



# **(Some) Research Data Management Best Practices**

James Doiron

Research Data Management Services Coordinator, University of Alberta Library  
Academic Director, UofA Research Data Centre  
Co-Chair, Portage DMP Expert Group

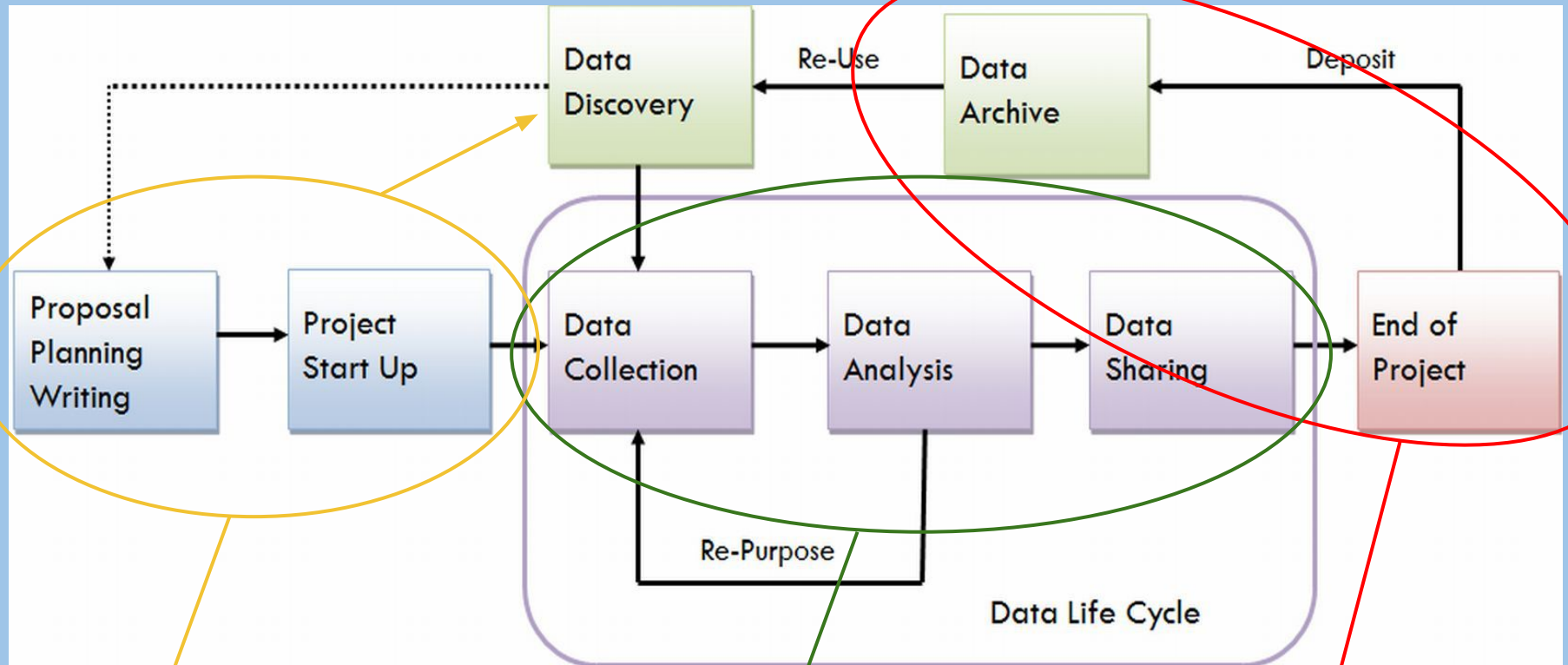
Jane Fry

Data Services Librarian, Carleton University  
Co-Chair, Portage RDM Training Expert Group

September 16, 2020

# RDM 'Best Practices': A bird's eye view

## 1. Overarching principles: FAIR; CARE; OCAP



## 2. Pre-Research:

- Funder requirements
- Including RDM into funding applications
- Data Management Plans
- Participant consent & Information Letters
- Ethics applications

## 3. During Research:

- Primary data collection
- Data storage
- Data transferring
- Data access
- Data documentation/metadata
- File naming conventions

## 4. Post-Research:

- Publication requirements
- Data deposit

# 1. Overarching Principles

## FAIR

**FAIR** is a set of guiding principles focused towards making data **F**indable, **A**ccessible, **I**nteroperable and **R**eusable:

- **Findable** - Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.
- **Accessible** - Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.
- **Interoperable** - Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- **Reusable** - Data and collections have clear usage licenses and provide accurate information on provenance.



**\*Key Reading:** Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*3:160018 doi: 10.1038/sdata.2016.18 (2016).

# FAIR Principles: Key Resource

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

- Detailed information across the FAIR principles
- Implementation Networks
- News
- Event
- Resources!



**GO FAIR**

FAIR Principles Implementation Networks News Events **Resources** About GO FAIR

## Resources

Home » Resources

This page is dedicated to resources that you might find useful in your FAIR endeavours. We have collected our GO FAIR materials as well as interesting papers & publications, tools and more for your information and use:

- **GO FAIR Materials**
- **GO FAIR Workshop Series**
- **FAQ**
- **Starter Kit for Research Data Management**
- **More on FAIR**
- **Glossary**

GO FAIR Materials

- > Materials for INs
- > Materials for Countries
- > Materials from GO FAIR meetings
- > Media & Communications Material

GO FAIR Workshop Series

- > Metadata for Machines Workshops
- > Germany GOes FAIR Workshops
- > Pillar-Specific Workshops
- > Manifesto Writing Workshop

FAQ

- > RDM Starter Kit
- > More on FAIR
- > Glossary

# 1. Overarching Principles

## CARE

**CARE** is a set of guiding principles for Indigenous data governance:

- **Collective benefit** for inclusive development and innovation, improved governance and citizen engagement, and equitable outcomes
- **Authority to control** - Recognizing rights and interests, data for governance, and governance of data
- **Responsibility** for positive relationships, expanding capability and capacity, and Indigenous languages and worldviews
- **Ethics** for minimizing harm and maximizing benefit, justice, and future use of data



\*Key Readings available at:

<https://www.gida-global.org/resources>

# CARE Principles: Key Resource

[gida-global.org/care](https://gida-global.org/care)

- Detailed information across the CARE principles
- Foundational readings & publications
- News
- Events
- Resources!



HOME ABOUT US CARE PRINCIPLES OF INDIGENOUS DATA GOVERNANCE RESOURCES

## CARE Principles for Indigenous Data Governance

The CARE Principles for Indigenous Data Governance can be downloaded here in [summary](#) or [full](#)

### CARE Principles for Indigenous Data Governance

The current movement toward open data and open science does not fully engage with Indigenous Peoples rights and interests. Existing principles within the open data movement (e.g. FAIR: findable, accessible, interoperable, reusable) primarily focus on characteristics of data that will facilitate increased data sharing among entities while ignoring power differentials and historical contexts. The emphasis on greater data sharing alone creates a tension for Indigenous Peoples who are also asserting greater control over the application and use of Indigenous data and Indigenous Knowledge for collective benefit.

This includes the right to create value from Indigenous data in ways that are grounded in Indigenous worldviews and realise opportunities within the knowledge economy. The CARE Principles for Indigenous Data Governance are people and purpose-oriented, reflecting the crucial role of data in advancing Indigenous innovation and self-determination. These principles complement the existing FAIR principles encouraging open and other data movements to consider both people and purpose in their advocacy and pursuits.

# OCAP

The First Nation Principles of OCAP are a set of standards that establish how First Nations data should be collected, protected, used or shared:

- **Ownership:** refers to the relationship of First Nations to their cultural knowledge, data & information - a **community/group collectively owns information** in the same way that an individual owns his/her personal information
- **Control:** affirms that **First Nations communities have rights in seeking control over all aspects of research** - from start to finish - that impact them. This extends to control of resources and review processes and management of information.
- **Access:** First Nations must have access to **information and data about themselves and their communities** regardless of where it is held, and have the **right to manage and make decisions** regarding access to their collective information.
- **Possession:** Refers to the physical control of data - the **mechanism by which ownership can be asserted and protected**.



*"OCAP® is a registered trademark of the First Nations Information Governance Centre (FNIGC)"*  
[www.FNIGC.ca/OCAP](http://www.FNIGC.ca/OCAP)

# OCAP Principles: Key Resource

## First Nations Information Governance Centre

(<https://www.FNIGC.ca/>)

- Fundamentals of OCAP online training program
- FNIGC data online
- First Nations Data Centre (data by request)
- First Nations surveys (i.e., regional health, early childhood, education, labour, oral health)
- FNIGC online library



Home About Our Work Data Access Media Room News OCAP® English Français LOGIN Search

Home FNIGC | CGIPN  
First Nations Information Governance Centre  
Le Centre de gouvernance de l'information des Premières Nations

OCAP®  
HOME / OCAP®

Follow us   

### The First Nations Principles of OCAP®

#### What is OCAP®?

The First Nations principles of OCAP® are a set of standards that establish how First Nations data should be collected, protected, used, or shared. They are the *de facto* standard for how to conduct research with First Nations.

Standing for ownership, control, access and possession, OCAP® asserts that First Nations have control over data collection processes in their communities, and that they own and control how this information can be used.

#### What do the four “OCAP®” principles mean?

There are four components of OCAP®: Ownership, Control, Access and Possession.

**Ownership** refers to the relationship of First Nations to their cultural knowledge, data, and information. This principle states that a community or group owns information collectively in the same way that an individual owns his or her personal information.

**Control** affirms that First Nations, their communities, and representative bodies are within their rights in seeking to control over all aspects of research and information management processes that impact them. First Nations control of research can include all stages of a particular research project from start to finish. The principle extends to the control of resources and review processes, the planning process, management of the information and so on.

The fact that First Nations must have access to information and data

**The First Nations Information Governance Centre presents our newest online course:**  
**The Fundamentals of OCAP®**  
**LEARN MORE & REGISTER** 

**THE FIRST NATIONS DATA CENTRE**  
A new data-access service from FNIGC

**Latest Tweets**

 Very proud to be co-hosting w/ @FNIGC the First Nations Data Governance Strategy Summit in Calgary this week. An L...  
<https://t.co/c0BmwReBNG> - 1 day 11 hours ago

 Condolences to the family and friends of Sayisi Dene Elder,

## 2. Pre-Research

### **RDM Funder requirements**

Awareness of funder requirements helps to identify:

- specific supports needed;
- collaborative opportunities;
- RDM supports to leverage.

**When involved in funded research projects, know who the funder is and:**

- 1) What general RDM related policies they may have; and
- 2) If there are any RDM related requirements pertaining to the specific call for funding.

# 2. Pre-Research

## General RDM Funder requirements

Example:

[‘Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans’](#) (TCPS-2)

Chapter 5 .....	55
PRIVACY AND CONFIDENTIALITY.....	55
Introduction .....	55
A. Key Concepts .....	55
B. Ethical Duty of Confidentiality .....	55
C. Safeguarding Information .....	55
D. Consent and Secondary Use of Identifiable Information Research Purposes .....	55
E. Data Linkage.....	55
Chapter 8 .....	97
MULTI-JURISDICTIONAL RESEARCH.....	97
Introduction .....	97
A. Review Mechanisms for Research Involving Multiple Institutions and/or Multiple Research Ethics Boards .....	97
outside the Institution .....	101
Chapter 9 .....	105
RESEARCH INVOLVING THE FIRST NATIONS, INUIT AND MÉTIS PEOPLES OF CANADA .....	105
Introduction .....	105
A. Key Concepts and Definitions.....	105
B. Interpreting the Ethics Framework in C. Applying Provisions of This Policy in	105
Chapter 13 .....	181
HUMAN GENETIC RESEARCH .....	181
Introduction .....	181
A. Application of Core Principles to Genetic Research .....	181
B. Plans for Managing Information Revealed through Genetic Research .....	182
C. Genetic Counselling.....	184
D. Genetic Research Involving Families .....	184
E. Genetic Research Involving Communities and Groups .....	185
F. Genetic Material Banks .....	186
G. Gene Transfer .....	186

## 2. Pre-Research

### Specific RDM Funder requirements

Example: [CRAft Digital Research Archive grants](#)



**Kule Institute for Advanced Study**

ABOUT FUNDING PROJECTS PEOPLE PARTNERS **NEWS & EVENTS**

[Home](#) / [News & Events](#) / [KIAS News Collection](#) / Funding opportunity CRAft digital research archive grant

### Funding opportunity: CRAft Digital Research Archive Grants

Announcing grants for the creation of digital archives for research in the social sciences, humanities and arts. The Kule Institute for Advanced Study, the Digital Initiatives unit of the Library, and the Arts Resource Centre are partnering to support small CRAft grants with a value of \$10,000-15,000. Deadline for applications is December 1st.

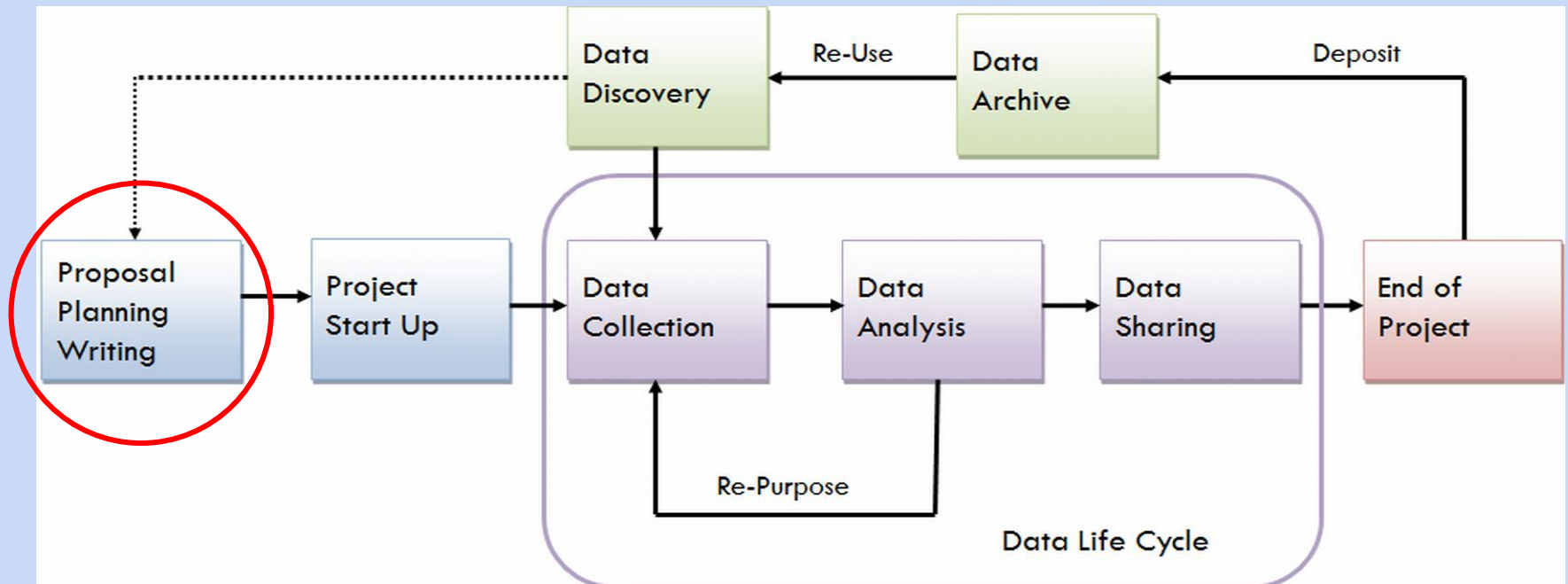
- Research data management plan with a focus on data accessibility and stewardship. \*\*\*Please note that the RDM plan is not considered part of the 5-page proposal and should be included as an attachment. We recommend using the [Portage DMP Assistant](#) to generate an RDM plan.\*\*\*

## 2. Pre-Research

### Including RDM into funding applications

Including RDM within funding applications can help to strengthen funding applications by identifying:

- areas where essential RDM support may be needed
- collaborative opportunities before the research begins
- specific RDM supports that research projects may leverage



## 2. Pre-Research

### Including RDM into funding applications

#### Example RDM statement:

**Research Data Management:** The University of Alberta Libraries system will provide research data management training and support for project researchers on a one-on-one and group basis (including HQP), host project research data in Dataverse, UAlberta's data repository, and host project papers and publications, learning objects, digital images, etc. in its open-access Education & Research Archive.<sup>[i]</sup> Data management will extend beyond the project itself to ensure sustainability of the data for future researchers. We will form a Research Data Management Committee (reporting to the EC; see *Governance*) and use the Portage DMP Assistant<sup>[iii]</sup>, a web-based open source application, to develop a Research Partnership Data Sharing Agreement at the beginning of the formal partnership. One of the major outcomes of this project will be well-documented, well-preserved data sets which can be used by future researchers and are themselves a form of scholarship.

*\*Special thanks to [Dr. Carla Peck](#) from the Faculty of Education, UofA, for permission to use this text from her recently successful application for SSHRC Partnership funding - "Thinking Historically for Canada's Future"*

## 2. Pre-Research

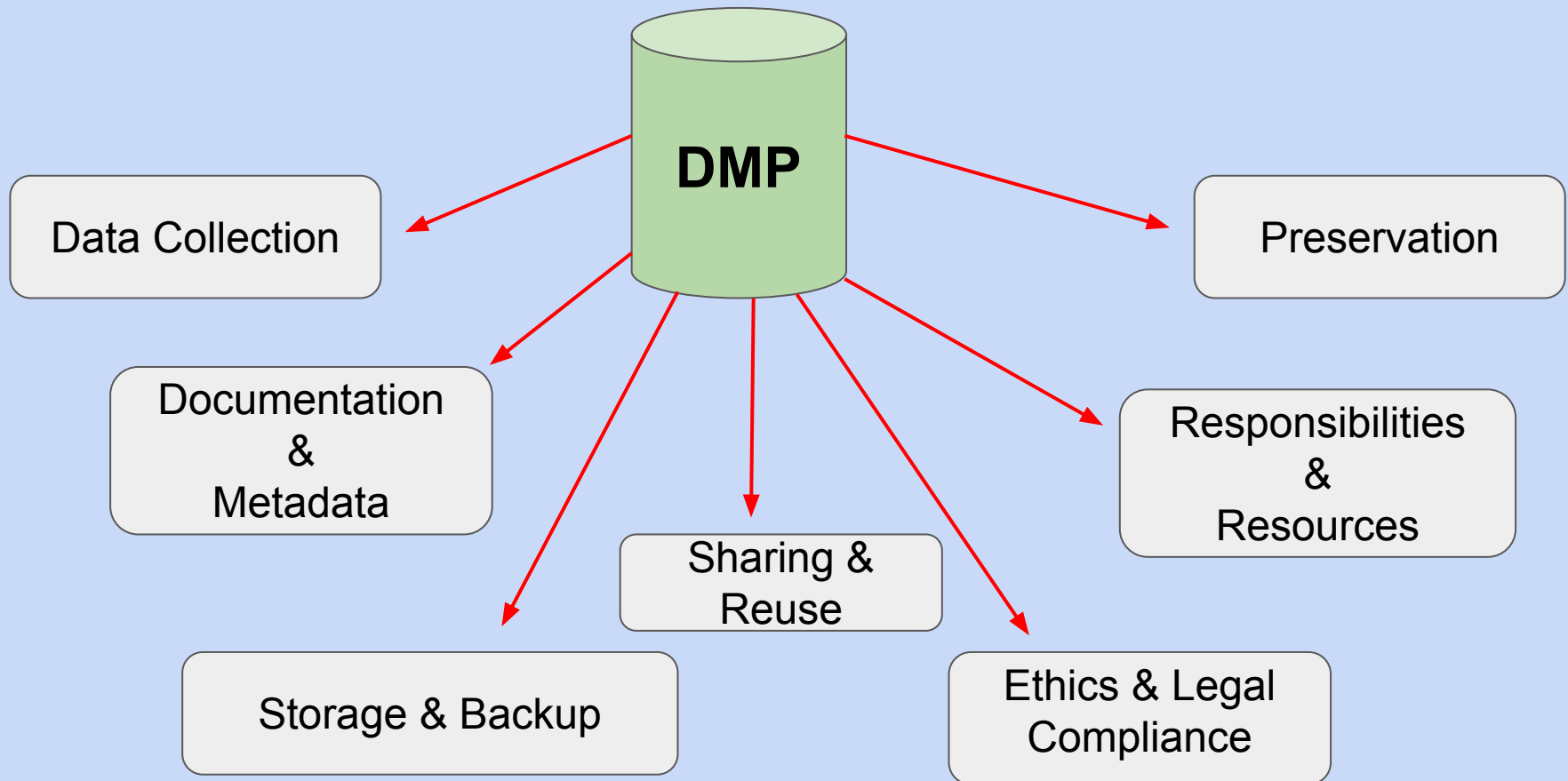
### What *is* a data management plan (DMP)?

#### A DMP:

- Is a formal document which clearly articulates the strategies and tools you will implement to effectively manage your data.
- Speaks to the management of data both **during** the active phases of your research and **after** the completion of the research project.

The objective of a DMP is to address issues related to data management prior to starting your research project!

A DMP provides information across key research lifecycle categories:

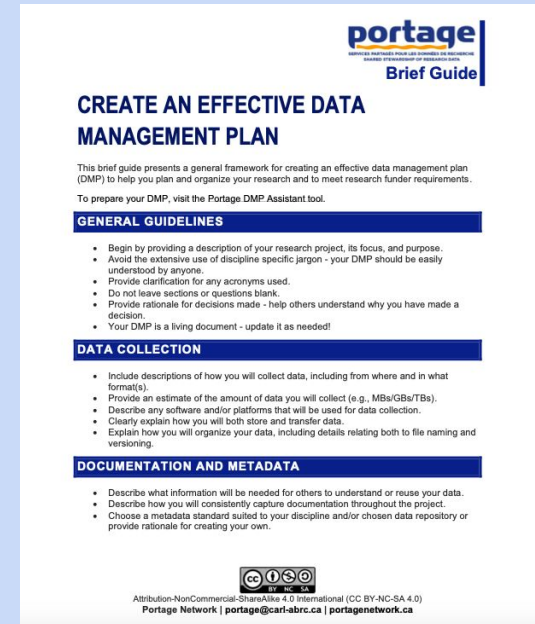


# DMP guidance resources

## Creating an effective DMP ([English](#)/[French](#))

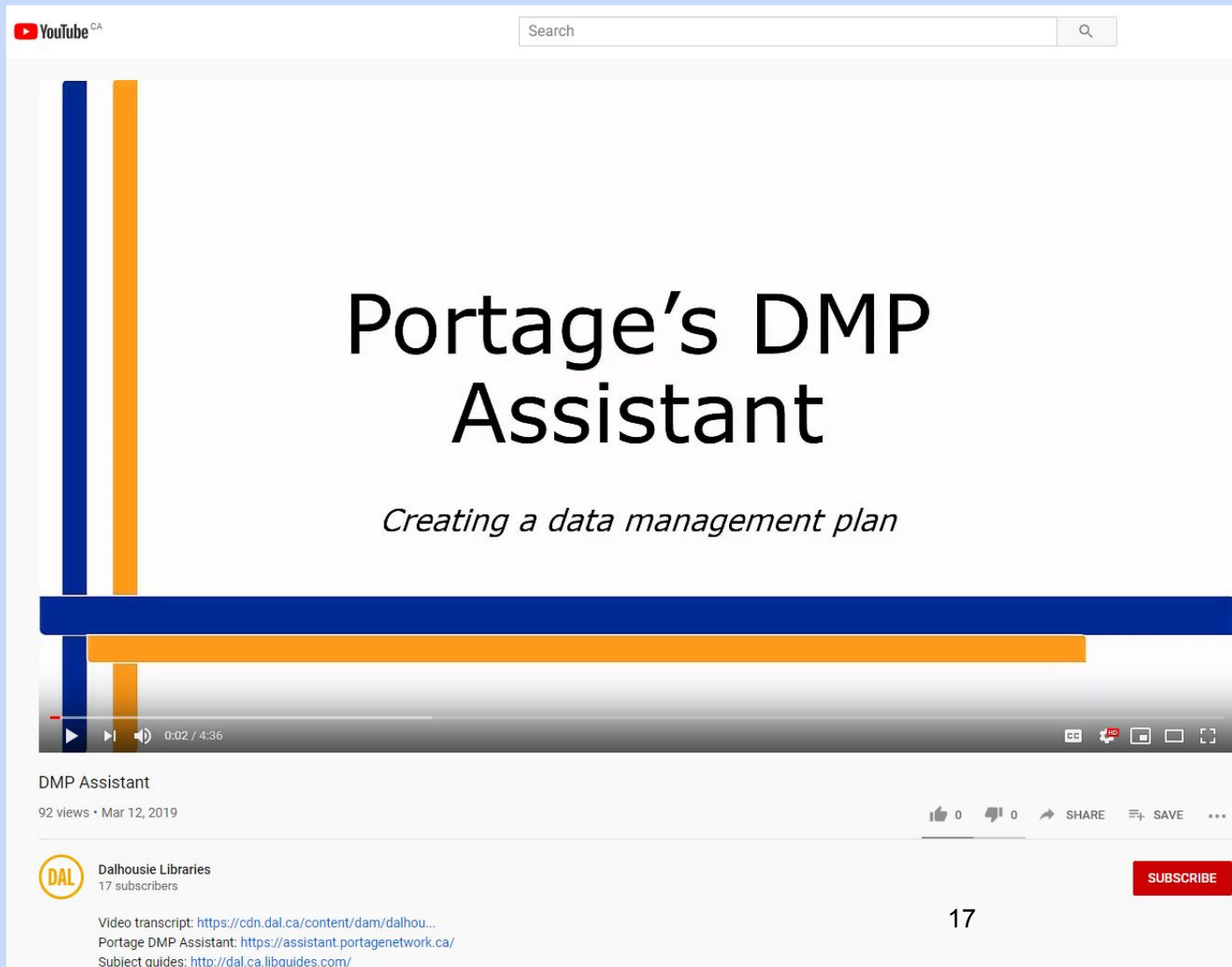
### Exemplar DMPs:

- #1: Digital Humanities ([English](#)/[French](#))
- #2: Digital Humanities and Secondary Data ([English](#)/[French](#))
- #3: Mixed Methods ([English](#)/[French](#))



# Short tutorial video

[https://libcasts.library.dal.ca/Portage/DMP\\_Assistant/](https://libcasts.library.dal.ca/Portage/DMP_Assistant/)



The image shows a YouTube video player interface. At the top, there is a search bar with the text "Search" and a magnifying glass icon. Below the search bar, the video title "Portage's DMP Assistant" is displayed in a large, bold, black font. Underneath the title, the subtitle "Creating a data management plan" is shown in a smaller, italicized font. The video player itself shows a progress bar at the bottom with a play button, a volume icon, and a timestamp of "0:02 / 4:36". To the right of the progress bar, there are icons for closed captions, a red heart, a share icon, a save icon, and a full screen icon. Below the video player, the video title "DMP Assistant" is repeated, followed by "92 views • Mar 12, 2019". To the right of this information, there are icons for likes (0), dislikes (0), share, save, and a menu icon. Below the video player, the channel name "Dalhousie Libraries" is displayed with a logo and "17 subscribers". To the right of the channel name is a red "SUBSCRIBE" button. At the bottom of the page, there is a video transcript section with three links: "Video transcript: https://cdn.dal.ca/content/dam/dalhou...", "Portage DMP Assistant: https://assistant.portagenetwork.ca/", and "Subject guides: http://dal.ca.libguides.com/".

YouTube CA

Search

## Portage's DMP Assistant

*Creating a data management plan*

0:02 / 4:36

DMP Assistant

92 views • Mar 12, 2019

0 0 SHARE SAVE ...

**Dalhousie Libraries**  
17 subscribers

Video transcript: <https://cdn.dal.ca/content/dam/dalhou...>  
Portage DMP Assistant: <https://assistant.portagenetwork.ca/>  
Subject guides: <http://dal.ca.libguides.com/>

**SUBSCRIBE**

## 2. Pre-Research

### Participant consent & Information letters

Research involving human participants requires informed consent.

Information letters must describe *how data are handled during active phases and beyond*

What can/can't be done will in part be determined by what is said in:

- Participant information letters/consent forms
- Approved ethics applications

**NOTE:** It can be very difficult, or even impossible, to go back to participants to revise their consent, so getting things right at the start of your research project is important!

## 2. Pre-Research

### Participant consent & Information letters

Outline such things as:

- project background
- purpose of the study
- study procedures
- benefits/risks
- data preservation/destruction
- security/confidentiality
- voluntary participation
- freedom to withdraw

#### **\*FUTURE USE OF DATA:**

Participant consent is required in order for data to be used beyond the scope of the immediate project.

#### **Example statement:**

*“By participating in this research I hereby give consent for my \*de-identified\* information to be used for research purposes beyond this immediate project.”*

## 2. Pre-Research

### **Ethics applications**

**An ethics application addresses such things as:**

Research design/methodology, risks/benefits, security/confidentiality, participant information, informed consent, data sensitivity, data collection, & data storage, retention and disposal.

Most research projects involving human participants require ethics approval.

Multi-institutional/regional projects require multiple ethics approvals.

Your institutional ethics office offers essential supports & services to help guide and support the ethical management of your research data!

## 2. Pre-Research

### Ethics applications

#### **\*FUTURE USE OF DATA:**

If there is potential for future use of data outside of the immediate project then this should be clearly stated within the ethics application.

#### **Example statement:**

*“There are no plans to destroy these data. Data will be securely stored **\*enter details of storage methods\*** **\*for \*i.e., the minimum 5 years\***.”*

*\*De-identified\* data may be deposited into an institutional repository for discovery and possible repurposing. Any future use of these data outside of the immediate research project will occur only with all ethical and contractual obligations met.”*

### 3. During Research

#### **Data collection: immediate storage**

A safe definition of 'storage' = 48+ hours

Storing data on mobile devices is not considered best practice.

Any electronic devices used for collecting/storing data should always be encrypted (*i.e., laptops, digital voice recorders, tablets, etc*).

#### **Best Practice:**

- Clear & succinct data collection policies and protocols *define when and how data are transferred off of data collection devices.*
- Be aware of any data storage policies imposed by institutions, funders, data providers, etc.

### 3. During Research

#### Data collection: immediate storage

##### \$1M study at U of M paused after personal health info of 420 participants is breached



Study was led by Peter Jones, who was suspended by the university last year



Kristin Annable · CBC News · Posted: Mar 06, 2019 5:23 PM CT | Last Updated: March 6



#### Best Practice:

- Clear how

It informed them their health information was "not handled, stored or secured properly." According to the letter from the university's access and privacy office, an audit of the program found data was being stored off-site, without encryption, allowing the potential for it to be accessed by a third-party company.

- Be when and  
The announcement of the PHIA breach means all the tests and data from the 420 participants funders, data  
pro will be destroyed and will no longer be used in the study.

### 3. During Research

#### **Data collection: 'longer term' immediate storage**

Sometimes it is necessary to store data on electronic devices - i.e., laptops, portable hard drives - for longer periods (collecting data in rural/remote areas).

Beyond security risks, these also introduce risk of data loss and/or corruption.

#### **Best Practice:**

*If the use of laptops/desktops/hard drives is deemed necessary for longer term data storage, use the 3-2-1 rule:*

*At least **3** independent copies of your data:  
store copies on **2** different types of media; and  
keep **1** backup copy offsite.*

### 3. During Research

#### **Data transferring**

Transferring of data is a critical stage of the data collection process.

Regardless of whether data are collected from primary or secondary sources, transferring of data is a necessity.

#### **Some Risks:**

Data transfers may occur:

- from field (real world settings)
- from data providers
- between researchers
- between researchers & stakeholders

### 3. During Research

#### Data transferring

Transferring of data is a critical stage of the data collection process.

Regardless of whether data are collected from primary or secondary sources, transferring of data is a necessity.

#### Some Risks:

A photograph of a single, dark-colored rolling suitcase standing on the side of a long, straight, empty road that stretches into the distance under a hazy sky. The road is flanked by dry, grassy fields.

**Loss of data**

A photograph showing a large warehouse or storage area filled with numerous red, duffel-style bags or suitcases. They are stacked in rows, filling the space from the foreground into the background.

**Unintended copies**

A close-up photograph of a black, textured object, possibly a piece of electronic equipment or a case. A red fabric tag is attached to it with the words "WRONG ONE" printed in black capital letters.

**Unintended recipients**

### 3. During Research

#### **Data transferring: Risks**

**Definite don'ts:** e-mail, unencrypted devices

**Typical do's:** Secure FTP; Securely supported MS Sharepoint; secure extranets

#### **Best practice:**

- Identify data transfer methods that you will use before the research begins.
- Talk to your local IT support to identify secure methods available.

### 3. During Research

#### **Data storage: Cloud services**

##### **What is 'Cloud Storage'?**

- Physical storage typically spans multiple servers (sometimes in different locations).
- Data are easily available and remotely accessible → typically 24/7.
- Best practice → built in physical, technical, & administrative safeguards.

### 3. During Research: **Cloud Services**

#### **Do you know where your cloud is?**

‘Clouds’ are clusters of servers → servers need to live somewhere.

#### Find out:

- Where servers are physically located before using a cloud service → local, provincial, Canada?
- What security policies and procedures are in place → disaster recovery, back-ups, etc.



→ [Research Portal](#)

- [Rapid Access Service](#)  
→ Freely available cloud space for Canadian researchers
- [Resource allocation competitions](#)  
→ for greater needs
- **High performance computing**
- **Big data transfers** → i.e., TBs of data
- **Portals for specialized software & tools**
- **Data storage & back-up**

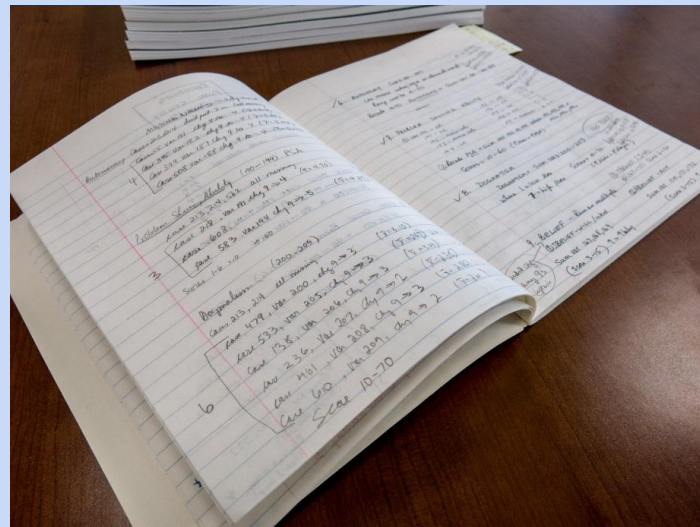
### 3. During Research: **Data documentation & metadata**

‘Metadata’ essentially refers to ‘data about your data’.

Descriptive information that describes your data, as well as to help others (and machines) to locate it and make it readable and useable.

**Some key examples of metadata documentation can include:**

- Data codebooks/dictionaries
- Data management and processing protocols
- Readme files
- Analytic plans
- Code



### 3. During Research: **File naming**

**Having clear and standardized file naming helps to support:**

- **Organization**
- **Quality assurance**
- **File versioning**
- **Collaborative use**
- **Data analysis**
- **Dissemination**
- **Preservation & archiving**
- **Staff/Student training**

**Elements of file names can include:**

- Project name/acronym
- File version
- Data type
- Participant codes/  
pseudonyms
- Geographic location
- Context information
- Date information
- Interviewer codes/  
initials

### 3. During Research: **File versioning**

## Working with data: File versioning

File versioning is an important component of research - it supports such things as participant confidentiality, organization, work efficiency, quality control, analysis, ...

As data are processed (cleaned), new 'versions' are created - from raw data, to the versions which will be used for analysis, and beyond.

Qualitative research example:

'Raw' audio data = the original digital audio recording

'Raw' transcript data = the original and unaltered transcript (text)

'Master' transcript data = the processed transcript → e.g., further de-identified, interviewer comments, typos fixed, etc.

'Analytic' transcript = working copy used for analysis/importing into analytic software

## Transcript File Naming Convention Table

This references the file names at the header on all of the transcripts.

File Version <i>*NOTE: This is only applicable for the Interviews (IN) and Transcripts (IVT)</i>	Illness	Data Type	Dyad Number (Numerically as they enter the study; two digits)	Dyad	Interview / FN/ Photo TYPE <i>*NOTE: Interviews with multiple parts (Part 1, 2, etc) use decimals (i.e., 1.1, 1.2)</i>	Interviewer (first and last initial)	Location (first three letters of city)	Day	Month	Year
Raw	HF=Heart Failure	IV = Audio Interview	01-20=HF	PA=Patient	1= 1st in-person interview / 1st photo	Joanna=JC	EDM=Edmonton	Two digits	First three letters	Use four Digits
Master	LD = COPD/Lung Disease	IVT = Interview Transcribed	21-40=LD	FA=Family	2=2nd in-person interview / second photo	LW=Lacie	VIC = Victoria			
Draft (use when cleaning transcript)	RD=End Stage Renal Disease	FN=Field Notes	41-60=RD		3=Phone interview / third photo	LD= Lindsay				
	CA = Advanced Cancer	PH=Photo	61-80=CA		OTH#= Other (e.g. field note during intake call)	LC=Lynn				
						MA=Marcy				
						RE=Rebecca				
EXAMPLE 2: MASTER_LD_IVT_39_PA_01_MA_VIC_15DEC2015										
Master Transcript	COPD	Interview	Dyad 01	Participant	First Interview	Marcy	Victoria	December 15, 2015		
EXAMPLE 1: Raw_CA_FN_01_FA_OTH1_LD_VIC_15Apr2016										
Raw Transcript	Advanced Cancer	Field note	Dyad 01	Family Member	Field note made at a time not specific to an interview (e.g. dropping off the camera)	Lindsay	Victoria	April 15, 2016		

Antonio, M. G., Schick-Makaroff, K., Doiron, J. M., Shields, L., White, L., & Molzahn, A. (2019). Qualitative data management and analysis within a data repository. *Western Journal of Nursing Research*, 19394591988170.

[doi:10.1177/0193945919881706](https://doi.org/10.1177/0193945919881706)

CA	12/2/2016 11:25 AM	File folder
HF	9/26/2016 8:43 AM	File folder
LD	12/2/2016 11:20 AM	File folder
RD	7/11/2016 8:57 AM	File folder

ality > Data > RD > D42 > PA > Raw Transcripts		
Name	Date modified	
Raw_RD_IV_42_PA_2_LW_EDM_11Mar2016.docx	3/17/2016 7:40 AM	
Raw_RD_IV_42_PA_1.1_RE_EDM_29Jul2015 1&2.docx	8/7/2015 3:43 PM	
Raw_RD_FN_42_PA_3_LW_EDM_09Jul2016.docx	7/26/2016 10:11 AM	
Raw_RD_FN_42_PA_2_LW_EDM_14Mar2016.docx	3/17/2016 7:40 AM	
Raw_RD_FN_42_PA_1_RE_EDM_29July2015.docx	8/7/2015 3:43 PM	
Draft_RD_IV_42_PA_2_LW_EDM_11Mar2016.docx	11/15/2016 3:53 PM	
Draft_RD_FN_42_PA_2_LW_EDM_14Mar2016.docx	4/1/2016 12:05 PM	

alberta.ca > NDR > Liminality > Data > RD		
ity > Data > RD > D42 > PA > Master Transcripts		
Name		
D41		Folder
D42		Folder
D43	Master_RD_FN_42_PA_1_RE_EDM_29July2015.docx	Folder
D44	Master_RD_FN_42_PA_2_LW_EDM_14Mar2016.docx	Folder
D45	Master_RD_IV_42_PA_2_LW_EDM_11Mar2016.docx	Folder
D46	Master_RD_IV_FN_42_PA_3_LW_EDM_09Jul2016.docx	Folder
D51	Master_RD_IVT_42_PA_1.1-1.2_RE_EDM_29Jul2015.docx	Folder
D52		File folder
D53		File folder
D54		File folder
D55		File folder

berta.ca > NDR > Liminality > Data > RD > D41		
Name	Date modified	Type
FA	6/9/2016 2:22 PM	File folder
PA	1/19/2016 10:41 PM	File folder

Liminality > Data > RD > D41 > PA		
Name	Date modified	Type
Audios	12/8/2016 2:11 PM	File folder
Master Transcripts	12/8/2016 2:16 PM	File folder
Photos	1/21/2016 8:06 AM	File folder
Raw Transcripts	12/8/2016 2:06 PM	File folder

## 4. Post-Research: **Publication requirements**

- Increasingly, journals are requiring research data to be made openly accessible.
- Talk to researchers early about journals to which they may be interested in submitting their articles.
- Find out what the journals' data policies are.
- Most journals requiring data to be made openly available will have exceptions for data with legal and/or ethical considerations.
- Refusal to share data are grounds for rejection.

EXAMPLE: [PLOS ONE](#)

**PLOS One** is a peer-reviewed open access scientific journal published by the Public Library of Science (**PLOS**) since 2006. The journal covers primary research from any discipline within science and medicine.

## 4. Post-Research: **Data deposit**

### **What *is* a ‘research data repository’?**

**A research data repository is a technology-based platform that allows for research data to be:**

- Deposited & described
- Stored & archived
- Shared & published
- Discovered & reused

**There are different types of repositories including:**

- Proprietary (paid for services)
- Open source (free to use)
- Discipline specific

**RESOURCE!** → [Re3data.org](https://re3data.org) is an online registry of data repositories, which can be searched according to subject, content type and country. Find a list of [Canadian research data repositories](#).

# Dataverse Features

**Digital Object Identifier (DOI) assigned**

Persistent identifier assigned when upload data

Automatically registered with DataCite  
<https://www.datacite.org/>

**User controlled data access**

From fully open to restricted access

**Many types of data**

Tabular (CSV, SPSS, R, etc)  
Documentation (pdf, doc, text)  
Geospatial data  
Multimedia (audio-visual)

**Built in data citations**

**Usage metrics**

**File Versioning**

Madueke, Ijeoma Sylvia, 2018, "NIGERIAN LITERATURE IN FRENCH TRANSLATION (NILIFT)",  
<https://doi.org/10.7939/DVN/CHNOHA>

**Customized Terms of Use**

Including a customizable guestbook

**Portage Training Materials in development**

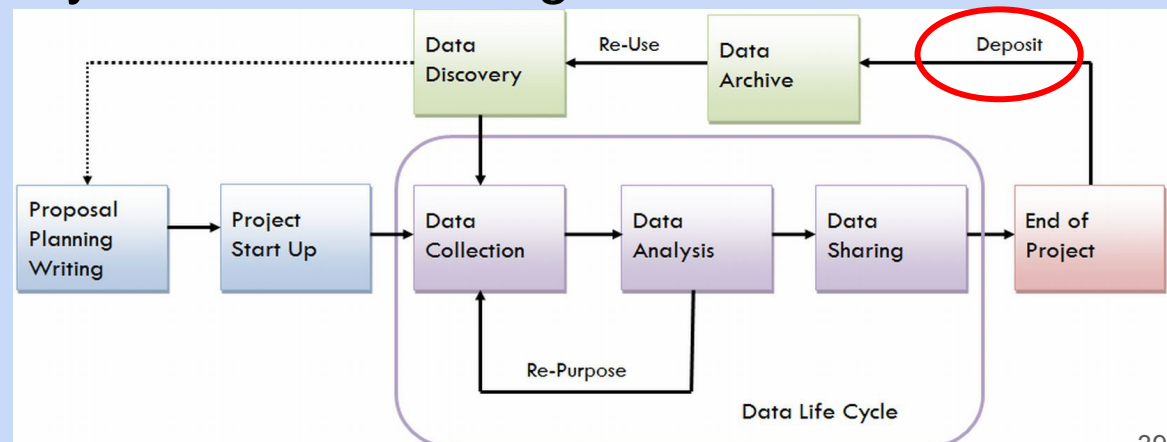
Retain previous versions and metadata

Create Terms of Use for data

High level as well as in-depth web based training modules

# Why should I consider depositing my data?

- Helps you to meet both funding and journal requirements.
- Increases the impact and visibility of your research .
- Digital Object Identifier (DOI) = your data are discoverable and citable.
- Your data are made available, as appropriate, to others → supporting science.
- Obtain metrics on *how* your data are being used.



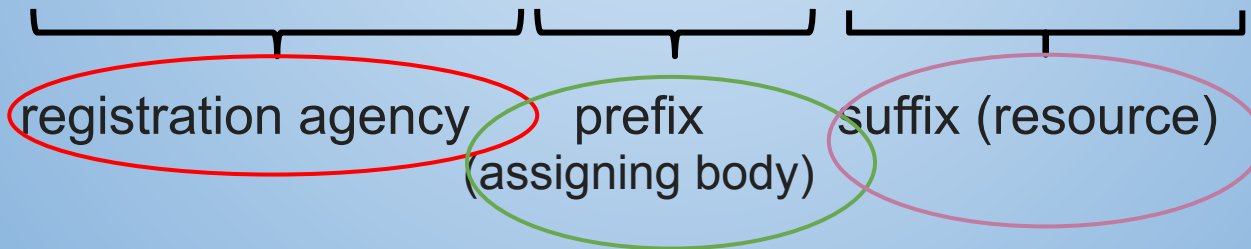
## 4. Post research: DOI

# What is a Digital Object Identifier?

A **digital object identifier** (DOI) is a unique persistent identifier assigned by a registration agency to identify digital content and provide a persistent link to its location.

## Anatomy of a DOI:

<https://doi.org/10.7939/DVN/10591>



## 4. Post research: DOI

# What does a DOI do?

**DOIs** help data producers take control of the management of their research in that they:

- Support the discovery & appropriate sharing of data;
- Support data producers in receiving credit for their data through data citations;
- Make research data easier to access, reuse and verify; and
- Help to meet funder requirements → data deposit and assignment of DOIs are becoming increasingly expected, or even required, in academia.

## 4. Post Research: **DOIs**

# **DOIs - supporting data usage metrics**

Without DOIs our understanding of data usage and impact is extremely limited:

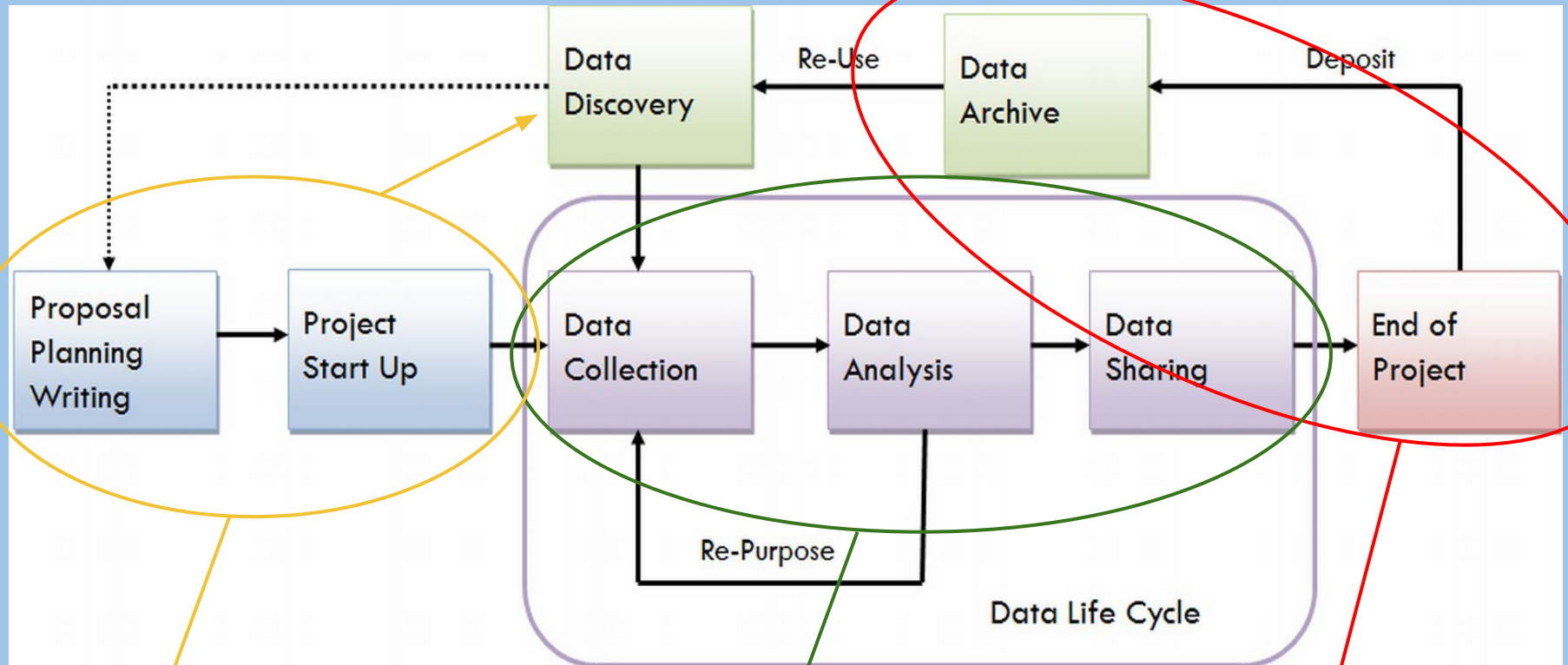
- # of downloads → does not tell us how data are used
- Possible citations, but not easily tracked

A DOI supports data usage metrics by providing:

- Persistent identifier to the data
- Ready-made citation which includes the DOI
- Bi-directional discovery → data to outputs and back to data
- Ability to 'scrape' the web using the DOI

# RDM 'Best Practices': A bird's eye view

## 1. Overarching principles: FAIR; CARE; OCAP



## 2. Pre-Research:

- Funder requirements
- Including RDM into funding applications
- Data Management Plans
- Participant consent & Information Letters
- Ethics applications

## 3. During Research:

- Primary data collection
- Data storage
- Data transferring
- Data access
- Data documentation/metadata
- File naming conventions

## 4. Post-Research:

- Publication requirements
- Data deposit

# Questions & Discussion

James Doiron

Research Data Management Services Coordinator, University of Alberta Library  
Academic Director, UofA Research Data Centre  
Co-Chair, Portage DMP Expert Group  
[jdoiron@ualberta.ca](mailto:jdoiron@ualberta.ca)

Jane Fry

Data Services Librarian, Carleton University  
Co-Chair, Portage RDM Training Expert Group  
[Jane.Fry@carleton.ca](mailto:Jane.Fry@carleton.ca)