



Research Data Management: What Is It & What's Happening In Ontario

OLA - January 30, 2019

Jane Fry
Meghan Goodchild



Agenda

- What is RDM?
- Where to find resources?
- OCUL RDM Needs Assessment Study
- Case studies

What is RDM?
Where to find
resources?

Why the Library?

- Research partner
 - Information
 - Consultation
- RDM expert
- Partner with the Research Office (in some unis)
- The scholarly lifecycle
- Discipline-agnostic
- Challenge!
 - How to help researchers advance their research

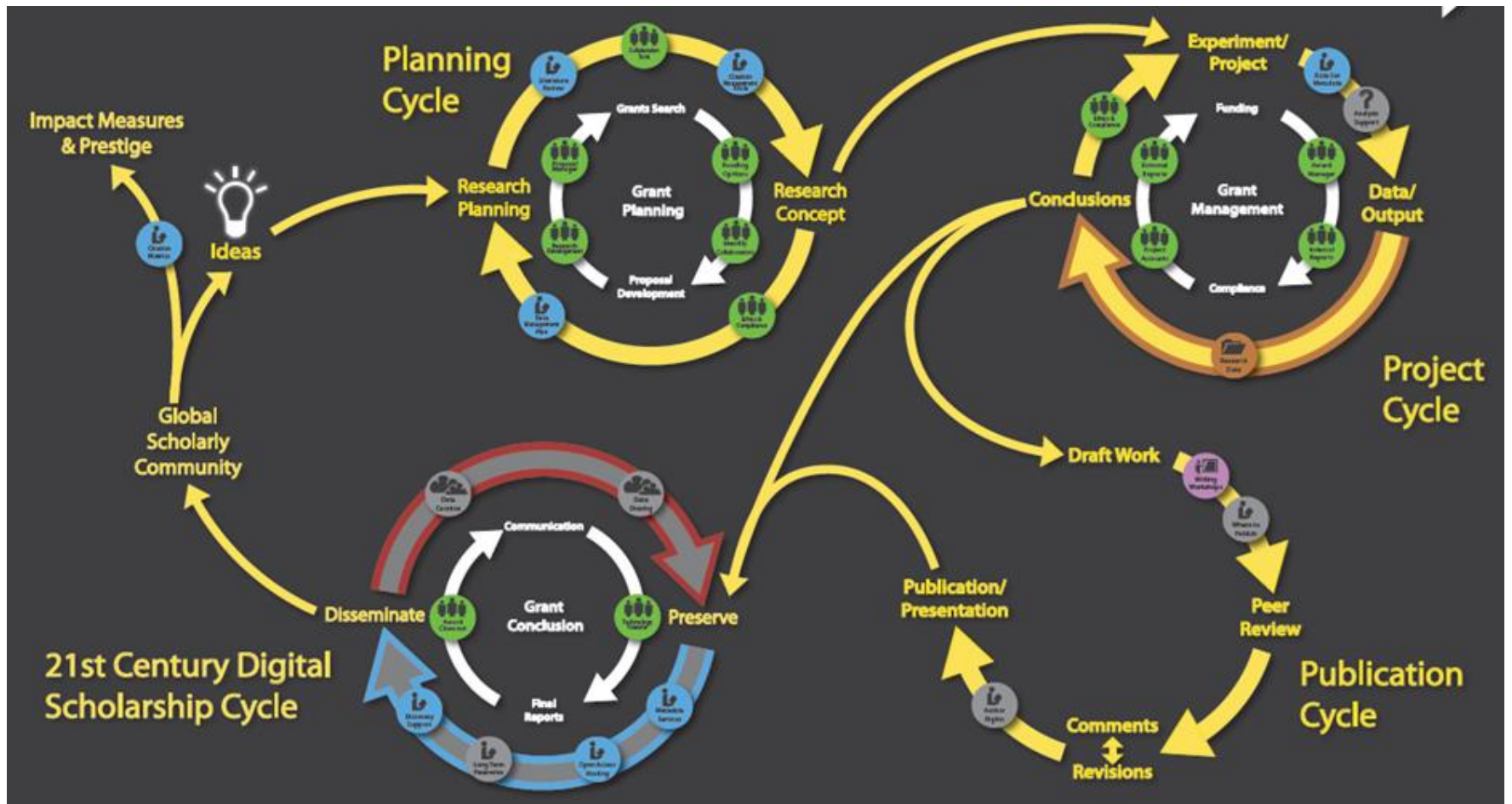
What are Research data?

“Research data can be defined as “a collection of facts, measurements, or observations used to make inferences about the world we live in” (Eisner and Vasgird, n.d.). In practice, research data can be a great many things, from DNA samples to interview transcripts to photographs.”

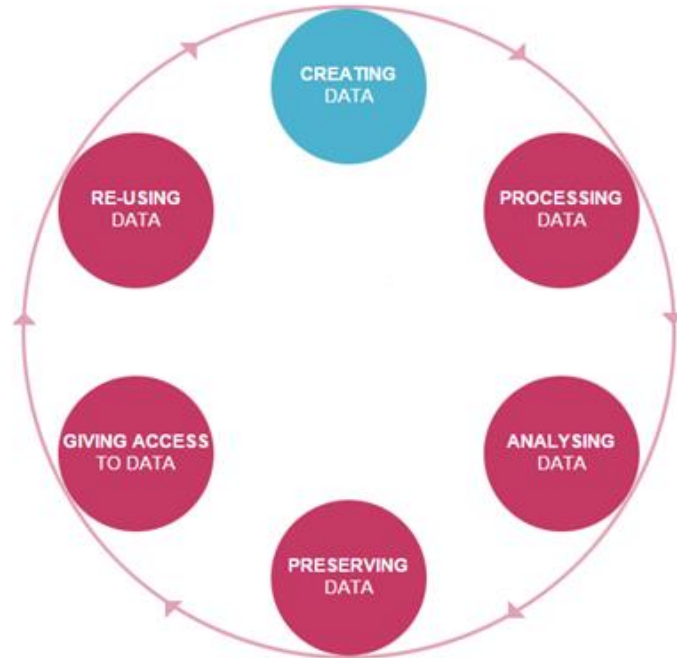
What are Research data?

“Research data are the original sources or material that you have created or collated to conduct your research project. They can be digital or non-digital. The response to your research question is based on the analysis of these research data.”

Source: <https://blogs.ucl.ac.uk/rdm/2015/09/what-is-research-data/>



The Research lifecycle



Source: <http://www.data-archive.ac.uk/create-manage/life-cycle>

What is RDM?

- The process of controlling the information generated during a research project
- An integral part of the research process
- Can be challenging
- Depends on
 - Types of data involved
 - How data is collected and stored
 - How it is used
- Occurs throughout the research lifecycle

Reference: <https://up-za.libguides.com/c.php?g=356288&p=2428309>

What is RDM? (cont'd)

- Includes
 - Sound practices
 - Data curation
 - Data stewardship

Why is RDM important?

- Becoming a requirement for funders
- Researchers are tending to collaborate more than ever before with international partners
- Will benefit the researcher
- Deterrent to research fraud
- Bonus ...

How do I find resources?

portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Home Contact Us Français

Training Resources Communities of Practice How to Manage Your Data Infrastructure Platforms Work with Portage About Portage

THE PORTAGE NETWORK is dedicated to the shared stewardship of research data in Canada through:

- Developing a national research data culture
- Fostering a community of practice for research data
- Building national research data services and infrastructure

Launched in 2015 by the [Canadian Association of Research Libraries](#), Portage works within the library community to coordinate expertise, services, and technology in research data management.

RDM PRIMER
For RESEARCHERS

DMP Assistant
Sign In Sign Up
If you have an existing account with DMP Assistant or previous version of DMP Builder. New to DMP Assistant? Sign up today.
Terms of Use

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

NEWS

How do I find resources?

portage
SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

Accueil Contactez-nous English

Ressources de formation Communautés de pratique Comment gérer vos données Plateformes d'infrastructure Collaborer avec Portage À propos de Portage

LE RÉSEAU PORTAGE se consacre à l'intendance partagée des données de recherche au Canada en favorisant :

- le développement d'une culture de gestion des données de recherche;
- l'établissement d'une communauté de pratique pour la gestion des données de recherche;
- la création de services et d'une infrastructure de gestion des données de recherche à l'échelle nationale.

GDR
Informations de base

Pour LES CHERCHEURS

11010010001

Assistant PGD

Se connecter
Si vous avez déjà un compte avec l'Assistant PGD ou avec une version antérieure de l'Assistant PGD.

S'inscrire
Nouvel utilisateur de l'Assistant PGD? Inscrivez-vous aujourd'hui.

FRDR
DFDR
FEDERATED RESEARCH DATA REPOSITORY
PÔT FÉDÉRÉ DE DONNÉES DE RECHERCHE

RDM equivalents in French and English

Search: <input type="text"/>	
Term related to RDM (english)	Expression liée à la GDR (français)
access [verb, to access data]	consulter
aggregate	regrouper, rassembler
aggregation	service regroupé
API	API
application program interface (API)	interface de programmation d'application [utiliser surtout API]
code book	dictionnaire de données

"Good Enough" Research Data Management

(a brief guide for busy people)

This brief guide presents a set of good data management practices that researchers can adopt, regardless of their data management skills and levels of expertise.

1

Save your raw data in original format

- 1.1 Don't overwrite your original data with a cleaned version.
- 1.2 Protect your original data by locking them or making them read-only.
- 1.3 Refer to this original data if things go wrong (as they often do).

2

Backup your data

- 2.1 Use the 3-2-1 rule: Save three copies of your data, on two different storage mediums, and one copy off site.
- 2.2 Do not backup or store sensitive data on a commercial cloud (Dropbox, Google Drive, etc.).

3

Describe your data

- 3.1 **Machine Friendly:** Describe your dataset with a metadata standard for discovery.
- 3.2 **Human Friendly:** Describe your variables, so your colleagues will understand what you meant. Data without good metadata is useless. Give your variables clear names.
- 3.3 Do not leave cells blank - use numeric values clearly out of range to define missing (e.g. "99999") or not applicable (e.g. "88888") data, and describe these in your data dictionary.
- 3.4 Convert your data to open, non-proprietary formats.
- 3.5 Name your files well with basic metadata in file names.

4

Process your data

- 4.1 Make each column a variable.
- 4.2 Make each row an observation.
- 4.3 Store units (e.g. kg or cm) as metadata (in their own column).
- 4.4 Document each step processing your data in a README file.

5

Archive and preserve your data

- 5.1 Submit final data files to a repository assigning a persistent identifier (e.g. handles or DOIs).
- 5.2 Provide good metadata for your study so others could find it (use your discipline's metadata standard, e.g. Darwin Core, DOI, etc.).

portage

This guide was developed by the Portage Training Dept. Copyright permission from the original creator, Eugene Searles, is granted by the University of British Columbia. It can be modified, printed, used freely under the CC BY license.



De bonnes pratiques en gestion des données de recherche

(Un guide sommaire pour gens occupés)

Ce petit guide présente un ensemble de bonnes pratiques que les chercheurs peuvent adopter, et ce, indépendamment de leurs compétences ou de leur niveau d'expertise.

1

Enregistrer vos données brutes dans leur format original

- 1.1 N'écrivez pas vos données avec une version corrigée.
- 1.2 Protégez vos données originales en verrouillant vos fichiers (avec un mot de passe) ou en définissant un accès en lecture seule.
- 1.3 Référez-vous à ces données originales lorsque les choses tournent mal (comme c'est souvent le cas).

2

Sauvegarder vos données

- 2.1 Utilisez la règle du 3-2-1: enregistrez trois copies de vos données sur deux supports de stockage différents et conservez-en une dans un lieu hors site.
- 2.2 Ne sauvegardez pas ou ne stockez pas des données sensibles dans un nuage commercial (Dropbox, Google Drive, etc.).

3

Décrire vos données

- 3.1 **Pour qu'elles soient facilement lisibles par les ordinateurs:** Décrivez votre ensemble de données selon une norme de métadonnées facilitant la découverte.
- 3.2 **D'une manière conviviale pour les humains:** Décrivez vos variables de façon à ce qu'elles soient facilement compréhensibles pour vos collègues. Des données sans bonnes métadonnées sont inutilisables. Donnez des noms clairs à vos variables.

3.3

- 3.3 Ne laissez pas les cellules vides - utilisez des valeurs numériques clairement hors limites pour définir les données manquantes (p. ex. "99999") ou non applicables (p. ex. "88888"), et décrivez-les dans votre dictionnaire de données.
- 3.4 Convertissez vos données dans des formats ouverts et non-propriétaires.
- 3.5 Nommez bien vos fichiers en utilisant des métadonnées de base dans le nom des fichiers.

4

Traiter vos données

- 4.1 Faites de chaque colonne une variable.
- 4.2 Faites de chaque rangée une observation.
- 4.3 Stockez les unités (tels que kg ou cm) en tant que métadonnées (avec leur propre colonne).
- 4.4 Documentez chaque étape du traitement de vos données dans un fichier « LISEZ-MOI ».

5

Archiver et préserver vos données

- 5.1 Soumettez vos fichiers de données finaux dans un dépôt de données en leur attribuant un identificateur permanent (DOI, ou autres Handles).
- 5.2 Fournissez de bonnes métadonnées à votre étude de manière à ce que les autres puissent la repérer plus facilement (utilisez les normes de métadonnées de votre discipline telles que Darwin Core, DOI, etc.).

portage

This guide was developed by the Portage Training Dept. Copyright permission from the original creator, Eugene Searles, is granted by the University of British Columbia. It can be modified, printed, used freely under the CC BY license.



RESEARCH DATA MANAGEMENT PRIMER

Research Data Management (RDM)

- RDM refers to the processes applied throughout the lifecycle of a research project to guide the collection, documentation, storage, sharing, and preservation of research data.
- RDM practices are integral to conducting responsible research and can help researchers save resources by ensuring their data is complete, secure, understandable, and secure.
- RDM practices also follow institutional and funding agency guidelines that protect their investments.
- The broader research community can derive maximum value from research data that can be accessed, shared, reused and repurposed.

The Research Data Lifecycle



* Life cycle model developed by the Leadership Council for Digital Research Infrastructure. For more information visit <http://lci.leadership.ca>

Defining Research Data

- Primary sources supporting research, scholarship or artistic endeavours
- Can be used as evidence to validate findings and results
- May take the form of experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data
- All other digital and non-digital content have the potential to become research data

*Research data. (n.d.) In CASRA's Dictionary. Retrieved from dictionary.casra.org/Research_data

portage

This guide was produced by the Portage Training Expert Group and can be modified and re-used freely under the CC BY license.

GESTION DES DONNÉES DE RECHERCHE INFORMATION DE BASE

Ressources canadiennes et internationales en GDR



Politiques canadiennes sur la GDR

- Le *Plan d'action du Canada pour un gouvernement ouvert* décrit le développement et l'adoption des politiques, des lignes directrices et des outils pour soutenir la gestion efficace des données scientifiques.
- Les trois organismes subventionnaires du gouvernement fédéral (IRSC, CRSNG et CRSH) ont adopté une *Déclaration de principes sur la gestion des données numériques* qui décrit les attentes et les responsabilités en matière de gestion des données produites grâce aux fonds publics. L'élaboration de plans de gestion de données peut aider à répondre à ces attentes et responsabilités.

Contact institutionnel:

portage

Le présent guide a été produit par le Groupe d'experts sur la formation en GDR de Portage et peut être modifié et réutilisé en toute liberté sous la licence CC BY.

Training Modules

- Research Data Management (RDM) 101
 - 4 parts
 - <https://portagenetwork.ca/training-resources/portage-training-resources/rdm-101/>
 - <https://bit.ly/2U7SluJ>
- CIHR Research Data Management Learning Modules
 - RDM 101 Lite!
 - <http://www.cihr-irsc.gc.ca/lms/e/app-rdm-mod1/>
 - <http://www.cihr-irsc.gc.ca/lms/f/app-rdm-mod1/>

OCUL RDM needs assessment study



What is OCUL?

- Established in 1967
- Consortium of 21 university libraries
- Innovative information services and collections
- “Collaborate, innovate, deliver”



OCUL

Ontario Council of
University Libraries

What is Scholars Portal?

- OCUL initiative established in 2002
- Shared technological infrastructure that provides access to and preserves information resources collected and shared by OCUL
- CRL Trusted Digital Repository (2013-) (ejournals)



Scholars Portal supporting RDM



<odesi> is a web-based data exploration and analysis tool for social survey and polling data



GeoPortal is an award-winning geospatial data discovery tool, providing access to a curated collection of maps and geospatial datasets



Dataverse is a research data repository to store, share, publish and discover research data



Ontario Library Research Cloud is a high capacity, geographically distributed storage and compute network



Permafrost is a hosted digital preservation service providing a suite of tools, training and resources for processing digital objects for long-term preservation

Needs assessment for RDM infrastructure

- Investigate current *strengths* and *gaps* of RDM infrastructure at local, regional, and national levels
- Inform service and platform developments that target needs of OCUL institutions

Needs assessment - Methodology

- Interview questions shared with OCUL institutions
- Participating groups bring together 2-4 people who oversee and/or who are involved with RDM
- Teams review the questions and choose to prepare responses in advance and/or answer at the time of interview
- Semi-structured interview approach
 - Interviewer may ask follow-up questions
 - Team can elaborate on open-ended questions

Needs assessment - Interview questions

General, institutional context

Active data

Data deposit

Digital preservation

RDM and OCUL within Canadian context

Needs assessment - Current status

- 16 interviews completed (October - January)
- 2 more interviews scheduled
- Groups consist of 2-4 people from library
 - e.g., University Librarian, Head/Director of Division/Unit, RDM Librarian, Liaison Librarian, Data Librarian
- Many groups also include people from outside library
 - e.g., CIO, IT, Research Office, VP-Research, Dean/Associate Deans

Needs assessment - Analysis

- Transcription of recordings
- Answers summarized and sent to groups for verification
- Responses coded and anonymized
- Qualitative analysis using *grounded theory*, an ethnographic approach
 - Inductive method to discover emerging themes
 - Iterative process to identify, combine, alter, or remove themes that reflect patterns and common ideas

Needs assessment - Preliminary Trends

- General challenges related to RDM infrastructure include:
 - Storage for active data and for preservation
 - Lack of sustainable resources
 - Dealing with sensitive data and data security
 - Lack of awareness related to RDM and library's role
- All institutions are currently responding in a proactive way to the changing landscape of RDM:
 - In process of developing policies/strategies related to RDM
 - Trying to build services and infrastructure with limited resources
 - Attempting to address needs at local, consortial, and national levels

Needs assessment - Next steps

Analysis!!

Report to OCUL directors

White paper

OCUL RDM symposium

Case studies

Case study 1 - Carleton U

- Training Materials
 - Portage training materials
- Presentations
 - from others - colleagues, unis, webinars, ...
 - When
 - Who
 - Why
 - Where
 - What

Case study 1 - Carleton U - Face to Face

- Consultations
 - When
 - Who
 - Why
 - Where

 - What
- Connections aka Networking
 - When
 - Who
 - Why
 - Where
 - What

Research Data Management Plan Template

PROJECT INFORMATION

Name of Researcher: [Click here to enter text.](#)

Project Title: [Click here to enter text.](#)

Funding Sources (if any): [Click here to enter text.](#)

DATA MANAGEMENT PLAN

- **Data Description:** Provide a brief description of the information to be gathered – the nature, scope, and scale of the data that will be generated or collected.

[Click here to enter text.](#)

- **Metadata Description:** What types of metadata will you produce to support the data? Will a metadata standard be used?

[Click here to enter text.](#)

- **Intellectual/Property Rights:** Who will hold intellectual property rights for the data and other information created by the project?

[Click here to enter text.](#)

- **Ethics and Privacy:** if applicable, how will you handle informed consent with respect to communicating to respondents that the information they provide will remain confidential when data are shared or made available for secondary analysis?

[Click here to enter text.](#)

- **Format:** Specify the anticipated submission, distribution, and preservation formats for the data and related files (note that these formats may be the same).

[Click here to enter text.](#)

- **Archiving and Preservation:** How will you ensure that data are preserved for the long term?

[Click here to enter text.](#)

- **Storage and Backup:** How and where will you store copies of your research files to ensure their safety? How many copies will you maintain and how will you keep them synchronized?

[Click here to enter text.](#)

- **Security:** How will you ensure that the data are secure?

[Click here to enter text.](#)

- **Responsibility:** Who will act as the responsible steward for the data throughout the data life cycle?

[Click here to enter text.](#)

- **Existing Data:** Are there existing data with a focus similar to the data that will be produced? If so, list what they are and explain why it is important to collect new data.

[Click here to enter text.](#)

- **Selection and Retention Periods:** Indicate how data will be selected for archiving, how long the data will be held, and what your plans are for eventual transition or termination of the data collection in the future.

[Click here to enter text.](#)

- **Access and Sharing:** Indicate how you intend to archive and share your data and why you have chosen that particular option. (eg. Institutional Repository like CURVE, self-archiving with deposit after a time period, dedicated website, domain repository)

[Click here to enter text.](#)

- **Audience:** Describe the audience for the data you will produce.

[Click here to enter text.](#)

- **Data Organization:** Indicate how the data will be managed during the project, with information about version control, naming conventions, etc.

[Click here to enter text.](#)

Research Data Management

Home » Services » Research Data Management

What is Research Data Management?

Why do I Need a Research Data Management Plan?

How to Develop a Research Data Management Plan

Share and Archive Your Data in Data Repositories

Advice and Training

Reference: <https://library.carleton.ca/services/research-data-management>

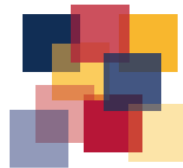
Case study 1 - Carleton U

- Remember
 - This is what I do
 - You are the one walking in your shoes!
 - Don't reinvent the wheel
 - Call me!

Case study 2 - Queen's University Library (QUL)

Team-based approach to RDM—integrated within library support for research

- Data Services (Data Librarian, Data Services Coordinator)
- Head, Open Scholarship Services
- Research Data Management Systems Librarian
- Open Scholarship Services Librarian
- Other Open Scholarship Services librarians
- Liaison librarians



QUEEN'S UNIVERSITY LIBRARY

Case study 2 - QUL - RDM Resources & Services

- Resources on the library website
 - RDM LibGuide
 - RDM Workflow
 - RDM Collection Development Policy
 - Data Deposit Form
 - Links to external resources (e.g., Portage)
- Researcher and research team consultations
- Outreach presentations and workshops



Research Data Management

Queen's University Library provides Research Data Management (RDM) services in order to support researchers in meeting grant requirements, produce a more competitive grant application, increase the impact and visibility of their research, encourage the discovery and use of existing data to explore new research questions, better guarantee the data are accurate, complete, authentic, and reliable, ensure long-term preservation of data for future researchers, and ensure compliance with ethics and privacy policies.

Information about RDM services can be found on the [RDM LibGuide](#).

RDM Resources

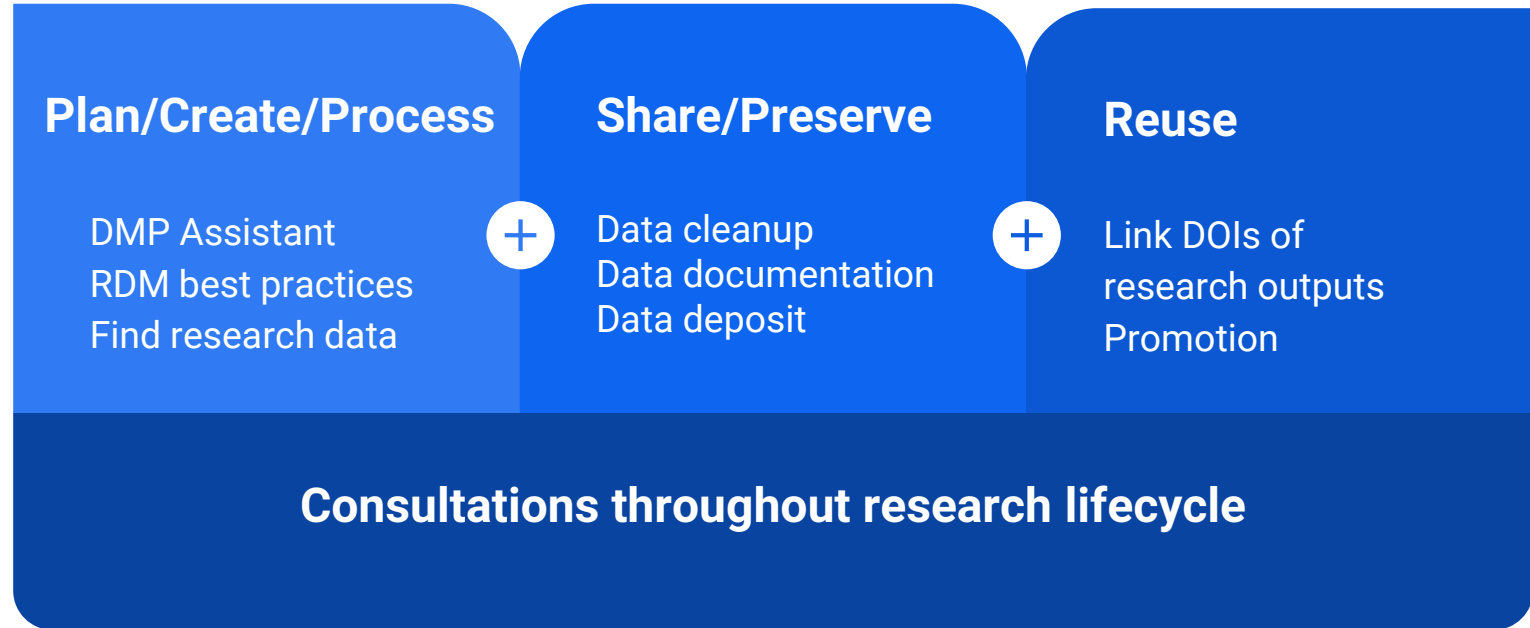
- [Data Deposit Form](#)
- [RDM Collection Development Policy](#)
- [RDM Workflow](#)

Related Links

- [Researcher Training Program](#)
- [Scholarly Publishing](#)
- [Portage Data Management Planning Assistant](#)

<https://library.queensu.ca/help-services/research-data-management>

Case study 2 - QUL - Consultations



Case study 2 - QUL - Outreach



Case study 2 - QUL - New RDM initiatives

- Researcher Training Program (RTP)
- “Train the trainer” sessions for librarians
- Resources for Researchers at Queen’s (R4R@Q)

Researcher Training Program

Welcome to the Queen's University Library researcher training program.

Queen's researchers and scholars generate globally impactful research and scholarship.

The researcher training program gives VOICE to research results according to the following principles: **Value, Openness, Inclusivity, Collaborative platforms, and Engaged researchers (VOICE)**.

The program consists of an expanding set of self-serve training modules that will help you to successfully manage your digital research and scholarship effectively throughout the research lifecycle and at your point of need:

- **Getting Published**
- **Open Publishing Support**
- **ORCID and your Research Profile**
- **Manage your Research Data**



<https://library.queensu.ca/help-services/researcher-training-program>

In Sum ...

We hope you have learned:

- What RDM is
- Where to find helpful resources
- Information about the RDM Needs Assessment study
- How RDM is conducted at Carleton and Queen's (two examples)

Thank you!

Questions????

Comments????

Anything else????

Contact Information

Jane Fry

Data Services Librarian

Carleton University

jane.fry@carleton.ca

Meghan Goodchild

RDM Systems Librarian

Queen's University & Scholars
Portal

meghan.goodchild@queensu.ca