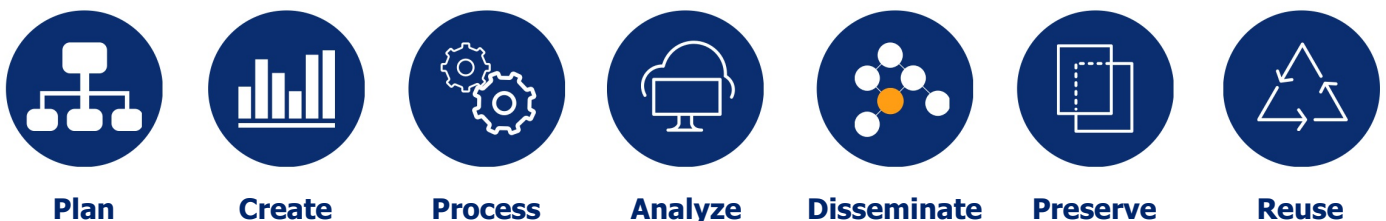


RESEARCH DATA MANAGEMENT PRIMER

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, 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://digitalleadership.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 CASRAI's Dictionary. Retrieved from dictionary.casrai.org/Research_data

"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, DDI, etc.).