



LIBRARY



Introduction to Research Data Management

FGS DEVELOPING YOUR RESEARCH SERIES

Tanya Nguyen, Library Assistant, Repositories Support
October 15, 2025

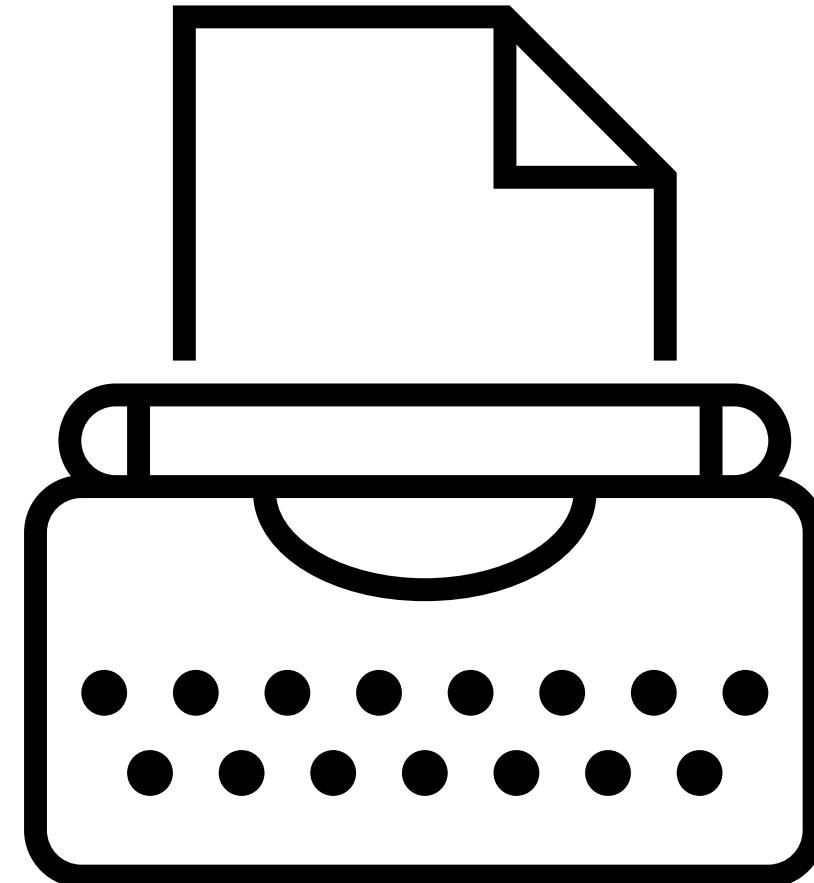
Overview

- Introduction to Research Data Management (RDM)
- Emerging Requirements for RDM
- Data Management Planning
- Creating a DMP (DMP Assistant Tool)

Activity: Research Context?

Please share your research context in the chat.

- Does your program have a research component? (project, thesis, dissertation, etc.)
- What is your research topic (and method(s) if you've decided on them)?
- What tools do you use to plan your research and manage your data?



Research Data Management (RDM)

AN INTRODUCTION

What is Research Data Management (RDM)?

Research data management is a general term for how you plan for and carry out the handling of your data for your research project from the planning stages through data gathering/generation, active analysis, and into data deposit or long-term storage and preservation as a project is completed.



5. Data Storage, Retention and Disposal

| # | Question |
|-----|---|
| 5.1 | Describe in detail how all types/forms of research data will be stored (e.g. digital files, hard copies, audio recordings, other). Specify the physical location and what means will be used to secure the data and protect confidentiality and privacy. If not applicable, enter N/A. |
| 5.2 | If you plan to destroy your data after the obligatory 5 year retention period, describe when and how this will be done. Indicate your plans for the destruction of the identifiers at the earliest opportunity consistent with the conduct of the research. If not applicable, enter N/A. |

RDM as Part of Research Culture

- Open scholarship movements 
- Making publicly funded research publicly accessible
- Responsible use of funds
- Recognizing investment and effort in data collection and curation
- Documenting rigour in research practice

SSHRC = CRSH



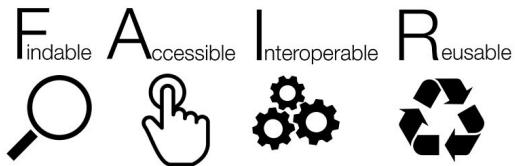
FAIR Principles

Findable Accessible Interoperable Reusable



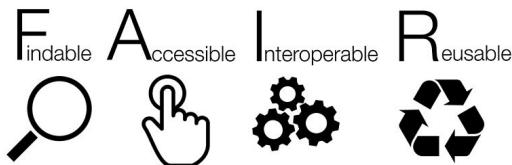
As open as possible. As closed as necessary.

FAIR Principles Knowledge Check



- Which of the FAIR principles apply to the following:
- 1) My dataset is stored in a repository and can be retrieved using a DOI
- 2) I've assigned a license to my dataset so others know how the dataset can be used.
- 3) My dataset is openly available to download from trusted repository.
- 4) My dataset is in CSV format so multiple programs can read it.

FAIR Principles Knowledge Check



- Which of the FAIR principles apply to the following:
- 1) My dataset is stored in a repository and can be retrieved using a DOI Findable
- 2) I've assigned a license to my dataset so others know how the dataset can be used. Reusable
- 3) My dataset is openly available to download from trusted repository. Accessible
- 4) My dataset is in CSV format so multiple programs can read it. Operable

CARE and OCAP Principles

[Global Indigenous Data Alliance](#); [First Nations Information Governance Centre](#)

CARE

- Collective Benefit
- Authority to Control
- Responsibility
- Ethics

OCAP

- Ownership
- Control
- Access
- Possession

Benefits & Advantages of RDM

Support efficiency and organization

- Standardizes ways data will be collected and dealt with
- Documents practices for team members' reference
 - Makes personnel transitions and communication easier

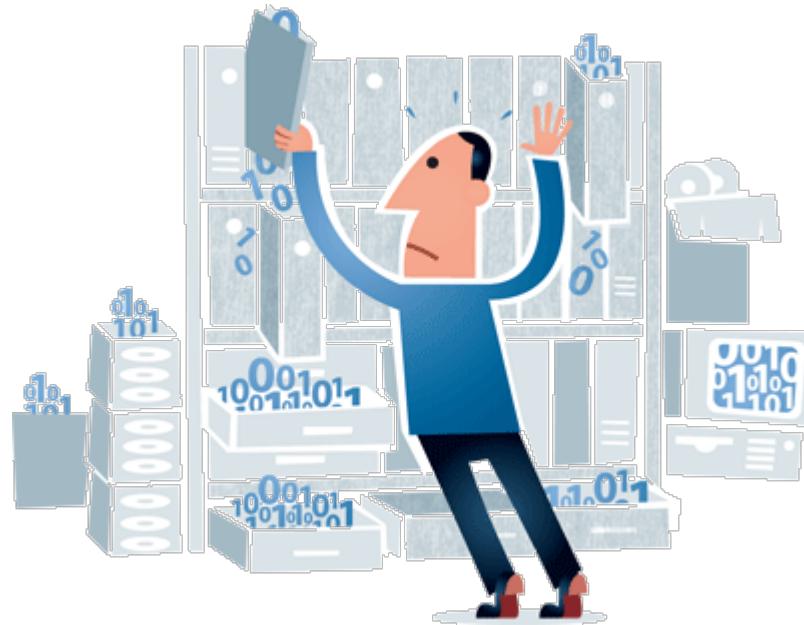
Increased visibility and recognition for work

- Planning for preservation and access
- Data availability statements and data citation

Contributes to Open Scholarship

- Potential for data reuse
- Increases replicability of studies

Comply with funder, institutional, and/or publisher policies



Takeaways: Tools and Resources

[AU Library Guide to Research Data Management](#)

[RDM PD course on PowerEd](#) (free)

- Access code: researchfree

Data Management Guides, by Eugene Barsky, UBC

- [Good Enough Research Data Management](#) (a brief guide)
- [Research Data Management Data Guide](#)

[FAIR Principles](#)

Indigenous Data Sovereignty

- [Global Indigenous Data Alliance](#)
 - [CARE Principles](#)
- [First Nations Information Governance Centre](#)
 - [OCAP Principles](#)



Emerging RDM Requirements

FUNDERS, PUBLISHERS, INSTITUTIONS

Tri-Agency RDM Policy

The policy was released in March of 2021. It has several requirements that will be phased in:

- **For institutions:**

- An institutional strategy
 - by March 2023
 - [AU's Strategy](#)

- **For researchers:**

- Data Management Plans (DMPs)
 - Requirement is being piloted
 - [Calls with requirement](#)
- Data Deposit
 - TBD

Tri-Agency Policy Requirements for Researchers

Data Management Plans (DMPs)

- Describe:
 - how data will be collected, documented, formatted, protected and preserved;
 - how existing datasets will be used and what new data will be created over the course of the research project;
 - whether and how data will be shared; and
 - where data will be deposited.
- Standardized tools are encouraged (e.g., [DMP Assistant](#))

Data Deposit

- All data, metadata, and code
- In a repository for safe storage, curation, and preservation
- Provide access “where ethical, legal, and commercial requirements allow” and according to disciplinary standards
- If possible, link data to publications with PIDs (persistent digital identifiers, e.g., DOIs)

Publishers' Requirements/Recommendations

Data availability statements explain where data associated with the research outcomes in a particular paper can be found, and what conditions need to be met to access the data.

- They ideally will also link to the dataset, particularly if the dataset is available in a repository.

Some publishers now **ask for a link to the dataset or its DOI** when you submit a manuscript.

In some disciplines there are also data journals that publish descriptions of datasets.

Data Availability Statements

Ex 1: uses T&F's template
Ex 2: direct quote from
[Springer Nature](#)

EXAMPLES

The data that support the findings of this study are openly available in Athabasca University's Data Repository Collection in Borealis at
[https://doi.org/\[doi\]](https://doi.org/[doi])

PRO-Seq data were deposited into the Gene Expression Omnibus database under accession number GSE85337 and are available at the following URL: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85337>.

Takeaways: Tools and Resources

Tri-Agency (CIHR, SSHRC, NSERC)

- [RDM Main Page](#)
 - [RDM Policy](#)
 - [Funding calls with DMP requirements](#)
- [Open Access Policy on Publications](#)
 - (note it's under review to be updated in 2025)

Publisher Data Availability Statements and Data Sharing

- [Taylor & Francis](#)
- [SAGE](#)
- [SpringerNature](#)
- [Wiley](#)
- [Elsevier](#)



Emerging RDM Requirements Knowledge Check

- 1) What are the three entities that provide guidance, recommendations, requirements related to data management?
- 2) True or False: Data should be closed as possible but as open as necessary.
- 3) A data management includes how data is:
 - A) Collected B) Documented C) Formatted D) Protected E) Preserved F) All the above
- 4) What can be included in a data deposit (3)?
- 5) What is missing in this data availability statement: The datasets generated during and/or analyzed during the current study are available in Borealis.

Emerging RDM Requirements Knowledge Check

- 1) What are the three entities that provide guidance, recommendations, requirements related to data management? **Funders, Publishers, and Institutions**
- 2) True or False: Data should be ~~closed~~ **open** as possible but as ~~open~~ **closed** as necessary. **False**
- 3) A data management includes how data is:
 - A) Collected B) Documented C) Formatted D) Protected E) Preserved F) **All the above**
- 4) What can be included in a data deposit? **Data, metadata, code, documentation**
- 5) What is missing in this data availability statement: The datasets generated during and/or analyzed during the current study are available in Borealis. **Persistent web link to the datasets.**

Data Management Planning

CONSIDERATIONS FOR DMPS

Data Management Plans (DMPs)

A DMP **describes** how data will be

- collected
- documented
- formatted
- protected
- preserved

A DMP **documents**

- if/how existing datasets will be used
- what new data will be created
- whether or how data can be shared
- where data will be deposited and/or how it will be retained and stored long-term

Ethics, Research Partners, and Legal Consideration

- Do you have human participants or require ethics approval?
- Do you have community or industry partners that have requirements about data access, handling, or ownership?
- Are there legal considerations?
 - E.g., are you accessing data through FOIP or does other privacy legislation apply?
- Have you signed or agreed to specific terms in order to access or collect data?

Ethics & RDM

Sensitive data and human subjects

Participant consent is required for data to be used beyond the current project

TCPS 2

- Privacy and Confidentiality
- First Nations, Inuit and Métis Peoples of Canada
- Qualitative Research
- Clinical Trials
- Human Biological Materials
- Human Genetic Research

TCPS 2 (2022) now includes a section on “**Broad Consent for the Storage of Data and Human Biological Materials for Future Unspecified Research**”

Takeaways: Tools and Resources

Ethical Conduct for Research Involving Humans – TCPS 2 (2022)

Portage Sensitive Data Tools

Portage (now under the Digital Research Alliance of Canada) created a three-part toolkit.

- [A Glossary of Terms for Sensitive Data used for Research Purposes](#)
- [A Human Participant Research Data Risk Matrix](#)
 - Intended to help researchers determine risk level and make decisions about data management (including deposit and appropriate access)
- [A document providing samples of Research Data Management Language for Informed Consent](#)
- [De-identification Guidance](#)
 - Created by the Portage COVID-19 working group as part of a 5-document series of guidance on sharing COVID-19 research data. Links to these documents are available in the [AU Library Guide to RDM here](#).



Document Your Data and Processes

- Do you use a data collection protocol?
- How will data be described and documented?
 - Are any metadata standards relevant?
- How will files and folders be named and organized?
- How will versioning be handled?
- What additional information is needed to understand the data?

File Naming

CONSIDER THESE IN YOUR FILENAMES

- Project or experiment name or
- Location/spatial coordinates
- Researcher name/initials
- Date or date range of experiment
- Type of data
- Conditions
- Version number of file

OTHER TIPS

Date_GroupCode_datatype
2019-06-29_Lilac_audio.mp3
2019-06-29_Lilac_video_webcam.mov
2019-06-29_Lilac_transcript.docx

- Use YYYY-MM-DD or YYYY-MM-DD format
- Use some form of version control (manual or tool based)

ctices and structure
s
e_20210929
20210929

Documentation Examples

- Data dictionaries
- Survey instruments
- Code books
- Lab notebooks (print or electronic)
- Field notes and memos
- Scripts or code
- Documented computing environments
- Readme files

Readme example

Project: Kristin's important chemistry project

Date: June 2013-April 2014

Description: Description of my awesome project here

Funder: Department of Energy, grant no: XXXXXX

Contact: Kristin Briney, kristin@myemail.com

ORGANIZATION

All files live in the 'ImportantProject' folder, with content organized into subfolders as follows:

- 'RawData': All raw data goes into this folder, with subfolders organized by date
- 'AnalyzedData': Data analysis files
- 'PaperDrafts': Draft of paper, including text, figures, outlines, reference library, etc.
- 'Documentation': Scanned copies of my written research notes and other research notes
- 'Miscellaneous': Other information that relates to this project

NAMING

As follows:

"YYYYMMDD_experiment_sample_ExpNum"

(ex: "20140224_UVVis_KMnO4_2.csv")

Raw data files will be named

STORAGE

All files will be stored on my computer and backed up daily to the shared department server. I will also keep a backup copy in the cloud using SpiderOak.

Takeaways: Tools and Resources



[File naming advice](#) (AU Library Guide) and [worksheet](#)

[Research Data Alliance Metadata Standards Catalog](#)

- Browse standards by disciplinary area, view use cases, and explore tools.

Readme Templates (for documenting your data)

- [FRDR's template](#) (for data deposit)
- [Cornell's Guidance and Template](#)
- [AU file naming and folder organization readme template](#)

[The Turing Way Guide for Reproducible Research](#)

- Has sections on documenting computational environments and code

Document Your Data Knowledge Check

How might you name this file based on these characteristics?

- It's a csv file
- It was created August 2, 2025
- Deidentified survey responses from adolescents about social media use using the project name ADSM
- Version 2

Document Your Data Knowledge Check

How might you name this file based on these characteristics?

2025-08-01_ADSM_deidentified_survey_v2.csv

(this is just one way example of many possible file names, but the important point is being descriptive and consistent when file naming)

[Active] Data Storage

- How will data be stored, accessed, and worked with?
 - Where will data be stored?
 - What appropriate security measures are in place?
 - Are different locations used for different versions (master copies, analytic copies, working copies, identifiable and de-identified copies, etc.)?
 - What is the process (or processes) for transferring data?
 - How will data be backed up, and where?
 - Who needs to access the data?
 - What software and platforms will be used?

Storage Types

- Networked Drives
- Cloud Storage
- Portable Storage Media
 - e.g., DVDs, CDs, USBs, external hard drives
- Your computer's hard drive



Storage Service Provider Considerations

- Consider terms of service
 - Check for policies on data use
- Remember that cloud services use geographically distributed data centres
 - Be aware of laws in different jurisdictions
- Check whether the provider uses recognized security standards
 - e.g., ISO 27001 or 27002
- Consult with AU IT (and/or IT departments of your collaborators).
 - Enterprise contracts go through security and privacy review processes.
 - Email helpdesk@athabascau.ca, and ask for help from “Research Tools Squad”
 - Note some services may require your supervisor to fund or sign off on.
- Digital Research Alliance of Canada has research computing services.

Backups

A useful general recommendation for backups is the 3-2-1 rule:

- Have at least 3 copies of your data.
- Have your data in at least two different locations
 - (i.e., local and off-site copies)
- Use more than one type of storage media

Cloud providers market the durability of their storage. Be sure you understand how it works (it's still possible to accidentally delete or overwrite your own data).

Takeaways: Tools and Resources

Technical Support (AU IT)

- Email helpdesk@athabascau.ca
- Ask for help from “research tools squad”

[Digital Research Alliance of Canada](#)

[ARC \[previously Compute Canada\]](#)

- Advanced Research Computing
- Data storage and back-up
- Note: students need to be sponsored by faculty supervisors



Computing and Storage Knowledge Check

What are some considerations to think about when using cloud service providers?

- What is the 3-2-1 backup rule?

Computing and Storage Knowledge Check

What are some considerations to think about when using cloud service providers?

- Data is held in geographically distributed data centres (held in different legal jurisdictions), terms of use (used to train models?), security protocols (encrypted data at rest and in transit)...
- What is the 3-2-1 backup rule?
 - Have at least 3 copies of your data.
 - Have your data in at least two different locations
 - (i.e., local and off-site copies)
 - Use more than one type of storage media

Long-term Storage / Data Deposit

- Explain what data will be kept.
- What value does the data have long-term?
- For what purpose(s) are you keeping the data?
- Are there other policy or ethical considerations that affect how long data should be kept or by whom?

Data Deposit

“Data deposit’ refers to when the research data collected as part of a research project are transferred to a research data repository. The repository should have easily accessible policies describing deposit and user licenses, access control, preservation procedures, storage and backup practices, and sustainability and succession plans. The deposit of research data into appropriate repositories supports ongoing data-retention and, where appropriate, access to the data.

Ideally, data deposits will include accompanying documentation, source code, software, metadata, and any supplementary materials that provide additional information about the data, including the context in which it was collected and used to inform the research project. This additional information facilitates curation, discoverability, accessibility and reuse of the data.”

Data Deposit

Deposit your data in a digital repository. Deposit makes your data more “FAIR”.

- **Disciplinary repositories** – some disciplines have repositories that are well established. It is a good idea to deposit in a repository recognized in your discipline because your data will be more likely to be found by people in your field.
- **Institutional repositories** – many institutions have data repositories. AU has an institutional repository you can use..
- **Generalist repositories** – non-discipline specific repositories. They typically take data of any type. The Federated Research Data Repository (FRDR) is an option for medium to large datasets with a default storage allocation of 1TB.

Data Deposit

When selecting a repository consider:

- Fit for your project and discipline
- Curation
- Provision of persistent identifiers
- Preservation
- Certification



Takeaways: Tools and Resources

- [AU Data Repository](#)
 - [Guide](#) for depositing data
- [Federated Research Data Repository \(FRDR\)](#)
 - A nationally supported repository
- [Registry of Research Data Repositories](#)
 - For finding repositories
- [NIH Desirable Repository Characteristics](#)
 - For evaluating repository options



Creating a DMP

USING DMP ASSISTANT

Using DMP Assistant

<https://dmp-pgd.ca/>

- Create an account
- Select a template
- Share
- Request feedback
- Download/Export

The screenshot shows the DMP Assistant web application. At the top, there's a navigation bar with links for 'My Dashboard', 'Create plans', 'Reference', 'Help', and 'About'. Below the navigation is a header with the AU Library logo and links for 'AU Library & Scholarly Resources', 'AU Library RDM Guide', and 'library@athabascau.ca'.

The main content area is titled 'Sample Plan VIII'. It features a navigation bar with tabs: 'Project Details', 'Contributors', 'Plan overview' (which is highlighted in yellow), 'Write Plan', 'Research Outputs', 'Share', 'Request feedback', and 'Download'. Below this is a toolbar with 'expand all | collapse all' and a progress bar showing '0/6'.

The main workspace contains two sections: 'Introductory Guidance (0 / 1)' and 'Plan Questions (0 / 5)'. Under 'Introductory Guidance', there's a question about ethical, legal, or commercial issues, followed by a rich-text editor and a note about applicable considerations related to the project and data. To the right, there's a sidebar titled 'Comments & Guidance' with a section for 'Alliance' containing text about managing data throughout the project lifecycle and a list of notable key considerations.

Takeaways: Tools and Resources

DMP Assistant

Examples:

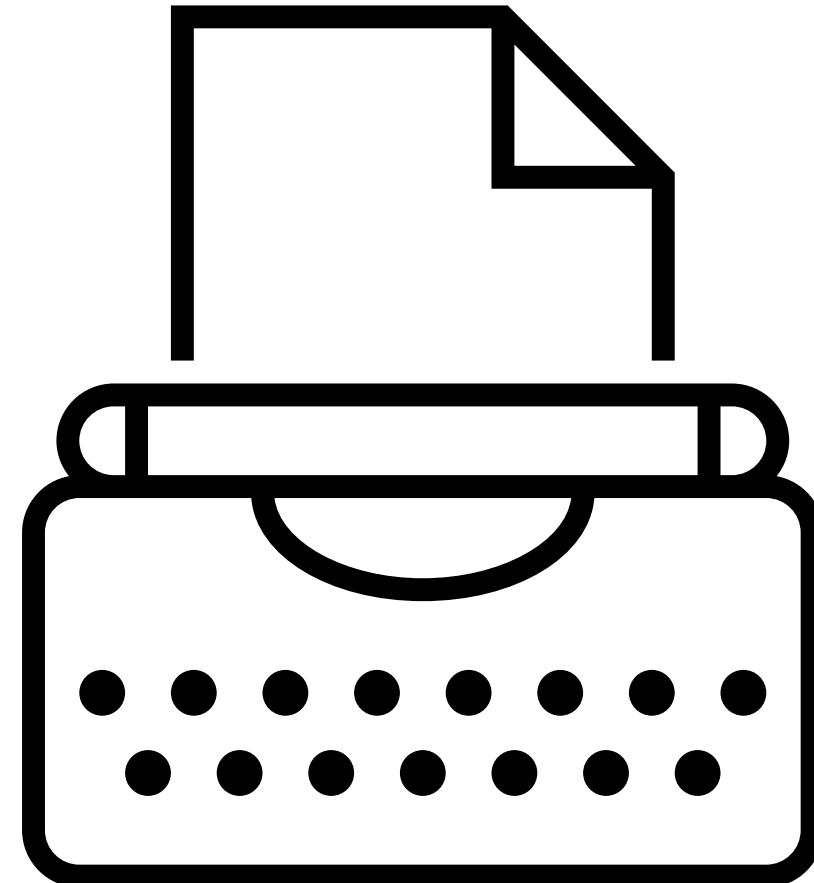
- McMaster University has created a [DMP Database](#)
- [Examples](#) have also been published by the Digital Research Alliance of Canada (the link is to Zenodo search results)



Activity: Research Context?

Based on the research context you shared earlier:

- Are there any tools or tips you want to try?
- Do you have other tips to share?



Thank-you! Any Questions?



Want to follow up? I can be reached at:
library@athabascau.ca

You can also check out our AU Library Guide:
<https://libguides.athabascau.ca/rdm>

Or, book an appointment:



From Memorial University?

Reach out to Alison Farrell at
alisonr@mun.ca

Memorial University Library's RDM
Guide:

<https://guides.library.mun.ca/datamanagement/rdm>

Love Data Week

Other AU Events

- Understanding Data Deposit
 - Thursday, February 13, 1:00PM, [registration link](#)

- Launch of the [Data and Statistics Library Guide](#)
 - Check it out and provide feedback.

ICPSR Love Data Week

- [Event calendar](#)



References

Bishop, B. W., & Hank, C. (2018). Measuring FAIR Principles to inform fitness for use. *International Journal of Digital Curation*, 13(1), 35–46. <https://doi.org/10.2218/ijdc.v13i1.630>

Briney, Kristin. (2014, February 25). Readme.txt. DATA AB INITIO: Managing data right, from the start. <http://dataabinitio.com/?p=378>

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2022, June 1). *Research Data Management*. Innovation, Science and Economic Development Canada. <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management>

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada“ (2021). *Tri-Agency Research Data Management Policy* Innovation, Science and Economic Development Canada. https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97610.html

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council. (2021b, October 28). *Tri-Agency research data management policy—Frequently asked questions*. Innovation, Science and Economic Development, Government of Canada; Innovation, Science and Economic Development Canada. <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy-frequently-asked-questions>

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2016, December 21). *Tri-Agency statement of principles on digital data management*. Innovation, Science and Economic Development Canada. https://www.ic.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html

First Nations Information Governance Centre (FNGIC). (n.d.). *The First Nations Principles of OCAP®*. Retrieved February 11, 2021, from <https://fnigc.ca/ocap-training/>

GO FAIR. (n.d.). FAIR Principles. Retrieved February 11, 2021, from <https://www.go-fair.org/fair-principles/>

Leadership Council for Digital Research Infrastructure (LCDRI). (2017). Data management position paper: For Innovation, Science, and Economic Development Canada. Retrieved August 11, 2020 from <https://wiki.queensu.ca/display/QDSRWG/Queen%27s+Digital+Scholarly+Record+Working+Group+Home>.

Portage Network, (n.d.). Portage DMP Assistant. Available when logged in at <https://assistant.portagenetwork.ca/>

Research Data Alliance International Indigenous Data Sovereignty Interest Group. (September 2019). “CARE Principles for Indigenous Data Governance.” The Global Indigenous Data Alliance. Retrieved February 11, 2021 from

https://static1.squarespace.com/static/5d3799de845604000199cd24/t/5d79c383e904c741c9e9cd86/1568260995760/CARE+Principles+for+Indigenous+Data+Governance_FINAL_Sept+06+2019.pdf

Simmons University Library. (2021). Research Data Management: File Organization. Retrieved September 29, 2021 from <https://simmons.libguides.com/c.php?g=814790&p=5983200>