

Research Data Management @ Queen's

R4R@Q Webinar Series | March 2nd, 2022 | 2-3pm ET

Presented By:

Rebecca Pero, Information & Project Coordinator, VPR Portfolio

Alicia Cappello, RDM Librarian, Queen's University Library

Meghan Goodchild, RDM Systems Librarian, Queen's University Library /
Scholars Portal

Slides shared under CC-BY



Agenda



1. Tri-Agency Research Data Management (RDM) Policy
2. Draft Queen's Institutional RDM Strategy
3. Data Management Plans (DMPs)
4. Data Deposits
5. Questions / Discussion

Poll



- How familiar are you with RDM?
- How familiar are you with the Tri-Agency RDM policy?
- Have you created a Data Management Plan before?
- Have you deposited your data in a repository before?

Introduction

What is Research Data Management (RDM)?

- Processes applied throughout the **lifecycle of a research project** to guide the collection, documentation, storage, sharing, and preservation of research data.



Plan



Create



Process



Analyze



Disseminate



Preserve



Reuse

- RDM practices are **integral to conducting responsible research** and can help you save resources by **ensuring your data are complete, understandable, and secure.**

Tri-Agency RDM Policy



Tri-Agencies jointly released their [policy on Research Data Management](#) in early 2021.

This policy contains three 'pillars' or requirements that will assist researchers to:

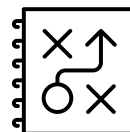
- responsibly and securely manage the research data they've collected via public funds, and;
- where possible, share that data for reuse.

Tri-Agency RDM Policy Pillars



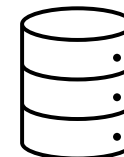
Institutional Strategy

Completed by
March 2023



Data Management Plans

Implementation
begins Spring 2022



Data deposit

Phased
implementation
beginning March
2023



Scholars Portal **Dataverse**

Pillar 1: Institutional Strategies

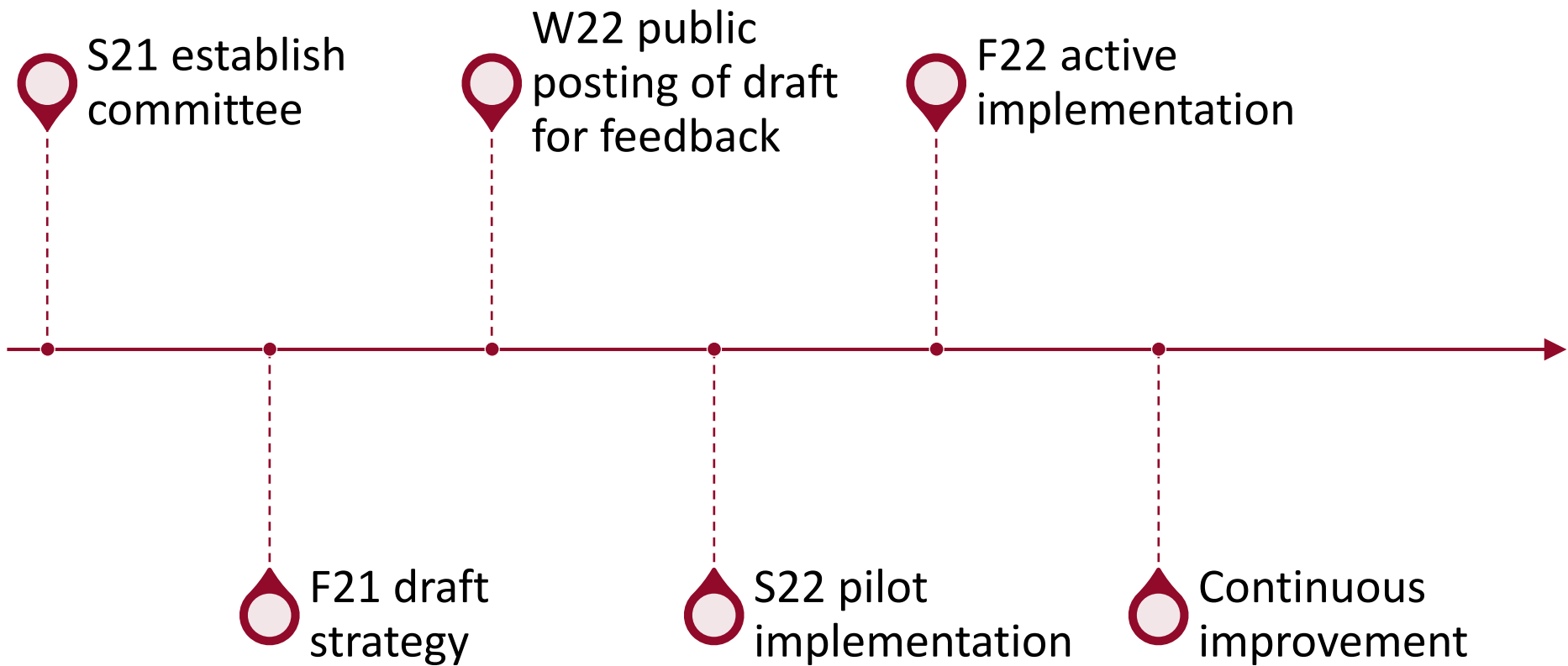
Institutional Strategy Requirements



- recognizing data as an important research output
- supporting researchers to establish and implement data management practices that are consistent with ethical, legal and commercial obligations, as well as tri-agency requirements and other relevant policies;
- promoting the importance of data management to researchers, staff and students;
- guiding their researchers on how to properly manage data in accordance with the principles outlined in the [Tri-Agency Statement of Principles on Digital Data Management](#), including the development of data management plans;
- committing to the adoption of established best practices when developing institutional standards and policies for data management plans;
- providing, or supporting access to, repository services or other platforms that securely preserve, curate and provide appropriate access to research data;
- recognizing that data created in the context of research by and with First Nations, Métis, and Inuit communities and organizations will be managed according to principles developed and approved by those communities and organizations, and in partnership with them;
- recognizing that a distinctions-based approach is needed to ensure that the unique rights, interests and circumstances of the First Nations, Métis, and Inuit are acknowledged, affirmed, and implemented.

Source: [Tri-Agency RDM Policy](#)

Queen's Institutional RDM Strategy



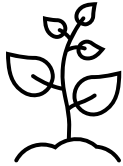
Shared Vision for RDM at Queen's



"Queen's University, we will provide a range of clear and accessible tools, technologies and support services to meet the needs of researchers throughout the research data life cycle. Our research-centered approach means we will promote wise practices around data management planning and data deposit, preservation and disposal. Our aim is to identify pathways to research data management, to build on areas of strength and expertise and improve on tools, technologies and service supports that allow Queen's to enhance research excellence and adapt to the research ecosystem."

Pillar 2: Data Management Plans

Data Management Plans – Tri-Agency Policy



Slated to begin in **spring 2022**, possibly with Partnership Grants.

Must reflect RDM **best practices**.



DMPs are considered **living** documents.

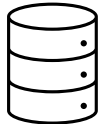
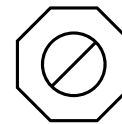
Must include:



HOW: collection, documentation, formatting, protection, preservation



WHAT: existing or new data, constraints



WHERE: shared and deposited

WHO: roles and responsibilities



Inclusion of Indigenous data governance principles and collaboration.

Data Management Plans - Lifecycle



What is a **Data Management Plan** or DMP?

- It's a formal, but living*, document that outlines the strategies and tools being used to manage data during the active phases of research, AND
- The mechanisms to be used to preserve and share research data once the project is completed.

The DMP is considered a **living** document because it can, and should, be modified throughout the life a project when changes occur. It is **not** set in stone.

Data Management Plans – Plan Ahead



Earl: *See you later Chang, we gotta schedule to keep.*

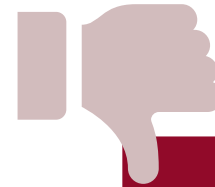
Val: *Oh yeah...see, we plan ahead, that way we don't do anything right now. Earl explained it to me.*

Data Management Plans – Benefits / Considerations



BENEFITS

- Compliance with policy
- Proactive / Forward-thinking
- Identify opportunities & challenges
- Easily adapt to unanticipated obstacles
- Engage partners
- Improved efficiency / increased productivity
- Everyone is on the same page



CONSIDERATIONS

- Time consuming up front (but saves time later)
- Review of DMP throughout project
- “Administrative” exercise
- No guarantee of good data management
- Not required after funding

Data Management Plans - Components



What are the components of a **Data Management Plan**?

1. Data collection
2. Documentation and metadata
3. Storage and backup
4. Preservation
5. Sharing and reuse
6. Responsibilities and resources
7. Ethical and legal compliance

1. Data Collection

- What types of data will you collect, create, link to, acquire, and/or record?
- What file formats will your data be collected in? Will these formats allow for reuse, sharing, and long-term access to the data?
- What conventions will you use to structure, name, and version your files?

2. Documentation & Metadata

- What documentation will be needed for the data to be understood in the future?
- How will you make sure the documentation is created consistently throughout the project?
- What metadata standard and/or tools are you using to document and describe your data?

3. Storage & Backup

- What are your anticipated storage requirements in terms of space?
- How long will you need to store this data?
- Where will you be storing data during your project?
- How are you going to back-up your research data during your project?
- How will project team members and collaborators access, modify, and contribute data to the project?


















4. Preservation

- Where do you plan to deposit your data for long-term preservation once your project is completed?
- How will you ensure your data will be accessible for the long-term (e.g., file formats, file integrity, anonymization/de-identification, documentation, etc.)?

Data Management Plans - Components

5. Sharing & Reuse

- What data will you be sharing and in what form?
- What end-user license do you want to use for your data?
- What steps will be taken to help the research community know that your data exists?

| CREATIVE COMMONS LICENSES | |  COPY & PUBLISH |  ATTRIBUTION REQUIRED |  COMMERCIAL USE |  MODIFY & ADAPT |  CHANGE LICENSE |
|---|---------------|---|--|---|--|---|
|  | PUBLIC DOMAIN | ✓ | ✗ | ✓ | ✓ | ✓ |
|  | CC BY | ✓ | ✓ | ✓ | ✓ | ✓ |
|  | CC BY-SA | ✓ | ✓ | ✓ | ✓ | ✗ |
|  | CC BY-ND | ✓ | ✓ | ✓ | ✗ | ✗ |
|  | CC BY-NC | ✓ | ✓ | ✗ | ✓ | ✓ |
|  | CC BY-NC-SA | ✓ | ✓ | ✗ | ✓ | ✗ |
|  | CC BY-NC-ND | ✓ | ✓ | ✗ | ✗ | ✗ |
|  You can redistribute (copy, publish, display, communicate, etc.) | |  You have to attribute the original work | |  You can use the work commercially |  You can modify and adapt the original work |  You can choose license type for your adaptations of the work. |

6. Responsibilities & Resources

- Who will be responsible for managing the project's data during the project, and after the project is finished?
- What major data management tasks is this person responsible for?
- How will you transfer these responsibilities if members of your project team change?
- What resources will you need in order to implement your DMP?
- What are the estimated costs of your DMP activities?



7. Ethical & Legal Compliance

- Does your project include sensitive data? If so, how will you securely manage that data and only provide access to approved team members?
- How will you address secondary uses of sensitive data (if applicable)?
- Do any other legal, ethical, or IP items need to be considered for your data?

Data Management Plans - Tools



DMP Assistant – Canadian tool:

- Free and easy to use
- Uses Canadian terminology
- Maintained by University of Alberta
- Can share with others
- Export options
- Can access anywhere
- Public DMPs & templates



Data Management Plans – NEW!



Digital Research Alliance of Canada's Data Management Planning Expert Group Seeking Community Assistance with DMP Exemplars

[Announcement](#)

- Accepting Expressions of Interest for creation of a range of disciplinary and methodology specific DMP Exemplars.
- Collaborate with DMPEG; for DMP Assistant.
- Especially interested in the following:
 - Exemplars from Health and Bioscience, and Fine Arts
 - Exemplars using more than one discipline and/or methodology
 - Exemplars for historical or fictional projects

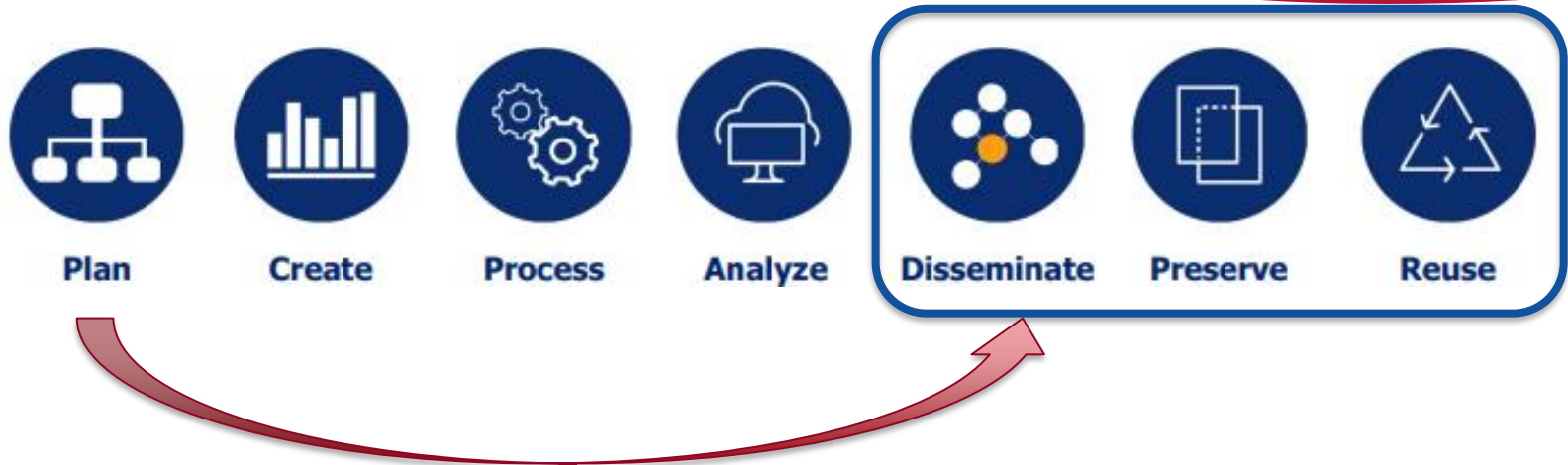
[Apply](#) by March 18th

Notified by April 6th

Exemplars due May 27th, posted by July 29th

Pillar 3: Data Deposit

Data Deposit



“Research data are **valuable assets**, which when properly managed, have the potential to be **reused** and recombined in innovative ways to derive greater value and advance research and scholarship.”

[Current State of Research Data Management in Canada](#) (2020)

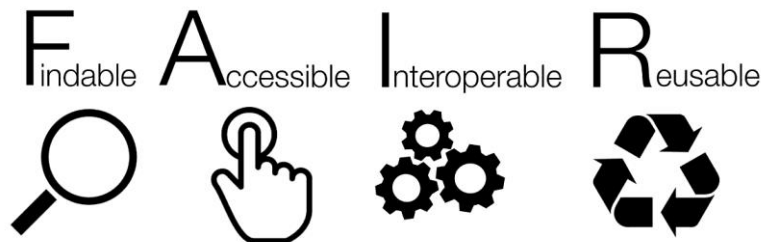
Data Deposit – RDM Policy



[Tri-Agency RDM Policy](#) (2021)

- “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.”
- Expectation that researchers to provide appropriate access to the data where **ethical, cultural, legal and commercial requirements allow**, and in accordance with the FAIR principles and the standards of their disciplines.

Data Deposit – Why deposit in a repository?



<https://www.go-fair.org/fair-principles/>

Findable

- Digital Object Identifier (DOI)
- Indexed in a searchable resource

Accessible

- Ensure controlled access, where appropriate

Interoperable

- Integrate with other data (metadata standards)

Reusable

- Clear and accessible data usage license
- Data are well-described (metadata)



Repositories
also offer
secure
storage and
long-term
stewardship

Data Deposit – Why deposit in a repository?

Citation Advantage



PLOS ONE advanced search

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

The citation advantage of linking publications to research data

Giovanni Colavizza, Iain Hrynaskiewicz, Isla Staden, Kirstie Whitaker, Barbara McGillivray

Published: April 22, 2020 • <https://doi.org/10.1371/journal.pone.0230416>

| Article | Authors | Metrics | Comments | Media Coverage | Peer Review |
|----------|-------------|-------------|-----------|----------------|-------------|
| 110 Save | 45 Citation | 11,850 View | 968 Share | | |

Download PDF Print Share

Check for updates

ADVERTISEMENT

How to Edit Your Work

Abstract

Introduction
Materials and methods
Results
Discussion
Conclusion
Data and code availability
Acknowledgments
References

Efforts to make research results open and reproducible are increasingly reflected by journal policies encouraging or mandating authors to provide data availability statements. As a consequence of this, there has been a strong uptake of data availability statements in recent literature. Nevertheless, it is still unclear what proportion of these statements actually contain well-formed links to data, for example via a URL or permanent identifier, and if there is an added value in providing such links. We consider 531, 889 journal articles published by PLOS and BMC, develop an automatic system for labelling their data availability statements according to four categories based on their content and the type of data availability they display, and finally analyze the citation advantage of different statement categories via regression. We find that, following mandated publisher policies, data availability statements become very common. In 2018 93.7% of 21,793 PLOS articles and 88.2% of 31,956 BMC articles had data availability statements. Data availability statements containing a link to data in a repository—rather than

“[A]n association between articles that include statements that link to data in a repository and up to 25.36% ($\pm 1.07\%$) higher citation impact on average, using a citation prediction model.”

Source: [Colavizza et al. \(2020\)](#)

Data Deposit – Which repository?

Disciplinary Repository

- Built to handle specialized datasets
- Storage likely outside of Canada
- Eligibility, pricing, repository functionality vary
- May only accept certain file types

Scholars Portal Dataverse

- Generalist
- Canadian storage
- File size <3 GB
- Open to researchers at participating Canadian institutions
- File-level restrictions possible
- Supports versioning

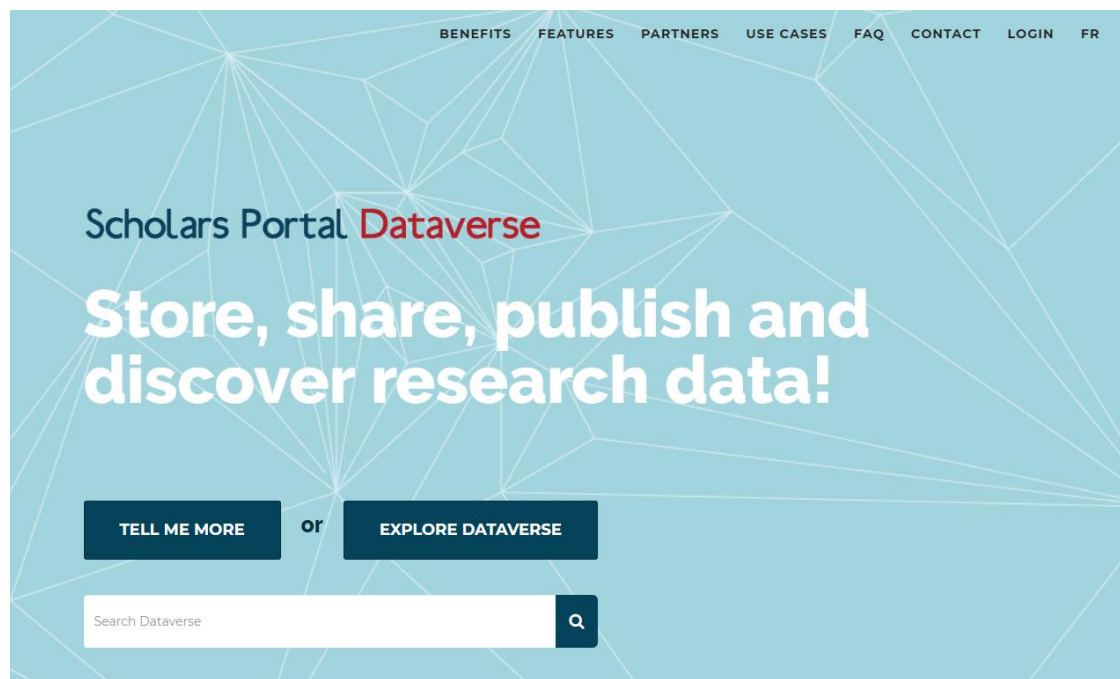
Federated Research Data Repository (FRDR)

- Generalist
- Canadian storage
- Big data support
- Open to faculty at Canadian institutions
- Open data only
- No versioning

Data Deposit – What is Scholars Portal Dataverse?



- Open-source repository platform to store, share, publish and discover research data
- Hosted by Scholars Portal on behalf of 59 subscribing institutions across Canada
- Data stored in the Ontario Library Research Cloud (OLRC)
- Bilingual platform
- Low barrier to access
- Queen's Dataverse Collection managed by your library!



Production: <https://dataverse.scholarsportal.info/>

Demo: <https://demodv.scholarsportal.info/>

A Comprehensive Global Database of Tailings Flows

Version 2.1



Rana, Nahyan M.; Ghahramani, Negar; Evans, Stephen G.; McDougall, Scott; Small, Andy; Take, W. Andy, 2021, "A Comprehensive Global Database of Tailings Flows", <https://doi.org/10.5683/SP2/NXMTX1>, Scholars Portal Dataverse, V2, UNF:6:Dy2b6QrNf1CeZOGEXI961g== [fileUNF]

[Cite Dataset](#) ▾[Learn about Data Citation Standards.](#)[Access Dataset](#) ▾[Contact Owner](#)[Share](#)[Dataset Metrics](#) ⓘ

3,944 Downloads ⓘ

Description

 ⓘ

Tailings are finely ground waste rock produced as a by-product of standard mining projects as well as some industrial and power plant operations. Tailings are conventionally impounded behind a dam that is raised perpetually insofar as operations continue, thus amassing large volumes of materials (sometimes including supernatant pond water) in the process. The failures of some tailings impoundments have triggered downstream mass movements that have caused human, economic and environmental impacts, thus inviting considerable public attention and scrutiny. Developing a detailed inventory of these "tailings flows" facilitates a better understanding of the magnitude-frequency statistics, preconditioning and trigger variables, breach-outflow processes and downstream runout behaviour. Upon screening over 350 historical waste impoundment failure incidents in pre-existing secondary datasets, we have developed a comprehensive global database of 83 tailings flows from

[Read full Description](#) [+]

Subject

 ⓘ

Engineering; Earth and Environmental Sciences

Keyword

 ⓘ

Tailings dam breaches, tailings flows, runout behaviour, liquefaction, mine waste, mass movements

Related Publication

 ⓘ

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., and Take, W.A., 2021. Catastrophic Mass Flows Resulting from Tailings Impoundment Failures. *Engineering Geology*, 106262. <https://doi.org/10.1016/j.enggeo.2021.106262>.

[Files](#)[Metadata](#)[Terms](#)[Versions](#)

Search this dataset...



Filter by

File Type: All ▾

Access: All ▾

File Tag: All ▾

[Sort](#) ▾

1 to 10 of 37 Files

[Download](#) ▾[00_README_RanaEtAl_June2021_Dataverse.pdf](#)

Adobe PDF - 89.1 KB

Published Jun 30, 2021

130 Downloads

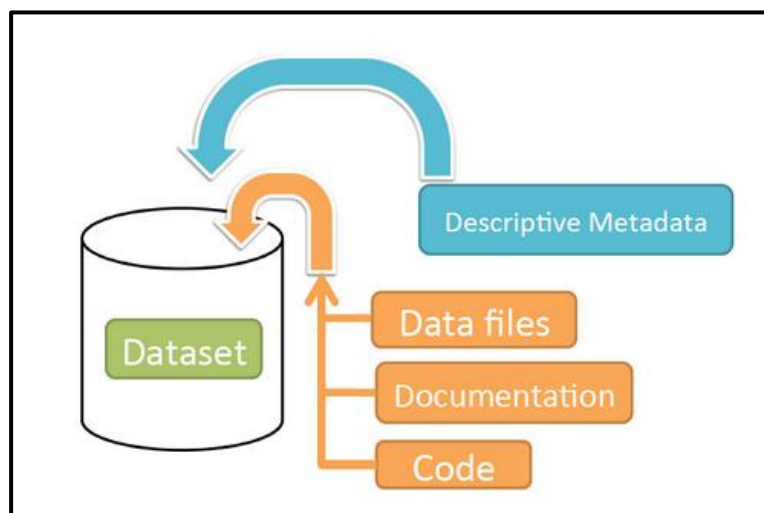
MD5: 790...eb7

[Documentation](#)[1965-Bellavista-Chile-Corona.jpg](#)

JPEG Image - 5.4 MB

Published Jun 30, 2021

124 Downloads



Data Deposit – What is Scholars Portal Dataverse?



Data Citation

- Automatic DataCite Canada DOI minting
- DOI used in standard data citations
- Cross-reference research outputs
 - Establish unbreakable link between scholarly output and associated data

Version 1.0



Hung, Jacqueline, 2021, "Soil nitrogen dynamics and carbon dioxide exchange in a High Arctic wetland", <https://doi.org/10.5683/SP2/V38NLF>, Scholars Portal Dataverse, V1, UNF:6:6Co+33HOyAx3Nwn6EQSN2w== [fileUNF]

Cite Dataset ▼

[Learn about Data Citation Standards.](#)

Data Deposit – What is Scholars Portal Dataverse?



Files and metadata

- All file formats are accepted
- Web browser and API data upload (up to 3 GB per file)
- Mandatory citation metadata fields; Optional disciplinary metadata for Social Sciences, Life Science, Geospatial, Astronomy

Filter by
File Type: All Access: All

Sort

| | | |
|--------------------------|---|---|
| <input type="checkbox"/> | 1 to 10 of 133 Files | <input type="button" value="Download"/> |
| <input type="checkbox"/> |  0_readme.txt Plain Text - 3.1 KB Published Jan 13, 2022 1 Download MD5: 55b...d50 |   |
| <input type="checkbox"/> |  aggradation_dataframe XES10_Data/ Unknown - 28.5 MB Published Jan 13, 2022 1 Download MD5: 122...1cf |  |
| <input type="checkbox"/> |  aggradation_mmsec.npy XES10_Data/ Unknown - 961.1 KB Published Jan 13, 2022 1 Download MD5: 6c4...fac |  |

Data Deposit – What is Scholars Portal Dataverse?



Licensing

- Templates for Creative Commons Licenses
- Default to CCO (open data)

OR

- Custom Terms of Use through “Data Usage Agreement”

Dataset Template ?

Changing the template will clear any fields you may have entered data into.

*Asterisks indicate required fields

Citation Metadata ^

Title * ?

Default CC0 Public Domain (CC0 1.0)

Default CC0 Public Domain (CC0 1.0)

CC Attribution 4.0 International (CC BY 4.0)

CC Attribution-Non-Commercial 4.0 International (CC BY-NC 4.0)

CC Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

CC Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

CC Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)


Custom License (See "Terms of Use" after saving)

Data Deposit – What is Scholars Portal Dataverse?





Dataset versioning


- Automatically track file and metadata changes

 **Canadian Media Concentration Research Project Dataverse** (Carleton University) [CMCRP Project Site](#)

Scholars Portal Dataverse > Carleton University Dataverse > Canadian Media Concentration Research Project Dataverse >
The Growth of the Network Media Economy in Canada, 1984-2016

 Metrics 3 Downloads [Contact](#) [Share](#)

 **The Growth of the Network Media Economy in Canada, 1984-2016** Version 2.0

Winseck, Dwayne, 2017, "The Growth of the Network Media Economy in Canada, 1984-2016",
[doi:10.5683/SP/I5ZHPR](#), Scholars Portal Dataverse, V2  Cite Dataset ▼

[Learn about Data Citation Standards.](#)

[Files](#) [Metadata](#) [Terms](#) [Versions](#)

| Dataset | Summary | Contributors | Published |
|---------|---|--|-------------------|
| 2.0 | Files (Added: 2); View Details | MacOdrum Library Cataloguing and Collections | December 13, 2017 |
| 1.0 | This is the first published version. | MacOdrum Library Cataloguing and Collections, MacOdrum Library | December 7, 2017 |

Data Deposit – What is Scholars Portal Dataverse?



Tools and integrations

- Data Explorer Tool
- File previewers
- Integrations:
 - Open Science Framework (OSF)
 - RSpace



Data Deposit – How do I find research data?



Search for data in the national discovery tool and repository

- Federated Research Data Repository (FRDR)
- Consult FRDR for an ever-expanding source of Canadian research data

A screenshot of the FRDR/DFDR website. The header features the FRDR and DFDR logos with the text "FEDERATED RESEARCH DATA REPOSITORY" and "DÉPÔT FÉDÉRAL DE DONNÉES DE RECHERCHE". Below the header is a large banner with the text "Find and Share Canadian Research Data". A search bar is located below the banner, with a "Deposit Data" button to its right. The main content area is divided into two columns: "Find Data" and "Deposit Data". The "Find Data" column contains text about searching for research datasets and a "Learn more" button. The "Deposit Data" column contains text about depositing data into FRDR and a "Learn more" button. The footer includes a "Privacy Policy" link, copyright information, and logos for Portage, CARL ABRC, and the Globus project.

FRDR DFDR
FEDERATED RESEARCH DATA REPOSITORY DÉPÔT FÉDÉRAL DE DONNÉES DE RECHERCHE

Find and Share Canadian Research Data

Search Deposit Data

Advanced search

Find Data

Search FRDR to find research datasets originating from researchers affiliated with Canadian institutions. Data deposited to other repositories across Canada can also be found by searching in FRDR. View the growing [list of collaborating repositories](#).

[Learn more >](#)

Deposit Data

Any researcher affiliated with a Canadian institution can deposit data into FRDR. The platform can efficiently ingest datasets of any size, and preservation processing is done automatically. Data professionals from the Portage Network and institutions across Canada work with researchers to curate and approve deposited items.

[Learn more >](#)

Privacy Policy
© 2017/2016 Canadian Association of Research Libraries & Compute Canada

portage CARL ABRC powered by globus

<https://www.frdr-dfdr.ca/>

Data Deposit - Resources



Tools and infrastructure

- re3data.org
- [Scholars Portal Dataverse](#) repository
- [Scholars Portal Demo Dataverse](#) - Try it out!
- [Federated Research Data Repository \(FRDR\)](#)

Resources

- Portage Research Data Repositories 101 - [Module 1](#)
- [Queen's Dataverse Deposit Guidelines](#)
- [FAIR principles](#)
- Portage Network [Sensitive Data Toolkit](#) (3 parts)

Conclusion

Key Takeaways



Institutional Strategy

- Feedback on Queen's draft RDM Strategy will be requested Spring 2022

Support for RDM at Queen's

- Data Management Plans
- Data Deposit
- Training, workshops, and consultations
 - [VP-R office](#)
 - [Queen's Library](#)
 - [Centre for Advanced Computing](#)
 - ...and more

RDM Resources



Policies

- Tri-Agency [Research Data Management Policy](#) (2021)
- Tri-Agency [RDM Policy FAQs](#)
- Tri-Agency [Statement of Principles on Digital Data Management](#) (2016)

Tools

- [DMP Assistant](#)
- [re3data.org](#)
- [Scholars Portal Dataverse](#)
- [Scholars Portal Demo Dataverse](#) – Try it out!
- [FRDR](#)

Resources

- Queen's Library RDM [Guide](#)
- [FAIR principles](#)
- The First Nations [Principles of OCAP](#)[?]
- The [CARE Principles](#) for Indigenous Data Governance
- [CARL YouTube Channel](#)

Summary resource document:
<https://guides.library.queensu.ca/rdm/>

Thank You for Attending!



Please feel free to contact any of us if you have additional questions:

Rebecca Pero, Information & Project Coordinator – pero.rebecca@queensu.ca

Alicia Cappello, RDM Librarian – alicia.cappello@queensu.ca

Meghan Goodchild, RDM Systems Librarian – meghan.goodchild@queensu.ca



Queen's
UNIVERSITY

Questions?