Spend less time managing your research data and more time on your research

For Health Sciences and Medicine Researchers

Laure Perrier
University of Toronto Libraries

WHAT WE'LL COVER

- RDM landscape *(10 mins)*

- 4 Simple Things *(40 mins with Q&A)*

- Small-group exercise *(30 mins)*

- Wrap up *(10 mins with Q&A)*

LEARNING OBJECTIVES

To provide participants with,

- An overview of current research data management requirements

- Strategies for implementing best practices in managing research data

WHY IS IT IMPORTANT?

Funding agencies

WHY IS IT IMPORTANT?

Funding agencies
Journals

FUNDING AGENCIES: TRI-AGENCY

CIHR

NSERC

SSHRC

FUNDING AGENCIES: TRI-AGENCY

CIHR

Bioinformatics, atomic, and molecular coordinate data

FUNDING AGENCIES: TRI-AGENCY

CIHR

Bioinformatics, atomic, and molecular coordinate data

All researchers: retain original data for a minimum of 5 years at end of grant

FUNDING AGENCIES: TRI-AGENCY

NSERC
None

FUNDING AGENCIES: TRI-AGENCY

SSHRC

Data preserved and made available within 2 years of completion of research project

TRI-AGENCY: DRAFT RDM POLICY

Data Management Plans

TRI-AGENCY: DRAFT RDM POLICY

Data Management Plans

Share research data

TRI-AGENCY: DRAFT RDM POLICY

Data Management Plans

Share research data

University: Create an institutional strategy

TRI-AGENCY: DRAFT RDM POLICY

Timeline: Summer / Fall 2019

TRI-AGENCY: DRAFT RDM POLICY

Timeline: Summer / Fall 2019

Implementation:
- phased
- incremental

JOURNALS

International Committee of Medical Journal Editors:

Require trials to have a data sharing statement (July 2018)

JOURNALS

International Committee of Medical Journal Editors:

Require trials to have a data sharing statement (July 2018)

PLOS One; BMJ

PLOS | ONE

Acceptable Data-Sharing
Methods

Unacceptable Data Access
Restrictions

Explanatory Notes and
Guidance

Recommended Repos...

Repository Inclusion ...

FAQs for Data Policy

PLOS Data Advisory Board

Give Feedback

Data Availability

PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction at the time of publication. When specific legal or ethical requirements prohibit public sharing of a dataset, authors must indicate how researchers may obtain access to the data.

...able without ...taset, authors must

...n PLOS's policy. If the article is accepted for publication, the data availability statement will be published as part of the accepted article.

Refusal to share data and related metadata and methods in accordance with this policy will be grounds for rejection.

...olicy will be grounds for rejection. PLOS journal editor... ...ning data from articles published in PLOS journa... ...the right to post a correction, to contact the autho...

Meth... ...accompanied by guidance for authors as to what must be indicated in their data availability statement and how to follow best practices in reporting. If authors did not collect data themselves but used another source, this source must be credited as appropriate. Authors who have questions or difficulties with the policy, or readers who have difficulty accessing data, are encouraged to contact the journal office (plosone@plos.org). If you have broader questions about the PLOS data availability policy, contact data@plos.org.

This data policy was implemented on March 3, 2014. Any paper submitted before that date will not have a data availability statement. However for all manuscripts submitted or published before this date, data must be available upon reasonable request per previous policy.
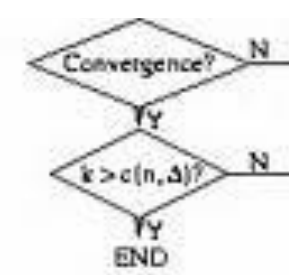
WHAT IS DATA?

✗ No

Published Table in journal: Aggregate data

Table 1:

Statistic	N	Mean	St. Dev.	Min	Max
rating	30	64.633	12.173	40	85
complaints	30	66.600	13.315	37	90
privileges	30	53.133	12.235	30	83
learning	30	56.367	11.737	34	75
raises	30	64.633	10.397	43	88
critical	30	74.767	9.895	49	92
advance	30	42.933	10.289	25	72
high.rating	30	0.333	0.479	0	1

✓ Yes

Raw data file: individual responses

	A	B	C	D	
1	REFID	Retirment Age	Class of worker	Income source	
2	4836	55-64	Public sector	Employment	L
3	4837	65-74	Self-employed	Investment	L
4	4838	55-64	Public sector	Gov't transfer	L
5	4964	75-84	Private sector	Private pensi	L
6					

WHAT IS DATA?

✗ **No** **Published Table in journal:** **Aggregate data**	Table 1: Statistic N Mean St. Dev. Min Max rating 30 64.633 12.173 40 85 complaints 30 66.600 13.315 37 90 privileges 30 53.133 12.235 30 83 learning 30 56.367 11.737 34 75 raises 30 64.633 10.397 43 88 critical 30 74.767 9.895 49 92 advance 30 42.933 10.289 25 72 high.rating 30 0.333 0.479 0 1

✓ **Yes**

Raw data file: individual responses

	A	B	C	D
1	REFID	Retirment Age	Class of worker	Income source
2	4836	55-64	Public sector	Employment
3	4837	65-74	Self-employed	Investment
4	4838	55-64	Public sector	Gov't transfer
5	4964	75-84	Private sector	Private pensi
6				

WHAT IS DATA?

Any material that will allow your results to be reproduced

WHAT IS DATA?

Any material that will allow your results to be reproduced

Includes notes and details

WHAT IS DATA?

Any material that will allow your results to be reproduced

Includes notes and details

Others should be able to understand

4 SIMPLE THINGS

Make a Plan

DATA MANAGEMENT PLANS

Type, format, volume of data

DATA MANAGEMENT PLANS

Type, format, volume of data

Storage, backup, access

DATA MANAGEMENT PLANS

Type, format, volume of data

Storage, backup, access

Ethical, legal issues

DATA MANAGEMENT PLANS

Type, format, volume of data

Storage, backup, access

Ethical, legal issues

Sharing

UOTTAWA: ETHICS REVIEW APPLICATION

REB submissions:

Length of data conservation

UOTTAWA: ETHICS REVIEW APPLICATION

REB submissions:

Length of data conservation

Safeguarding data

UOTTAWA: ETHICS REVIEW APPLICATION

REB submissions:

Length of data conservation

Safeguarding data

Issues of **anonymity** or **confidentiality**

TOOLS

DMP Assistant (Canada): portagenetwork.ca

DMP Online (UK): dmponline.dcc.ac.uk

DMP Tool (USA): dmptool.org

TOOLS

Consent Forms

- Two-tiered
  1) Taking part in study
  2) Use of data

- Examples: UK Data Service
  Google: UK Data Services consent forms

4 SIMPLE THINGS

Make a Plan

Create a System

Chris Buddle @CMBuddle · Following

Still se
from 2
"Super
tuned.

FAVORITES
2

2:49 PM - 1 S

Chris Buddle @CMBuddle · Following

in my "Su
expense
funny pho
#Science

FAVORITE
1

2:56 PM - 1 Sep 201

Chris Buddle @CMBuddle · Following

#protip
NEVER name folders "stuff", "miscellaneous", "important stuff", "misc stuff"

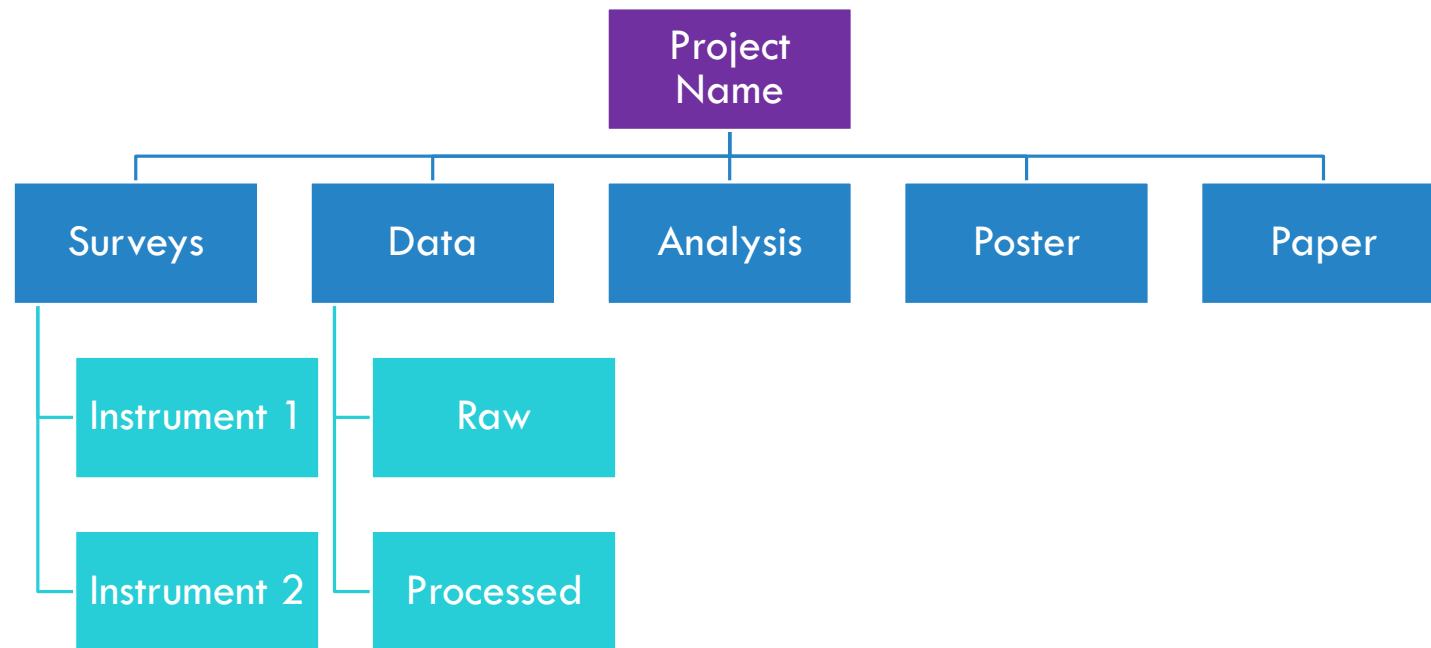Today, I realize this is what I do.
#ScienceInAction

RETWEETS FAVORITES
12   13

3:21 PM - 1 Sep 2015

ORGANIZE YOUR FOLDERS

ORGANIZE YOUR FOLDERS

FILE NAMING EXAMPLE

Project name	Standardized date format	Description of file content	Versioning Information

Project_YYYYMMDD_ContentDescription_Version.ext

Underscore

File extension
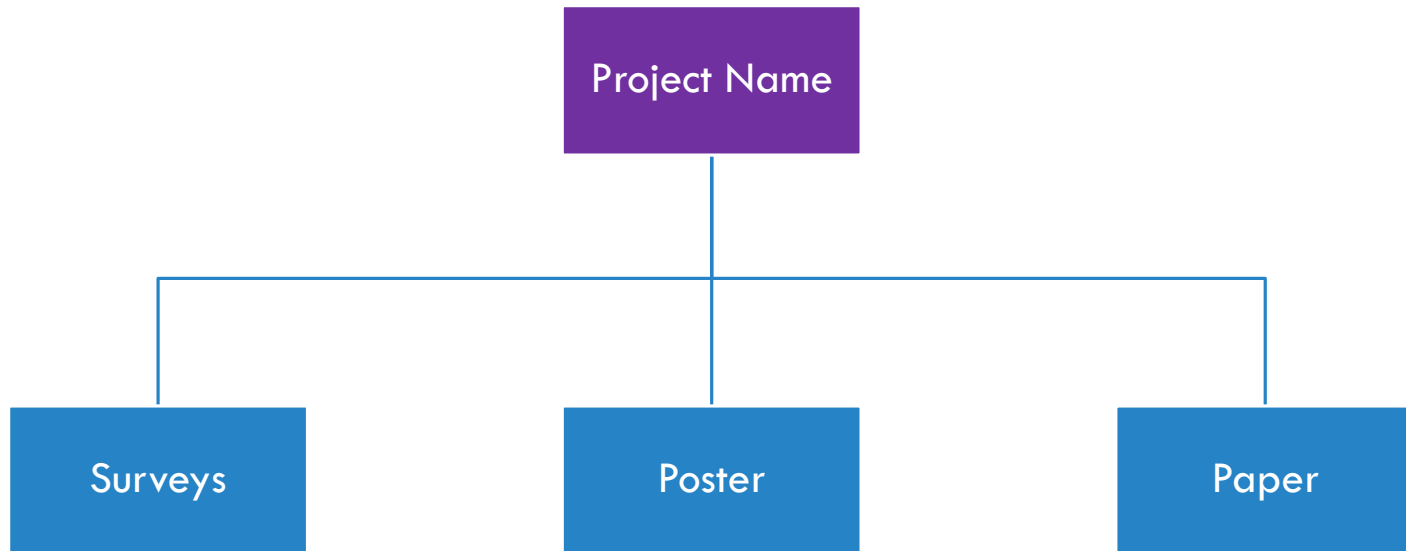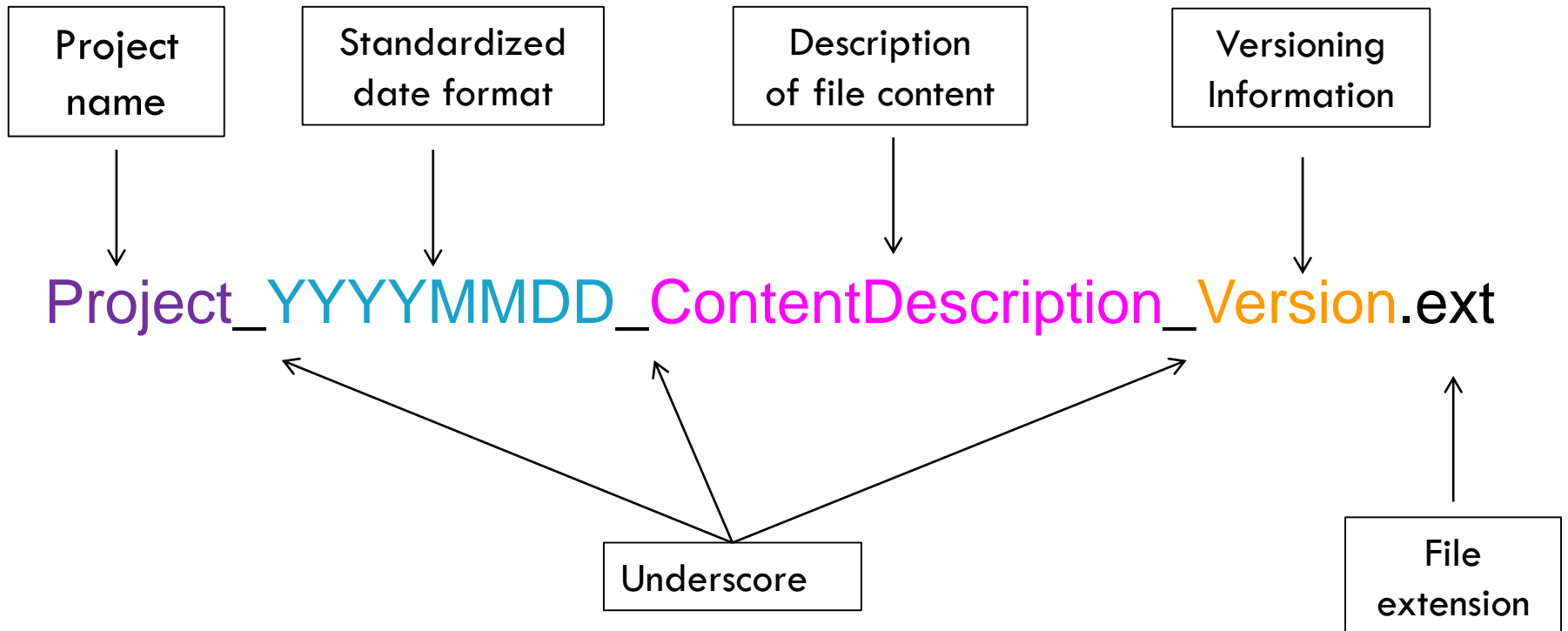
NAMING FILES

Versioning

FileName_**v1**.doc
FileName_**v2**.doc

NAMING FILES

Avoid special characters: #%

Use capitals or underscores
Example: FileName.doc

Standardized (ISO) date:
YYYYMMDD

DOCUMENTATION & METADATA

Important amongst researchers if sharing

DOCUMENTATION & METADATA

Important amongst researchers if sharing

Describe the data for future use

METADATA EXAMPLE

Batman movie

1-24 of 765 results for **Movies & TV Shows** : **Movies** : **English** : **"batman"**

Show results for

< Any Category
< Movies & TV Shows
  Movies

Refine by

Amazon Prime
☐ ✓prime

Shipping Option
FREE Shipping

Delivery Date
☐ Get It by Tomorrow

Video Format
☐ Blu-ray
☐ DVD
☐ VHS
☐ HD DVD
☐ Universal Media Disc
☐ Blu-ray 3D
☐ 4K

Movie & TV Show Original Language
< Clear
☑ **English**
☐ French
☐ Spanish
☐ Italian
☐ German

Blu-ray Region
Region A/1
Region B/2
Region C/3

Video Format: Blu-ray | DVD | VHS | HD DVD | Universal Media Disc | See more ▾

Amazon's Choice

Batman: The Motion Picture Anthology 1989-1997 2009
Starring: Various

Blu-ray Batman / Batman Returns / Batman Forever / Batman & Robin - Bilingual
CDN$ 50.92 ✓prime
Get it by **Tomorrow, Oct 5**

More buying choices
CDN$ 27.99 (12 used & new offers)

Eligible for FREE Shipping

★★★★☆ ▾ 128

Batman: Gotham by Gaslight 2018
Starring: Bruce Greenwood

Blu-ray BD/ DVD/ UV/ BIL
CDN$ 19.99 ✓prime
Get it by **Tomorrow, Oct 5**

More buying choices
CDN$ 17.99 (4 used & new offers)

Eligible for FREE Shipping

★★★★☆ ▾ 19

Batman vs. Two-Face 2017
Starring: Adam West

Blu-ray
CDN$ 14.99 ✓prime
Get it by **Tomorrow, Oct 5**

More buying choices

Batman Ninja 2018
Starring: Fred Tatasciore

Blu-ray BIL/ UV/ DVD/ BD - Steelbook
CDN$ 39.99 ✓prime
Only 8 left in stock - order soon.

More buying choices

Title: Batman Returns

Release Date: 19 June 1992

Alternate Title: Batman 2

Runtime: 126 minutes

Director: Tim Burton

Storyline: Batman battles Penguin. Catwoman shows up to help.

Movie & TV Show Release Decade

< Clear

- [ ] 2010 & Newer
- [ ] 2000 - 2009
- [x] **1990 - 1999**
- [ ] 1980 - 1989
- [ ] 1970 - 1979
- [ ] 1960 - 1969
- [ ] Up to 1960

Blu-ray Region

Region A/1

Region B/2

Region C/3

Region Free

DVD Region

Region 0

Region 1

Region 2

Region 3

Region 4

Region 5

Region 6

Avg. Customer Review

⭐⭐⭐⭐☆ & Up

⭐⭐⭐☆☆ & Up

⭐⭐☆☆☆ & Up

⭐☆☆☆☆ & Up

Price

Under $25

$25 to $50

$50 to $100

Batman 1996
Starring: Michael Keaton

DVD
CDN$ 28.81
Only 1 left in stock - order soon.

More buying choices
CDN$ 1.00 (29 used & new offers)
⭐⭐⭐⭐☆ ▾ 196

★STAR SELECTION

DVD * Batman 1998
CDN$ 15.56 (2 used offers)
⭐⭐⭐⭐☆ ▾ 196

Batman: le défi 1992
Blu-ray
CDN$ 35.66
Only 1 left in stock - order soon.

FREE Shipping
⭐⭐⭐⭐☆ ▾ 125

Batman Returns 1992
Blu-ray Import
CDN$ 46.81
Only 1 left in stock - order soon.

FREE Shipping
⭐⭐⭐⭐☆ ▾ 125

Title: Canada Health Survey

Creator: Tobias Jones

Dates Data Collected: 20160101 to 20161231

Subject: Health Science

Language: English

Location: Canada

Study Type: Cross-sectional

METADATA

Suitability for your study: understand data in detail

METADATA

Suitability for your study: understand data in detail

Other researchers can find, use, properly cite your data

CITING DATA

Data cited similar to a journal citation

4 SIMPLE THINGS

Make a Plan

Create a System

Secure Your Data

3-2-1 RULE

3 copies of your data

3-2-1 RULE

3 copies of your data

2 different locations

3-2-1 RULE

3 copies of your data

2 different locations

More than 1 type of storage hardware

BACK-UP OPTIONS

Personal computer hard drive

External hard drives

Departmental servers

uOttawa contact: Jarno van der Kolk
Senior Scientific Computing Specialist
jvanderk@uottawa.ca

ENCRYPTION

One option = programs such as:

- 7-zip (free at: 7-zip.org)

- Foxit PhantomPDF
  (request education discount $10 US per year at: support@foxitsoftware.com)

ENCRYPTION

- Computer: Disk encryption
  https://it.uottawa.ca/security/disk-encryption

- USB Drives (office supply retailers)

- Secure file transfer: uOttawa LiquidFiles
  https://it.uottawa.ca/email/liquidfiles

uOttawa contact: Jarno van der Kolk
Senior Scientific Computing Specialist
jvanderk@uottawa.ca

SAFE & SECURE STORAGE

Dropbox or Google Drive (generic) are **NOT** secure storage

SAFE & SECURE STORAGE

Dropbox or Google Drive (generic) are **NOT** secure storage

Option:

uOttawa Google Drive File Stream

uOttawa contact: Jarno van der Kolk / jvanderk@uottawa.ca

SAFE & SECURE STORAGE

Option:

Sync.com

- Canadian data residency
- Do not claim ownership to data

SAFE & SECURE STORAGE

Sensitive data:

- Encryption

- Do not use cloud services

4 SIMPLE THINGS

Make a Plan

Create a System

Secure Your Data

Long-Term Access

DATA MUST BE READABLE

✗ **No** **Proprietary formats**	.doc (Word) .xlsx (Excel) .mov (Quicktime) .gif or .jpg (images)
✓ **Yes** **Open formats**	.txt [text] .csv [spreadsheets] .m4a (M-PEG 4) [audio] .tif [images]

DATA MUST BE READABLE

✗ **No** **Proprietary formats**	.doc (Word) .xlsx (Excel) .mov (Quicktime) .gif or .jpg (images)
✓ **Yes** **Open formats**	.txt *[text]* .csv *[spreadsheets]* .m4a (M-PEG 4) *[audio]* .tif *[images]*

DATA MUST BE ACCESSIBLE

Data must include notes, details

DATA MUST BE ACCESSIBLE

Data must include notes, details

Others should be able to understand

DATA DICTIONARY: QUANTITATIVE

Variable	Variable Name	Variable Type	Variable Width	Values / Notes
Participant ID number	ID	Numeric	3	001-900
Date of birth	DOB	yyyy/mm/dd		1900-2010/1-12/1-31
Status	STAT	Numeric	1	1 = alive 2 = deceased
Hemoglobin	HB	Numeric	2.1	4.0 – 18.0

DATA DICTIONARY: QUALITATIVE

Code	Code Name	Explanation
Clarity	CLA	Coherence of components
Structure	STR	Arrangement between component parts
Navigation	NAV	Accurately ascertaining position and planning for movement through information
Saliency	SAL	Quality by which item stands out in relation to its neighbours
Flow	FLO	Moving along in a logical, steady manner

DATA REPOSITORIES

Secure place for research data

DATA REPOSITORIES

Secure place for research data

Can impose appropriate access restriction and/or embargoes

DATA REPOSITORIES

- Subject-specific: re3data.org

- Multidisciplinary (Zenodo; Figshare; Dryad)

- UOttawa: Dataverse

  https://dataverse.scholarsportal.info/dataverse/ottawa
  Contact: Felicity.Tayler@uottawa.ca, E-Research Librarian

ACTIVITY

Get into groups of 2 or 3

ACTIVITY

1) Read: Sample Data Management Plan

2) Fill out: Assessment Sheet

ACTIVITY

Section 1.3: '…how much data (volume) will be produced'

- Do not state sample size

- Do not describe the length of survey

ACTIVITY

Section 2.2: '…data formats created or used'

- "…data formats generally accepted…including SPSS, STATA"

- (Survey data not mentioned)

ACTIVITY

Section 4.4: '…communicating policies'

- Data management plan developed with guidance from UNC policies

- Not clear if these cover guidelines on reuse, redistribution, etc.

ACTIVITY

Section 5.4: '…identifies a timeframe for how long…archived'

- Implied given that repository (Dataverse) is at UNC

🏛 uOttawa

Programs and courses　Future students　Faculties　Research　Alumni　Support uOttawa

Library

Find ▾　**Research help** ▾　Use the library ▾　Locations ▾　My services ▾　About ▾

Research data management

biblio.uottawa.ca/en/services/faculty/ research-data-management

What is data management?

Data management involves ~~the~~ ... ~~at the~~ project's completion. It is an on-going activity throughout the data lifecycle

Read more

What is research data?

Research data are data "used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity"

Why manage research data?

Many publishers and granting agencies are asking authors and

Data management planning

Data management planning includes consideration for many

University of Ottawa Response to the Draft Tri-Agency Research Data Management Policy

uOttawa

Programs and courses   Future students   Faculties   Research   Alumni   Support uOttawa

Library

••• / Faculty / Research data management

Find ▾   **Research help** ▾   Use the library ▾   Locations ▾   My services ▾   About ▾

Research data management

uOttawa: rdm@uottawa.ca

What is data management?

Data management involves the organization and maintenance of data during the active stages of a project, and suitable archiving of the data at the project's completion. It is an on-going activity throughout the data lifecycle.

Read more

What is research data?

Research data are data "used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity"

Why manage research data?

Many publishers and granting agencies are asking authors and

Data management planning

Data management planning includes consideration for many

University of Ottawa Response to the Draft Tri-Agency Research Data Management Policy

QUESTIONS

Felicity Tayler,
Research Data Management Librarian
Felicity.Tayler@uottawa.ca

Laure Perrier

l.perrier@utoronto.ca