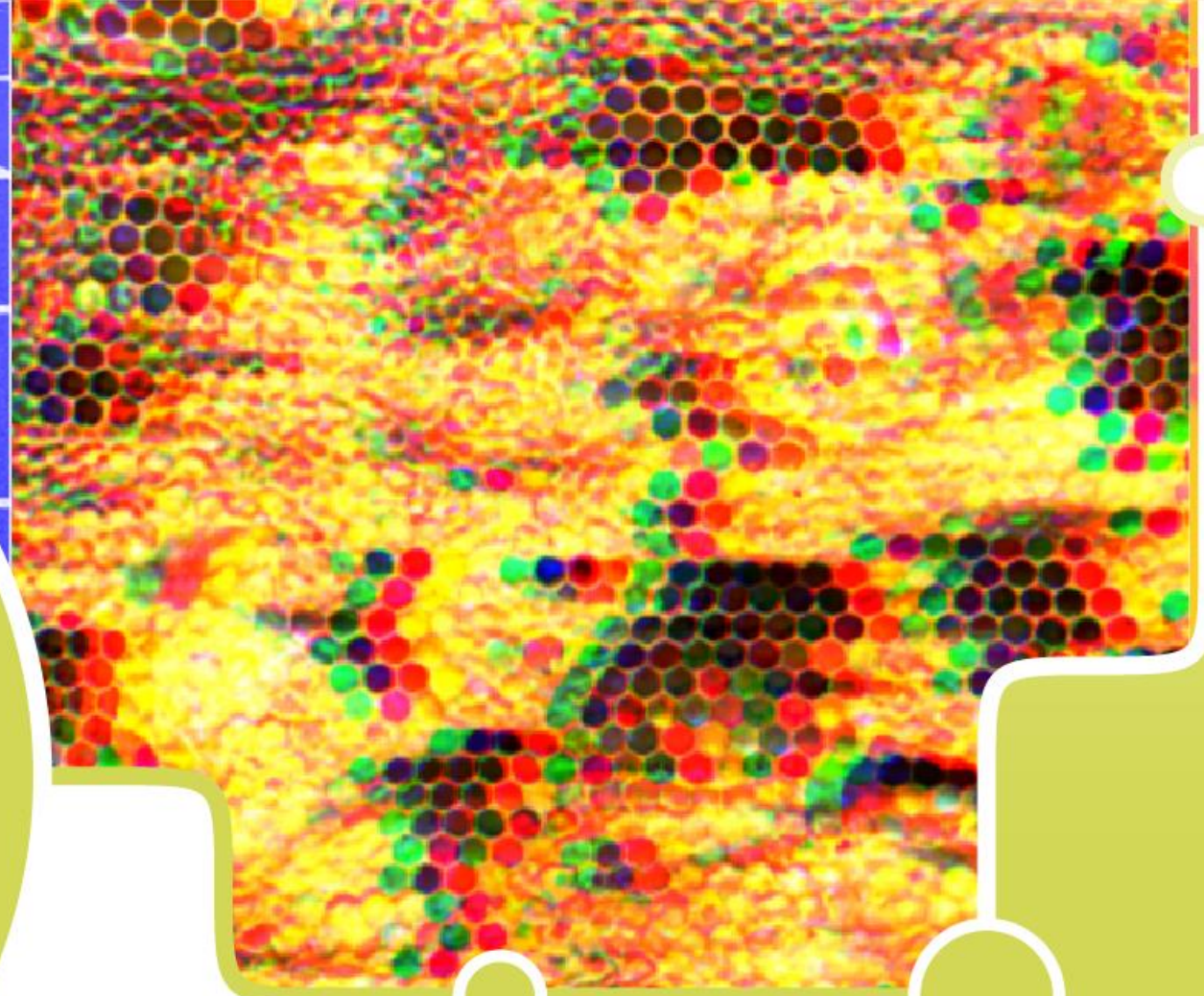




May 14, 2024 | 1-4pm
In-Person Workshop

Data Management Plan Boot Camp

u.mcmaster.ca/scds-events

RDM

SCDS

Library

McMaster University

Who are we?

Research Data Management Specialists

Isaac Pratt, PhD
My background is in **Biological Anthropology, Medical Imaging,** and **Human Anatomy**.

I have a PhD in **Anatomy & Cell Biology** from the University of Saskatchewan.

Danica Evering, MA
My background is in **social practice art, community-based research, communications studies**, and **medical laboratory healthcare**.

I have an MA in **Media Studies** from Concordia University.

McMaster University | Library

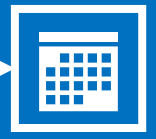
Outline

Goal Setting + Introductions [20 mins]

Data Management Plans (DMPs) Overview [15 mins]

Working Session [5 mins intro, 20 mins work time]

- Responsibility – Ethical, legal, and commercial
- Data Collection – Describe sources and formats
- Documentation – Make data understandable and reproducible
- Active Data Management – Store, secure, and back up data
- Long-Term Data Management – Preservation, discovery, accessibility



Next Steps – Sharing, Review, Use [15 mins]

DMP Assistant

- A web-based, bilingual data management planning tool
- Available to all researchers in Canada
- Walks you through relevant questions for data management.
- Exportable data management plans
- Send to RDM Services for review!
- Access at dmp-pgd.ca/plans

McMaster University | Library

Data Management Plans: Set up a system for best practices for your project.

A Data Management Plan (DMP) is a **living document** describing your plan for how you will create, store, organize, document, secure, preserve, and share your research data.

- DMPs are meant to help you **articulate your plans for managing data**; *they are not compel researchers to manage data differently.*
- DMPs outline how you will manage data both **during** the active phases of your research and **after** the completion of the research project.

Refreshment booth, Lake Shore Blvd W, elevations, plan & section June 18, 1936. Architect: M.A. Barten. City of Toronto Archives, Series 410, File 612, Item 1

Photo by National Cancer Institute on Unsplash.

DMPs **support collaboration** within your lab or research team.

- Engage your research team, partners, and collaborators in an **ongoing conversation** about how to best manage research data.
- Implement consistent data practices for a lab.
- Set up storage and security systems, with timelines for backups, transfers, and updates.
- Write out responsibility charts and contingency plans for unexpected events – illness, moving universities, ransomware attack.

Data Management Plans are **“living documents”**

- A living document reflects the inevitability of change.
- A living document is edited and updated on an ongoing basis.
- Update your DMP as your project evolves and book a time to review it every year
- Approached this way, a DMP can be a very useful research tool!

Waves breaking at the Fox Island boat launch with Shaws Cove in the background, animated. By Thorsten Lindner on Wikimedia Commons

What to keep in mind while creating a Data Management Plan

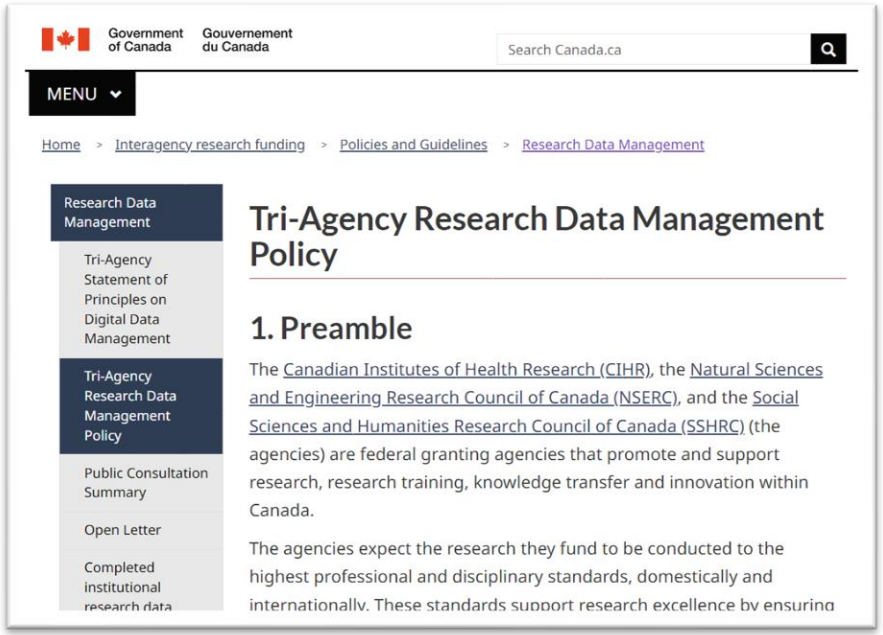
Set out consistent strategies for the whole research team on how data will be managed, shared, and archived.

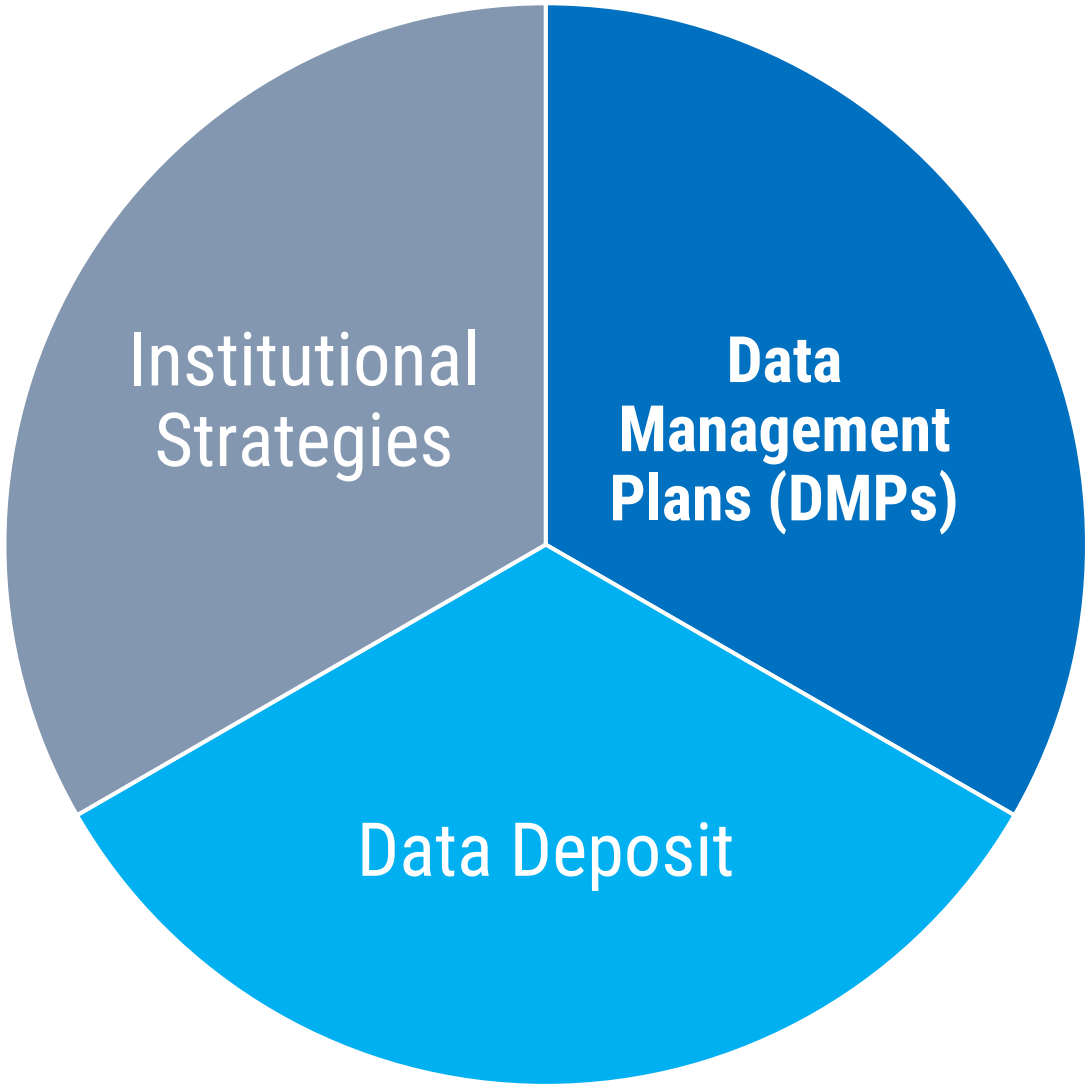
Set the **scope** of your DMP – is it for a single research project? For your whole lab? For an entire research centre?

Ensure **quality assurance** and decide how to integrate effective data management practices into your research.

Who is your audience? Write to an educated but unfamiliar reader

Tri-Agency RDM Policy 2021





McMaster University | Library

"All grant proposals submitted to the agencies should include methodologies that reflect best practices in RDM. **For certain funding opportunities, the agencies will require data management plans (DMPs) to be submitted to the appropriate agency at the time of application**, as outlined in the call for proposals; in these cases, the DMPs will be considered in the adjudication process.

Innovation, Science and Economic Development Canada. "Tri-Agency Research Data Management Policy." Government of Canada. Innovation, Science and Economic Development Canada, March 15, 2021. https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy

Tri-Agency Data Management Plan Requirements:

Grants that have recently rolled out DMP requirements:

Canadian Institutes of Health Research (CIHR)

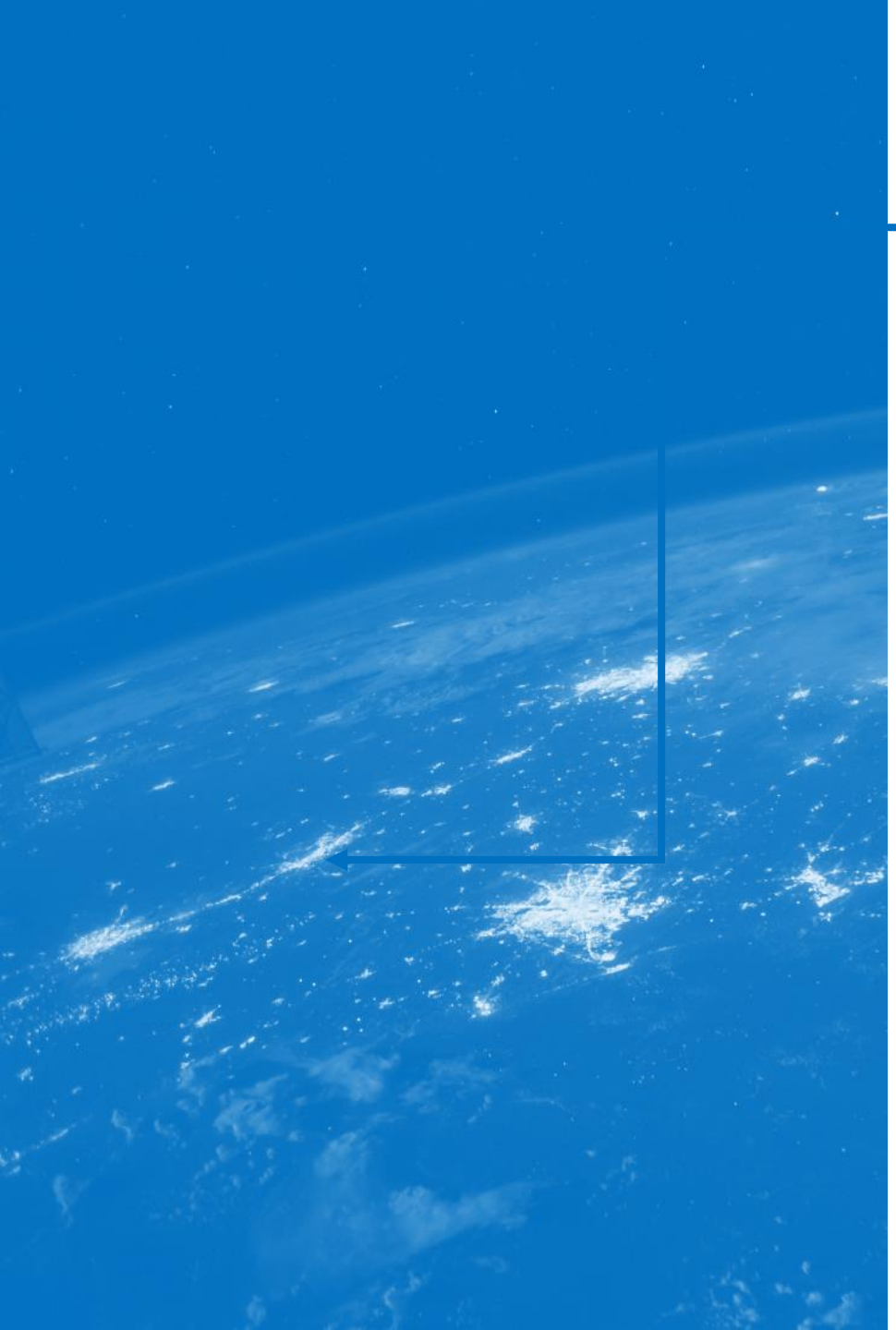
- Team Grants: HIV/AIDS and Sexually Transmitted and Blood-Borne Infections Community-Based Research (date to be determined)
- Directed Grant: Canadian Immunization Network Renewal (date to be determined)

Natural Sciences and Engineering Research Council of Canada (NSERC)

- Subatomic Physics Discovery Grants - Individual and Project

Social Sciences and Humanities Research Council of Canada (SSHRC)

- Partnership Grants Stage 2



Data Management Plans are a global practice.

- Most UK & EU funders, Wellcome Trust
- Other Canadian funders, such as the International Development Research Centre
- United States - National Science Foundation (NSF) Grants
- United States - National Institutes of Health (NIH) released their Data Management and Sharing Policy– requiring a **Data Management and Sharing Plan (DMSP)**

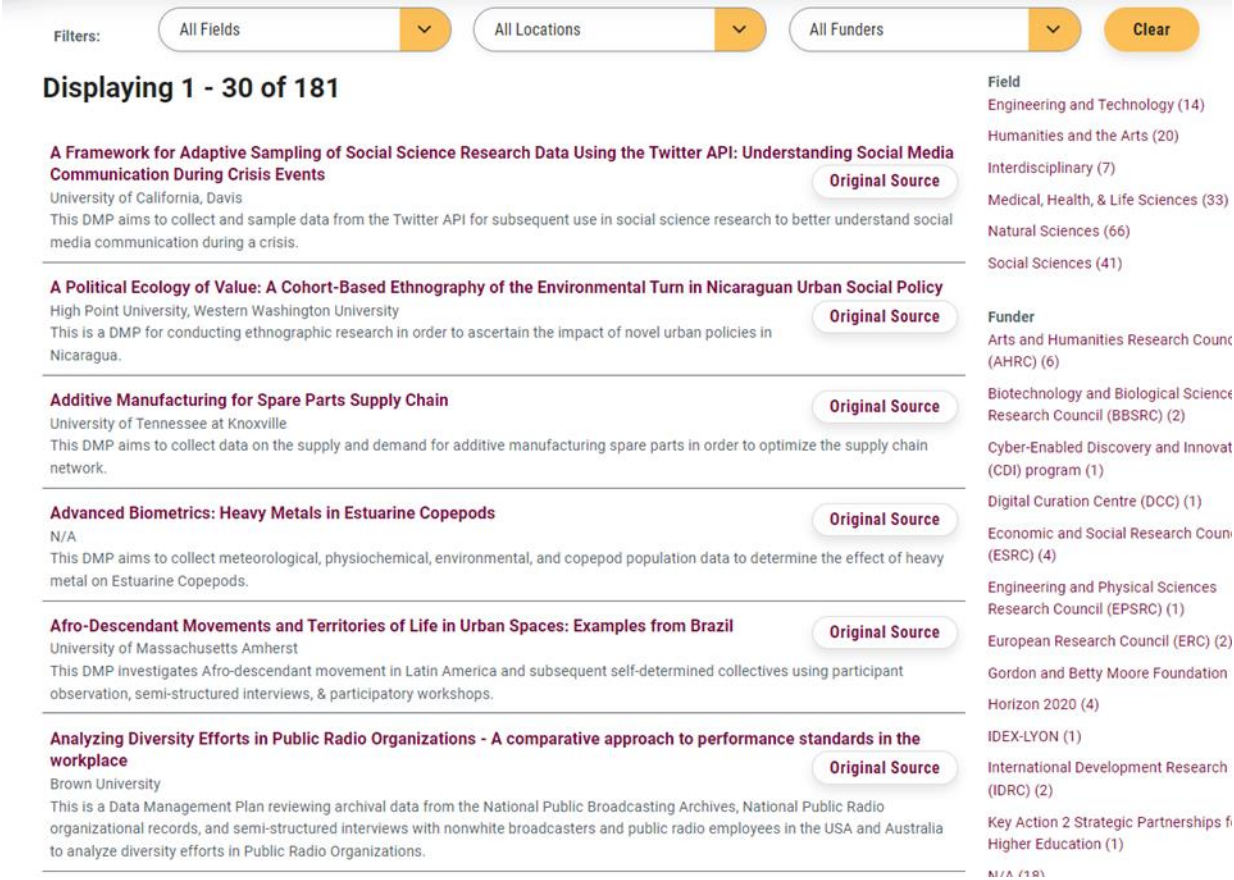
Funder Research Data Management Links:

- Tri-Agency RDM Policy
- Tri-Agency Funding opportunities requiring data management plans
- Tri-Agency Research Data Management Policy - Frequently Asked Questions
- SSHRC Guide to preparing a data management plan

- NIH Data Management & Sharing Policy Overview

McMaster DMP Database

- DMPs can vary across disciplines, methodologies, and data types.
- Over 180 example DMPs from resources across the world.
- Search by field, **location**, funder (use Canada to find SSHRC-specific DMPs).
- Submit your DMP for other researchers!
- Access at rdm.mcmaster.ca/dmps



McMaster University | Library

DMP Resources

The Digital Research Alliance of Canada has several DMP exemplars, available here. Our new DMP Database is available here.

We're going to look at the DMP for the *"People, Places, Policies and Prospects: Affordable Rental Housing for Those in Greatest Need"* project.

https://zenodo.org/record/4062466

Catherine Leviten-Reid, Jasmine Hoover, Cape Breton University.

Storage, Access, and Backup

Describe where, how, and for how long data will be securely stored during the active *phases* of the research project. If any data are to be collected through the use of electronic platforms, account for their usage within your data storage description. Include a description of any policies and procedures that will be in place to ensure that data are regularly backed-up.

All data storage and backup procedures will be clearly outlined within the project's data collection policies and procedures which will be developed prior to data collection. These procedures will also indicate where data will be stored throughout the active stages of the project.

As we are collecting survey data using REDCap software, the raw data will be transferred using a FTP process, and will be stored securely. REDCap servers are locally hosted by WCHRI within the Faculty of Medicine and Dentistry and undergo regular backups (incremental and full). Our virtual research project space has a regularly established scheduled incremental and full backup process in place to ensure no data loss occurs.

Qualitative interviews will be conducted using virtual recordings over Zoom. Upon completion of the interviews the data will be securely transferred within 48 hours to the virtual research project space located on Digital Research Alliance of Canada's cloud platform. Once the interviews are uploaded to the cloud platform they will be permanently deleted from the local computers on which the interviews were saved. This VRE undergoes backups on a regular schedule which include incremental and full backup processes.

Describe how members of the research team will securely access and work with data during the active phases of the research project.

All data will be securely stored on the Digital Research Alliance of Canada's (DRAC) cloud platform and these data will be accessible only by approved researchers, trainees, and project

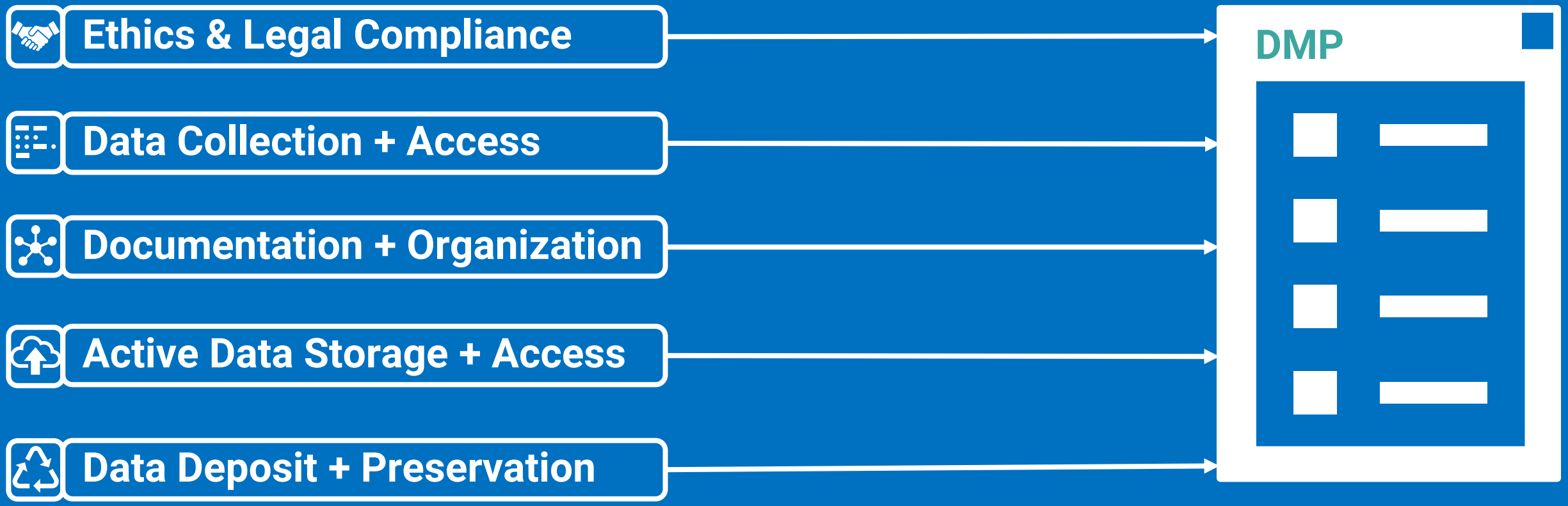
Today’s template: Alliance Pre-funding Template

- Simplified and condensed DMP Template focusing on the funding application stage
- Developed by the Digital Research Alliance of Canada’s Data Management Planning Expert Group (DMPEG)
- In DMP Assistant you will find it called the

  McMaster – Funding DMP Template (General Purpose)

Photo: McMaster University Georgia Kirkos

What goes in a Data Management Plan?



DMP Assistant

McMaster University | Library

Responsibility

- **What considerations will you take into account with respect to ethical, legal, or commercial issues?**
  - Describe any applicable ethical, legal, or commercial considerations related to your project and data.
  - Does your research involve Indigenous communities and knowledges, human subjects, legal and commercial considerations/agreements, partnerships or data with a high level of risk associated with it?
  - How will you support long-term data stewardship, including data deposit and appropriate sharing? Note - research involving human participants will very often require both informed consent and ethics approval with respect to data sharing.
  - How will data deemed as being sensitive be effectively safeguarded and protected across the lifecycle of your research project?

Ethics & Legal Compliance: **Resources**

Indigenous Data Management:

- First Nations OCAP Principles - https://fnigc.ca/ocap-training/
- CARE Principles for Indigenous Data Governance - https://www.gida-global.org/care
- McMaster Indigenous Research Primer - https://miri.mcmaster.ca/indigenous-research-primer/
- McMaster: McMaster Indigenous Research Institute (MIRI) and Indigenous Health Learning Lodge (IHLL)

Securing sensitive data:

- Sensitive Data Toolkit for Researchers
  - Glossary of Terms for Sensitive Data https://zenodo.org/record/4088946
  - Human Participant Research Data Risk Matrix
  - Research Data Management Language for Informed Consent
- McMaster RDM page 'Secure' https://rdm.mcmaster.ca/secure
- Trudi Wright, Privacy & Records Management Specialist – privacy@mcmaster.ca

Ethics & Legal Compliance

No sensitive data will be shared. Any sensitive data will be stored on secure servers for 5 years.

Research has been approved by the Research Ethics committees at the various institutions involved in the project. Participants are also required to sign the informed consent agreement. By mediating data requests and determining their own terms of access, researchers maintain their rights to the intellectual property.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

Data Collection

- **What data will you collect or otherwise bring into your project under this plan? Describe the data that will be collected, generated, and/or acquired.**
  - What types of data do you have? (e.g., image data, textual data, numerical data, audiovisual, etc)
  - What format is your data in?
  - How much data do you have?
  - How will the data will be collected or generated? - if existing data will be used, what is the source of the data and what data will be used?
  - Indicate if your data is, or may be, considered sensitive (e.g., health, administrative and/or clinical records, participant interviews).
  - Indicate if your data involves Indigenous communities and/or knowledges and information.

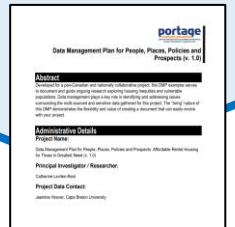
Data Collection: **Resources**

- File formats:
  - DataverseNO Prepare your data: Preferred file formats https://site.uit.no/dataverseno/deposit/prepare/#what-are-preferred-file-formats
  - DCN Data Curation Primers on preserving different file formats https://datacurationnetwork.org/outputs/data-curation-primers/
- McMaster RDM page on file naming and organization
  - https://rdm.mcmaster.ca/organize#tab-file-folder-organization

Data Collection

Data collected during our projects may include, but are not limited to, those gathered from surveys, in-depth interviews, focus groups, community conversations and arts-based methods such as photography. This means we will potentially generate numeric, audio, image, video and text-based data.

This research project is collecting a variety of types of data. Examples of these include XML and CSV for databases and spreadsheets, JPG or TIFF files for images, MP3 files for sound and TXT for text. Each of these file types are non-proprietary, ensuring ease and flexibility of reuse.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

Data Collection

All files will use a conventional **naming standard**. File names should include the **grant name** (in shortened form), a **summary** of the file's content, the **region** and the **date** (in the format YYYY/MM/DD).

An example is the following:

prospects_interviewguide_ON_20200617.

Document **versions** should be named sequentially (with file names ending in v1, v2 etc.).

An example is the following:

prospects_interviewguide_ON_20200617_v1.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

Documentation - Make data understandable and reproducible

- **How will you document data for future re-use or validation?**
  - Consider how you will ensure that your data is efficiently documented and captured through out your project (e.g. scripting, tool-generated, user-documented), including by whom, and identify needs, including staffing resources and expertise, that may possibly be built into your application and supported through grant funds.
  - If applicable, specify any data and/or metadata standards that are being used to support your research project.

Documentation & Metadata: **Resources**

Data Documentation

- McMaster RDM page https://rdm.mcmaster.ca/organize#tab-file-folder-organization
- McMaster README Template - https://rdm.mcmaster.ca/sites/default/files/YYYYMMDD_AUTHOR_DATASET_ReadmeTemplate.txt

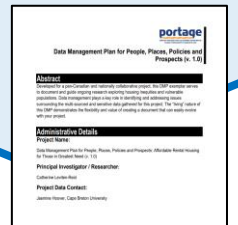
Metadata

- Dataverse North Metadata Best Practices Guide v 2.0 http://hdl.handle.net/2429/73609
- DCC list of disciplinary metadata standards https://www.dcc.ac.uk/guidance/standards/metadata

Documentation & Metadata

In order for data to be potentially reused, all data files should include a description of **team members** responsible for creating the data, **how** the data were collected, the **code book** (if involving survey data), the **interview guide** (if involving qualitative data), any issues affecting data quality and other pertinent background information which allows the content to be easily understood by others.

All files containing spreadsheets must include column names which are easily interpreted, even though they will be defined in a code book.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

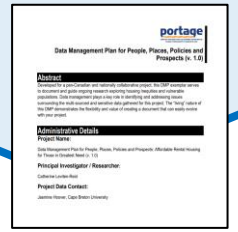
Principal Investigator / Researcher:

Project Data Contact:

Documentation & Metadata

Team researchers engaged in data analysis using software will create **logs and syntax files** to ensure that the steps leading to the final results are documented and saved. No identifying information of participants may be included in data files. Metadata must also include the grant name and funders (SSHRC and CMHC).

Since these data files will be deposited in the Scholars Portal Cape Breton University Dataverse, the **Data Documentation Initiative (DDI) metadata standard** will be applied.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

Active Data Management

- **How will data be stored, accessed and worked with?**
  - Describe where and how data will be stored, accessed, and worked with during the active phases of your research including as applicable:
  - All versions of data (e.g., raw, source, study, analytic, de-identified)
  - All activities (e.g., data collection, processing, analysis, dissemination)
  - All software and platforms
  - Who will have access to what data, including security measures (e.g., Investigators, research staff, collaborators, partners)
  - How data will be backed up to prevent data loss

Storage & Backup

Storage space is anticipated to be approximately **100 GB**. The data will be stored for **5 years locally**, with a permanent copy held in the Scholars Portal Cape Breton University **Dataverse**.

The **3-2-1 backup rule** will be followed for data storage and backup. All team members will upload their files to a cloud-based server located in Canada, to be identified by the project lead. Sensitive files are to be encrypted.

OneDrive is used to store, share, and work with data.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

Storage & Backup: **Resources**

Data Storage:

- McMaster Research Data Storage Finder
  https://u.mcmaster.ca/storagefinder

Backup:

Copies of your data (at least!)

Copies are on-hand (easily accessible) on different systems (internal hard drive, cloud storage, etc.)

Copy is in another location (“off-site”) from the others with a ***trusted*** service provider

Example:
1 copy stored locally on **hard drive** for analysis
1 copy stored on **cloud storage** platform
1 copy stored in a **secure campus drive**

Long-Term Data Management - Ensure data are preserved, discoverable, and accessible

- Describe plans for **long-term management** of your data after the active phases of your research have concluded, including data deposit and sharing. Consider and describe as applicable plans for:
  - All versions of data deposited (raw, master, analytic, published)
  - All activities (e.g., curation, preservation, ethical compliance, publishing etc.)
  - All software and platforms (e.g., data repositories)
- Will you **deposit** any versions for discovery, appropriate access, and potential reuse?
- Are there any **responsibilities** to funders, publishers, ethics boards, research participants that impact this?

Depositing Data - **Data Repositories**

Data Repositories:

- DataCite Repository Finder tool https://repositoryfinder.datacite.org/
- McMaster Dataverse https://borealisdata.ca/dataverse/mcmaster
- FRDR https://www.frdr-dfdr.ca/repo/
- Data Repository Guidance from *Nature Scientific Data* https://www.nature.com/sdata/policies/repositories

Anonymization and Licensing

Data Anonymization:

- Portage Network De-Identification Guidance https://zenodo.org/record/4270551
- RDM – Publish – Data Anonymization: https://rdm.mcmaster.ca/publish#tab-data-anonymization
  - Recommended tools and training

Data Licenses:

- Creative Commons (CC)
- Open Data Commons
- More about licensing

Centre for Longitudinal Study on Aging
Publication and Promotion Policy

Canadian Longitudinal Study on Aging (CLSA), Publication and Promotion Policy for CLSA Approved Users, November 29, 2022, Accessed April 17, 2024, https://www.clsa-elcv.ca/doc/2579

2.5 Data Availability Statement

Some journals stipulate that the data used for the analyses be deposited with the journal or otherwise be made available to replicate the findings. The CLSA's privacy and confidentiality requirements do not permit Approved Users to share CLSA data beyond their research team, so this requirement cannot be met. Approved Users should communicate this to the journal by providing the following statement:

"Data are available from the Canadian Longitudinal Study on Aging (www.clsa-elcv.ca) for researchers who meet the criteria for access to de-identified CLSA data."

Data Availability: The CLSA data used in our study is third-party data. The CLSA data are available from the Canadian Longitudinal Study on Aging (www.clsa-elcv.ca) for researchers who meet the criteria for access to de-identified CLSA data. We had no special access privileges to the data and other researchers will be able to access the data in the same manner as the authors. The specific data application requirements and process can be found at https://www.clsa-elcv.ca/data-access/data-access-application-process. Please email access@clsa-elcv.ca to request a Magnolia user account and other information on CLSA data application.

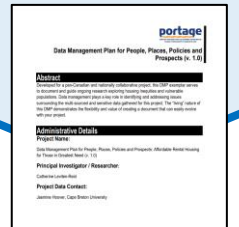
Zhang W, Sun H (2020) Formal and informal care received by middle-aged and older adults with chronic conditions in Canada: CLSA data. PLoS ONE 15(7): e0235774. https://doi.org/10.1371/journal.pone.0235774

Sharing & Re-use

The analyzed, de-identified data set or datasets will be put under **mediated access** in the Scholars Portal Cape Breton University **Dataverse**. Users will be required to request access to the data for reuse.

Requests will be **evaluated by the PI** and/or a backup member identified on the research team. Terms of access and use will be determined by the PI in consultation with the research team to ensure appropriate use of the data.

Data deposited in Dataverse will be assigned a **Digital Object Identifier** (DOI), a unique and persistent code that can be used by others to locate and access these data. **Metadata is harvested by the FRDR**, a Canada wide research repository, where data can be discovered, and then shared, at a national level. We will also link our dataset to the **publications** arising from this study.


portage

Data Management Plan for People, Places, Policies and Prospects (v. 1.0)

Abstract

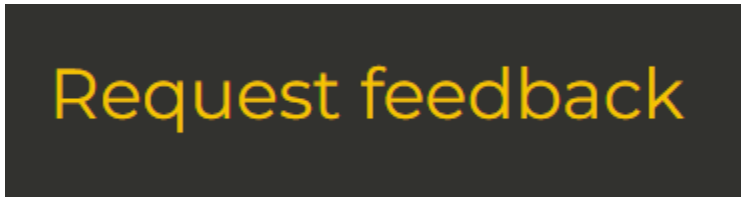
Administrative Details

Project Name:

Principal Investigator / Researcher:

Project Data Contact:

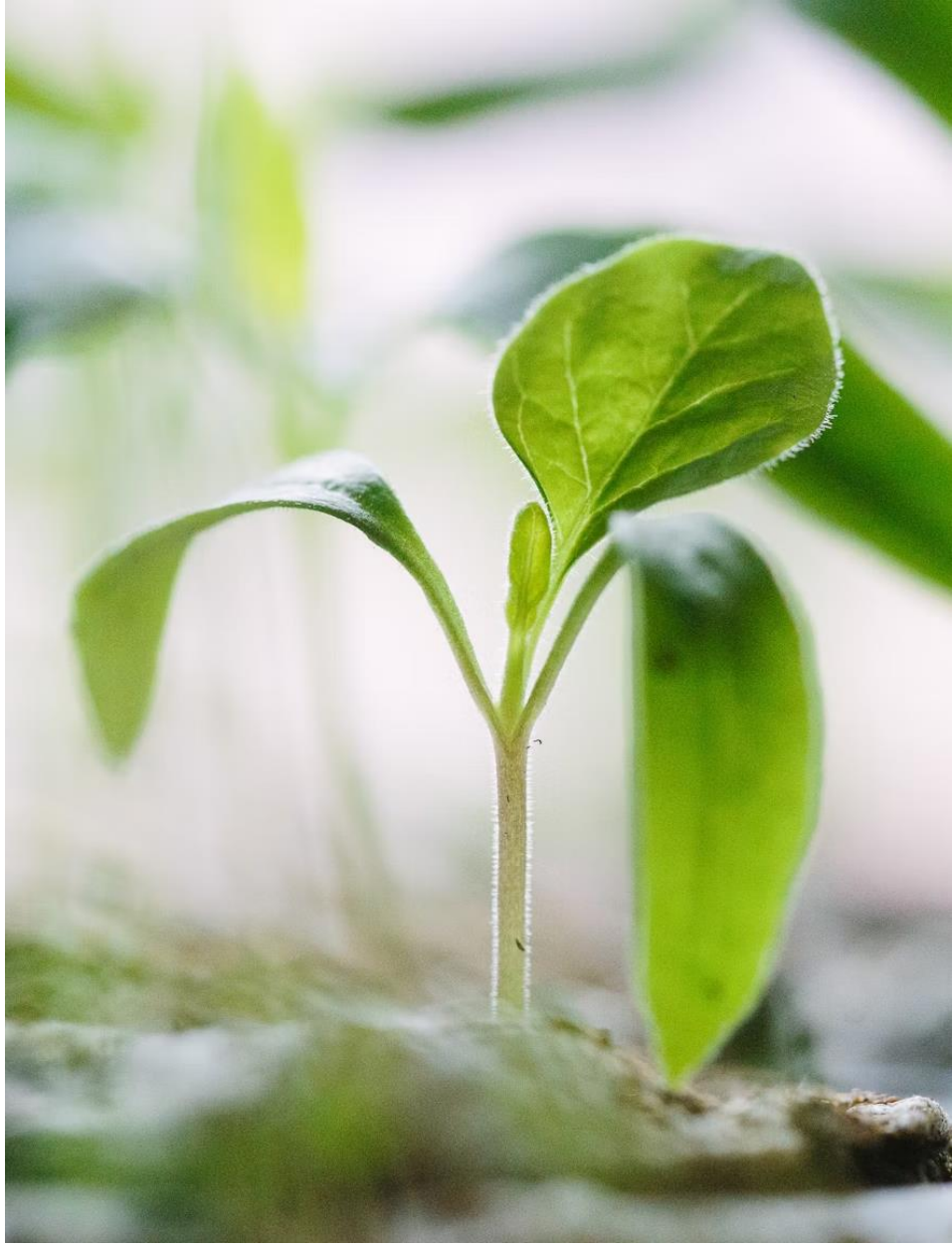
Next steps



- If you've finished your DMP, you can submit it to RDM Services for a review by hitting the Request feedback button in DMP Assistant or emailing a copy to rdm@mcmaster.ca
- If you have more questions, reach out to us by email or set up an appointment

DMPs are living documents

- Set up a review schedule for your new data management plan
- If it’s for a funding application, how will you fill out the details for a more in depth research plan
- Share it and talk about it with your students, supervisors, partners, and collaborators!



Francesco Gallarotti on Unsplash

McMaster RDM webpage: rdm.mcmaster.ca

Contact RDM services at: rdm@mcmaster.ca

Upcoming RDM webinars: rdm.mcmaster.ca/events

Recorded RDM webinars: u.mcmaster.ca/learn-rdm

Make an appointment with a Research Data Management Specialist:
u.mcmaster.ca/rdm-appointments