



**Queen's University Library
(QUL)
Academic Services -- Data Services**

Research Data Management

Collection Development Policy

QUL Research Data Working Group

Last reviewed: November 25, 2015

Contents

- Collection Development Policy 2
 - Introduction 2
 - Collection Criteria -- Details 3
 - Complete 3
 - Well-documented 3
 - Open..... 3
 - Ethics-compliant 3
 - Sensitive and responsive to intellectual property arrangements..... 3
 - In Preservation-friendly Formats 3
 - Other Considerations 4
 - Research Data Management Planning and Data Deposit 4
 - Data Repositories 5
 - Data Retention and Preservation 6
 - Appendix 1. Research Data Management Workflow..... 7
 - Appendix 2. Disclosure Control Flowchart..... 8

Collection Development Policy

Introduction

As part of Queen's University Library (QUL) Academic Services, Data Services provides Research Data Management (RDM) support¹ for the Queen's University research community (Queen's faculty, graduate students, and staff). As part of these services, the Library actively solicits researcher-generated data, from all disciplines, for deposit into one or more appropriate data repositories. The service is evolving, and we are currently experimenting with a wide range of data and levels of service. An overview of RDM workflow can be found in **Appendix 1**.

This policy is in keeping with the Government of Canada's "initiative to maximize access to federally-funded scientific research, to encourage greater collaboration and engagement with the scientific community, the private sector, and the public".

From a collection development perspective, our goal is to ensure that researcher-deposited data are:

- **Complete**
- **Well-documented**
- **Ethics-compliant**
- **Sensitive and responsive to intellectual property arrangements**
- **Open**
- **In preservation-friendly formats**
- **Stored in appropriate archives/repositories**

Our services are closely tied to collection development -- currently, we facilitate and advise on:

- **Planning for data management**
 - data management plans – on their own, or as part of research funding proposals
- **Preparing data for deposit**
 - coding, formatting, anonymizing, and more
- **Documenting data**
 - describing data so others can understand and use it
- **Archiving data**
 - finding an appropriate repository or archive
- **Making data discoverable**
 - promoting open access, indexing, assigning Digital Object Identifiers (DOIs)

¹ For more information, see Queen's Research Data Management Guide: <http://guides.library.queensu.ca/rdm>

Collection Criteria -- Details

QUL RDM Services will work with researchers to deposit data that meet the following criteria:

Complete

The QUL Research Data Management Service accepts datasets that are complete and not intended to be updated or revised. Revisions due to corrections/errors are accepted, but as a rule data should be in final, ready-to-use form. The Library can offer suggestions and support in this regard.

Well-documented

Data without good documentation is of limited value. Documentation consists of basic descriptive elements such as title, author, date, etc. but should also provide more detailed information, e.g. **User Guides, Codebooks, Questionnaires, etc.**, that ensure data can be understood and used by future users. Ideally, the process of documenting data files starts at the beginning of the research life cycle. The Library's Research Data Management Service can provide examples of good documentation and advice on how to approach the documentation process.

Open

Open access to data is an underlying principle of the QUL Research Data Management Service. The ultimate goal is to make deposited data available to other researchers. Reasonable **embargo periods** are accepted for *research* and *publication* purposes. Deposited files can be embargoed **fully** (e.g. *data and documentation*) or **partially** (e.g. *data only*), at the researcher's discretion. Once an embargo is established, it is the responsibility of researchers to inform the Library if an extension of an embargo is required. During the embargo, the Library can provide researchers and designates with controlled access to deposited data and/or documentation.

Ethics-compliant

QUL can work with researchers to ensure that data (e.g. *human-subject data*) are **de-identified** and **anonymized**. For more information, see '*Best Practices in Data Management*' (<http://guides.library.queensu.ca/best-practices>). For sensitive data, Data Services takes appropriate steps to ensure the protection of respondent privacy (see **Appendix 2**, for a 'Disclosure Control Flowchart').

Sensitive and responsive to intellectual property arrangements

In cases involving intellectual property, appropriate access restrictions will be established in consultation with the researcher. Contractual and other licensing issues will be taken into consideration.

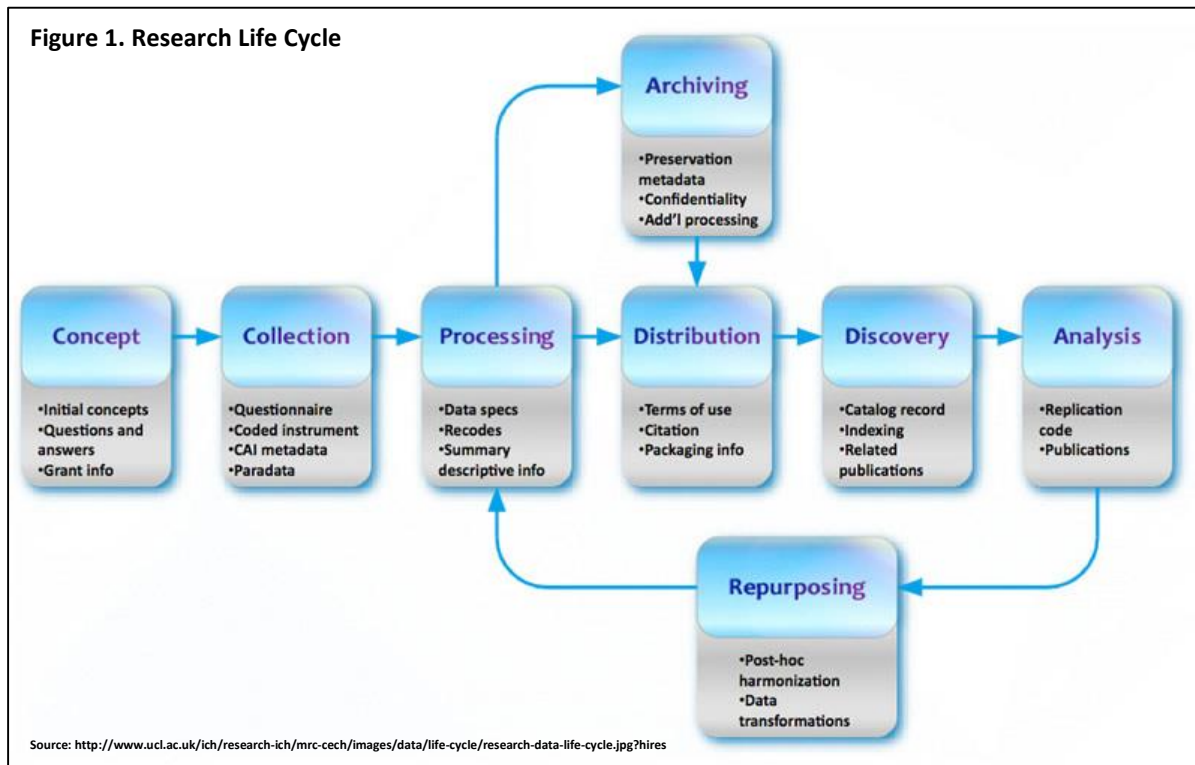
In Preservation-friendly Formats

QUL Research Data Management Service can advise researchers on '*cleaning up*' and formatting their data for preservation and future use. As far as possible, the Library encourages use of data formats that are hardware- and software-agnostic. The Library can provide advice and, in some cases, format-conversion services. Data in proprietary formats (e.g. collected using commercial scientific instruments) are

considered for deposit on a case-by-case basis. Ideally, these data should be converted to a more open and accessible format.

Other Considerations

Research data is part of a broader research life cycle (see **Figure 1**).



Researcher-generated data used in academic, peer-reviewed publications will be strongly considered for deposit. Subject to other collection criteria, data produced by Faculty, Graduate students, and Queen’s research teams, groups, projects, and/or departments will be considered. The Service is growing and evolving, and policies surrounding data selection are expected to change.

While current data holdings tend to be centred in the social sciences (e.g. survey data and statistics), data deposits from other disciplines are encouraged.

Research Data Management Planning and Data Deposit

To facilitate RDM Planning, researchers are encouraged to create a data management plan using the **Portage² DMP Assistant³**. This tool gathers and organizes information about the research

² **Portage** is a national organization, sponsored by the Canadian Association of Research Libraries (CARL), with a mandate to support Canadian innovation through shared expertise and stewardship of research data

³ **DMP Assistant** is a national, bilingual, data management planning tool developed and supported by Portage

project, the data that will be collected, and how the data will be organized, stored, preserved, and shared.

QUL Data Services also provides a convenient online ‘Data Deposit Form’ (<http://library.queensu.ca/data/data-deposit/>) that gathers basic metadata and establishes an embargo date (if applicable) and permission to release data at a specified level of access (restricted, Queen’s-only, Ontario, Canada, World). This form does not accept data as an attachment. Researchers are encouraged to make appropriate data transfer arrangements with QUL Data Services.

Data Repositories

Data deposited through Data Services can find its way into any of a number of data repositories. Use of domain or subject-specific repositories – those used by the majority of researchers in a particular field of study – are encouraged. By virtue of its membership in the Ontario Council of University Libraries (OCUL), QUL has access to two province-wide data repositories: the Scholars Portal **Dataverse** (<http://dataverse.scholarsportal.info/dvn/dv/qlrda>) and the Scholars Portal **ODESI Data Portal** (<http://odesi.ca>). In addition, QUL Data Services makes use of **QSpace** (<https://qspace.library.queensu.ca/handle/1974/7652>), our local institutional repository. These stable repositories offer a variety of features that, together, respond to a wide range of data deposit needs. Other repository options can be explored using the ‘Data Repositories and Archives’ tab in our RDM Guide (<http://guides.library.queensu.ca/repositories-archives>). A summary of repository features is provided in **Table 1**.

Table 1. Summary of Repository Features

	Local Control	Flexible Access Control	Data-specific Metadata	Online Statistical Analysis Tools	Search Features	User-mediated Deposit	Archive Tailored to Discipline
QSpace (Dspace-based Institutional Repository)	yes	no	no	no	basic	no	no
Dataverse (Scholars Portal)	yes	yes	yes	some	advanced	yes	flexible
ODESI Data Portal (Scholars Portal)	no	no	yes	yes	advanced	no	somewhat flexible
Discipline-based Archive	no	unknown	yes	unknown	unknown	in some cases	yes

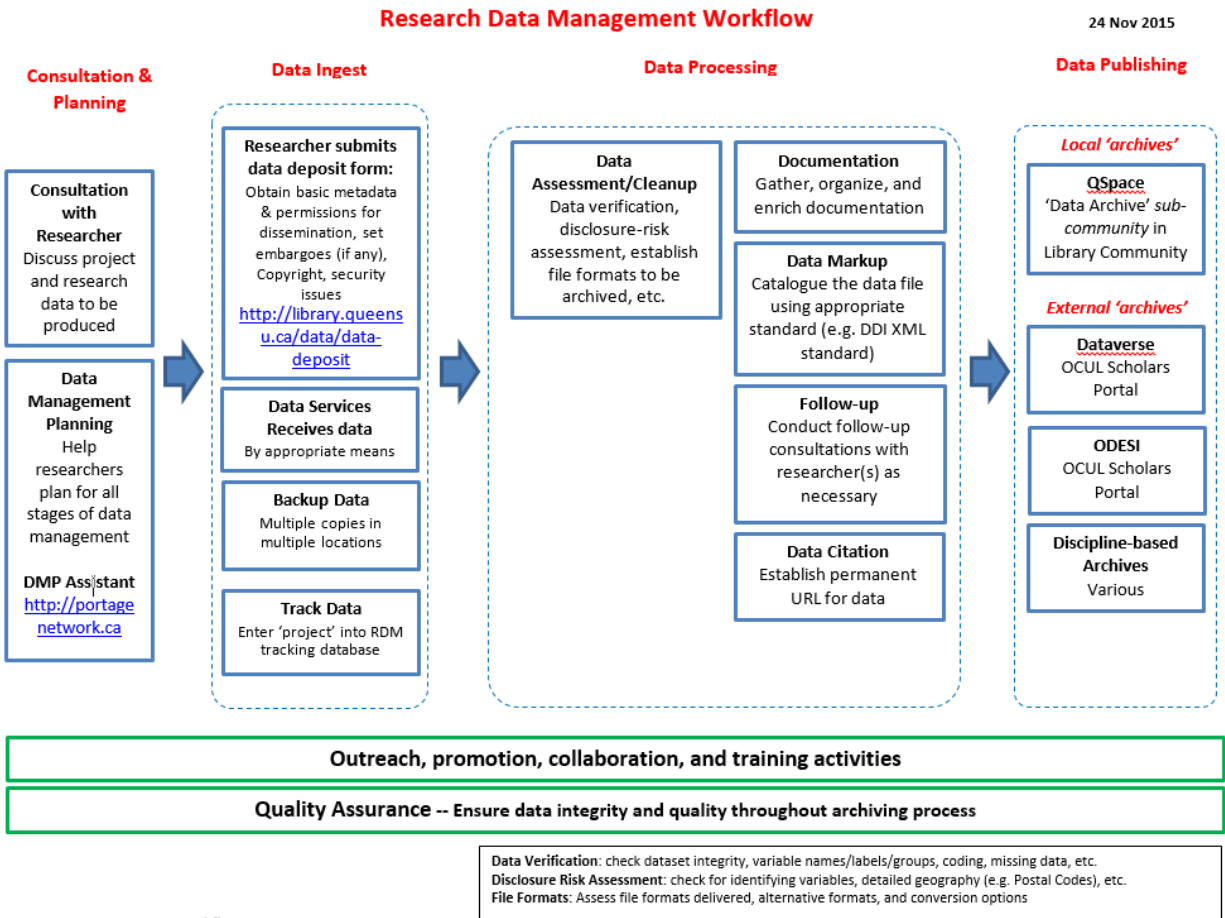
Data Retention and Preservation

As a rule, data that is accepted for deposit will be retained indefinitely, unless there are compelling reasons to remove them. Data Services stores research data in multiple repositories, in keeping with the **LOCKSS** principle (“**L**ots **O**f **C**opies **K**ee**P** **S**tuff **S**afe”) of data preservation. Data archival tools and standards are being explored and implemented in Ontario and Canada. For example, Ontario’s Scholars Portal, and the Portage national data management network, are both looking at the question of long-term preservation of research data. QUL Data Services is actively participating in these initiatives with the intent of taking advantage of one or more of these solutions as they become available.

Behind these initiatives is an Open Archival Information System (OAIS) model. OAIS “*is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information over the long term*”⁴. QUL Data Services is committed to using best practices and standards to ensure long-term preservation and availability of archived data.

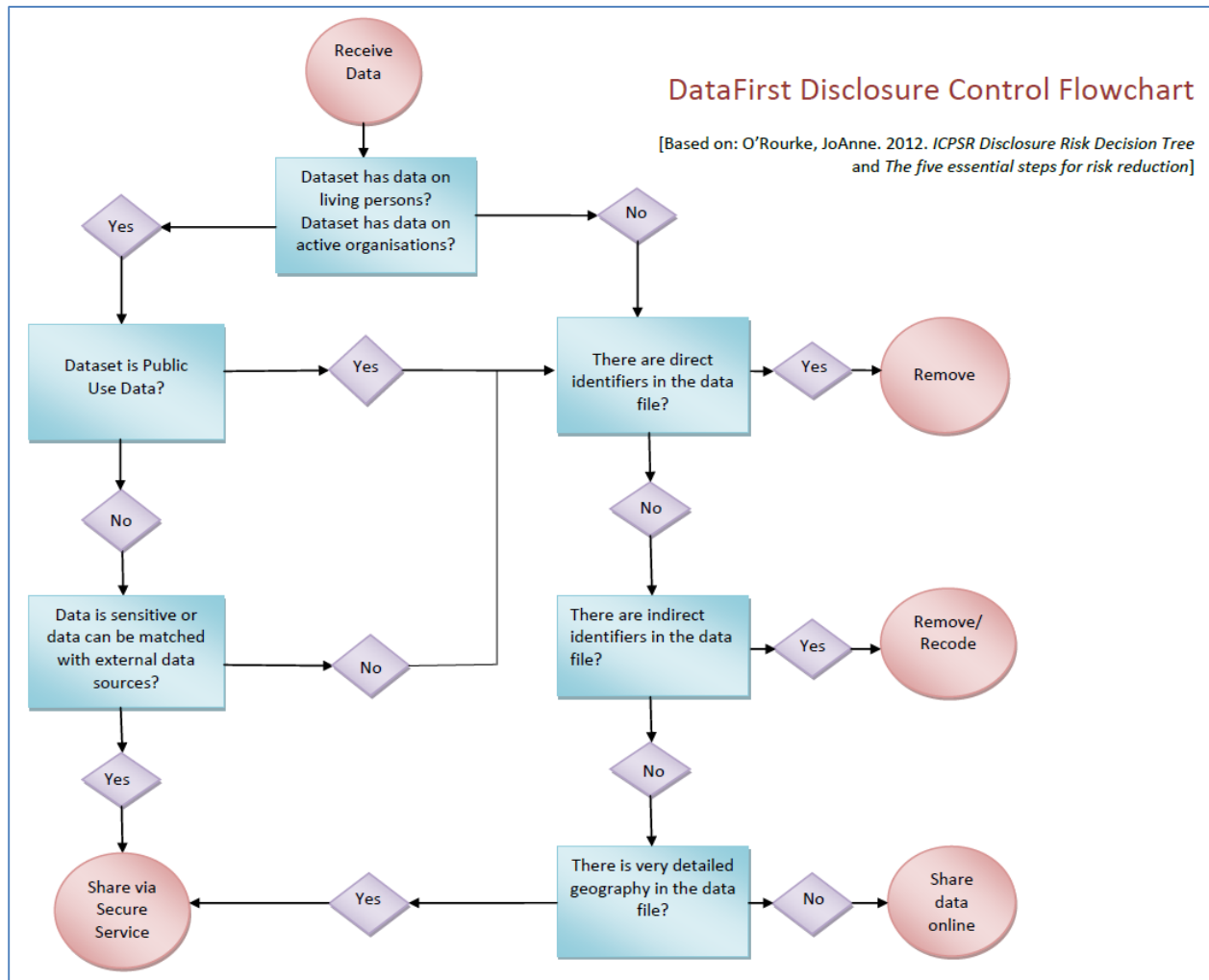
⁴ <http://www.oclc.org/research/publications/library/2000/lavoie-oais.html>

Appendix 1. Research Data Management Workflow



QUL RDM Workflow Diagram

Appendix 2. Disclosure Control Flowchart



[Based on: O'Rourke, JoAnne. 2012. *ICPSR Disclosure Risk Decision Tree and The five essential steps for risk reduction*]