

Research Data Management
Laure Perrier
University of Toronto Libraries

February 25, 2016





SO YOU MIGHT BE ASKING YOURSELF...

HOW IS SOMETHING AS MAGICAL & WONDROUS AS **PEACE OF MIND OBTAINED** WHILE DOING RESEARCH?

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE.
ALL YOU NEED
IS 4 KEY THINGS



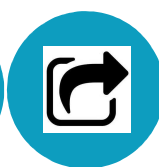
**Make
a Plan**



**Create a
System**



**Secure
Your Data**



**Shareable
Formats**

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

Have you ever had
to prepare a data
management plan?

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

WHY CARE?



- Funding agencies
- Journals

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

WHY CARE?



Canada: Major funding agencies

- **CIHR** (Cdn Institutes of Health Research)
- **NSERC** (Natural Sciences and Engineering Research Council)
- **SSHRC** (Social Sciences and Humanities Research Council)

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

WHY CARE?



CIHR: Current Obligations

- bioinformatics, atomic, and molecular coordinate data
- retain original data sets (all data) for a minimum of 5 years at the end of the grant

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

WHY CARE?



NSERC: Current Obligations

- none

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

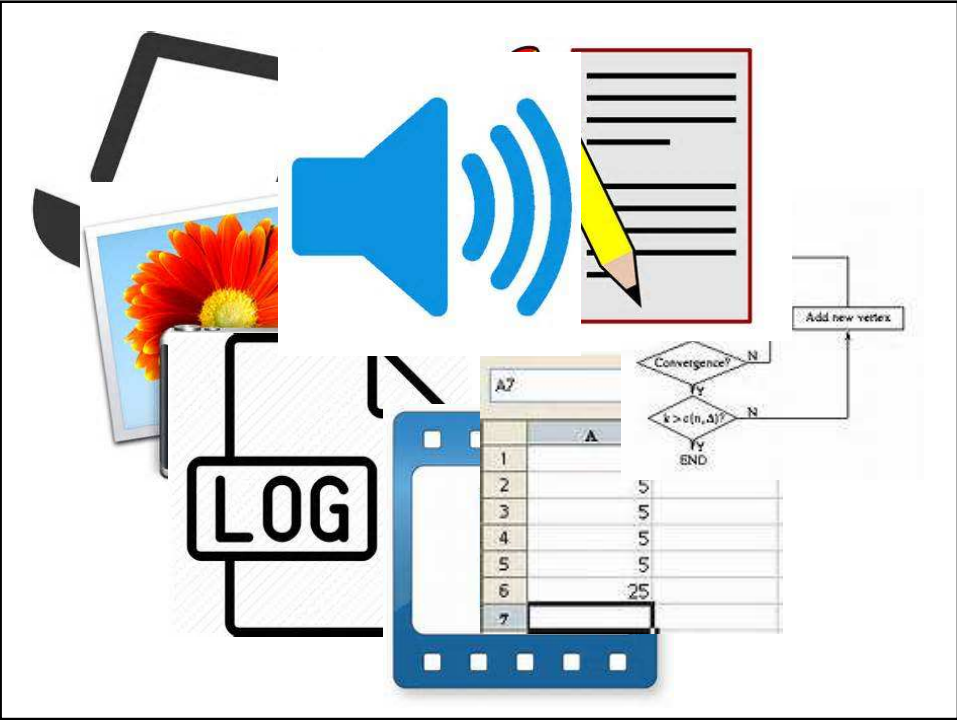
WHY CARE?



SSHRC: Current Obligations

- data must be preserved and made available within two years of the completion of the research project

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016



WHAT IS DATA?



- Any materials that will allow your results to be reproduced
- “What would you need to prove you did not falsify your data?”
- Data must include notes, details
- Others should be able to understand it

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE. ALL YOU NEED IS 4 KEY THINGS



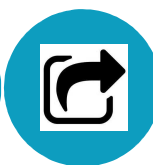
**Make
a Plan**



**Create a
System**




**Secure
Your Data**




**Shareable
Formats**

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 1



Make a Plan



UNIVERSITY OF TORONTO
Office of the Vice-President, Research and Innovation
Human Research Ethics Program

Office Use Only
Application Number:

ETHICS REVIEW APPLICATION FORM FOR SUPERVISED AND SPONSORED RESEARCHERS
(For use by graduate students, postdocs, residents, external investigators, and visiting professors/researchers)

SECTION A – GENERAL INFORMATION

Before you start, familiarize yourself with:
[TCPS2 Application Instructions](#)
[Office FAQs](#)


1. TITLE OF RESEARCH PROJECT

25. DATA SECURITY, RETENTION AND ACCESS

(a) Describe how data (including written records, video/audio recordings, artifacts and questionnaires) will be protected during the conduct of the research and dissemination of results.

(b) Explain how long data or samples will be retained. (If applicable, referring to the standard data retention practice for your discipline) Provide details of their final disposal or storage. Provide a justification if you intend to store your data for an indefinite length of time. If the data may have archival value, discuss how participants will be informed of this possibility during the consent process.

RESEARCH DATA MANAGEMENT PART 1



Make a Plan

25. DATA SECURITY, RETENTION AND ACCESS

(a) Describe how data (including written records, video/audio recordings, artifacts and questionnaires) will be protected during the conduct of the research and dissemination of results.

(b) Explain how long data or samples will be retained. (If applicable, referring to the standard data retention practice for your discipline) Provide details of their final disposal or storage. Provide a justification if you intend to store your data for an indefinite length of time. If the data may have archival value, discuss how participants will be informed of this possibility during the consent process.

(c) If participant anonymity or confidentiality is not appropriate to this research project, please explain.

(d) If data will be shared with other researchers or users, please describe how and where the data will be stored and any restrictions that will be made regarding access.

RESEARCH DATA MANAGEMENT PART 1



Make
a Plan

- Tools available for drafting plans
 - DMP Assistant: portagenetwork.ca
 - DMP Online: dmponline.dcc.ac.uk
 - DMP Tool: dmptool.org

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 1



Make
a Plan

ACTIVITY: Handout

Take 5 minutes to list:

- Project Title

1. Type of Data Produced

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE. ALL YOU NEED IS 4 KEY THINGS



Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



Create a System

Chris Buddle (@CMBuddle) Following

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

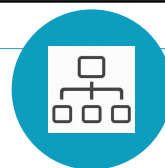
Today, I realize this is what I do. #ScienceInAction

RETWEETS 12 FAVORITES 13

3:21 PM - 1 Sep 2015

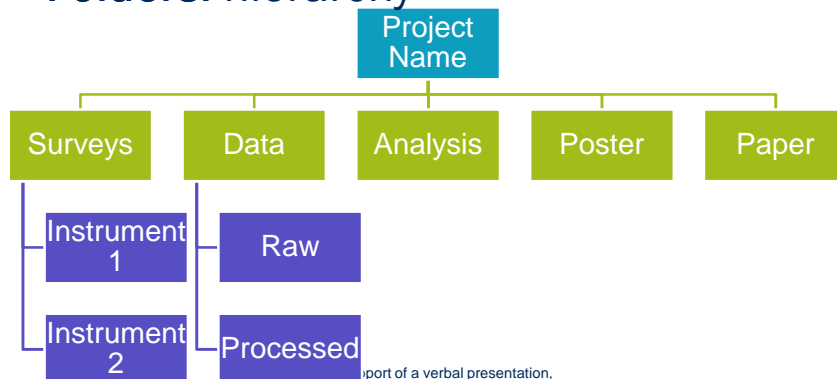
Source: Kristin Briney. Data Management 101 (2015). Retrieved from: <http://www.slideshare.net/kbriney/data-management-101-2015>

RESEARCH DATA MANAGEMENT PART 2



Create a
System

Folders: hierarchy



Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016
Source: Cook C. Love your data week – Day 2. February 9, 2016 [Web log post]. Retrieved from: <http://researchdata.wisc.edu/datapoints-the-rds-blog>

RESEARCH DATA MANAGEMENT PART 2



Create a
System

Folders: hierarchy

- May need to list which files belong in which folders

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



Create a
System

Folders: piling

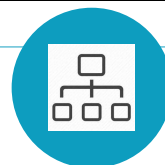
Study

Poster

Paper

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016
Source: Cook C. Love your data week – Day 2. February 9, 2016 [Web log post]. Retrieved from: <http://researchdata.wisc.edu/datapoints-the-rds-blog>

RESEARCH DATA MANAGEMENT PART 2



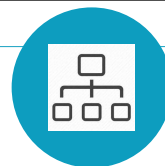
Create a
System

Folders: piling

- Less hierarchy = file names need more detail

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



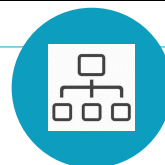
Create a
System

Naming Files

- Be consistent

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



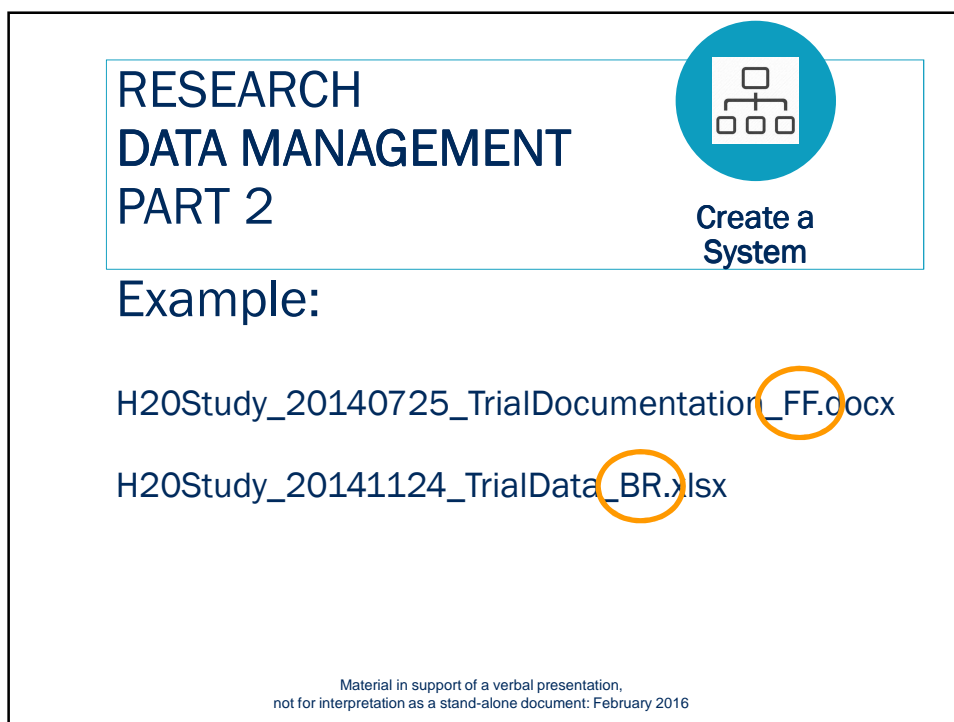
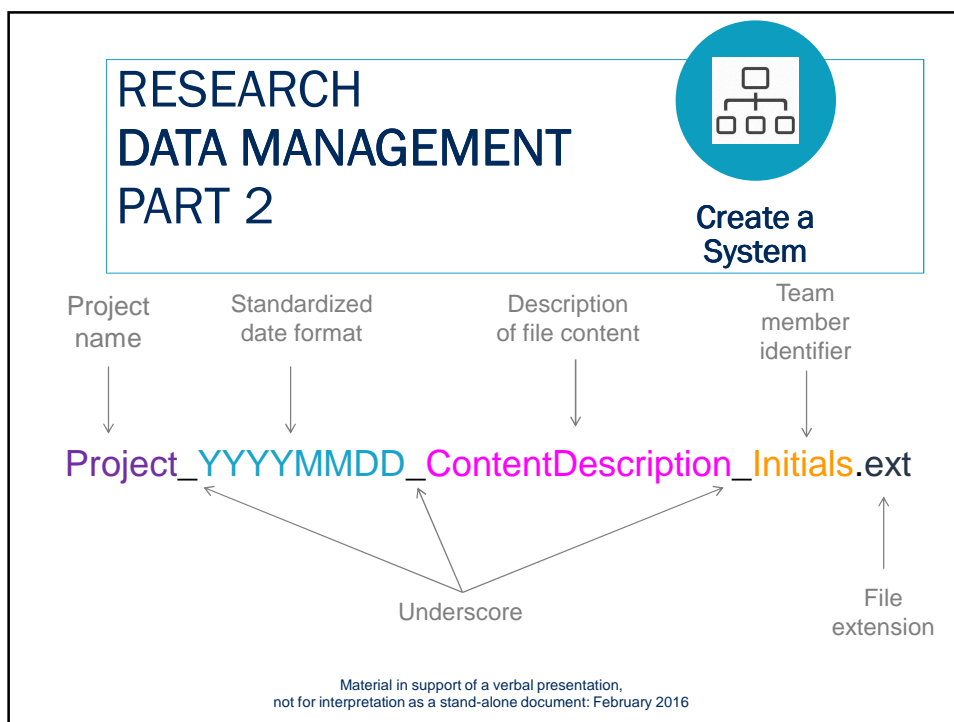
Create a
System

Naming Files

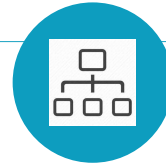
Best practices:

- Avoid special characters (#\$%)
- Capitals or underscores (*FileName.xxx*)
- Date (ISO): YYYYMMDD
- Version information

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016



RESEARCH DATA MANAGEMENT PART 2



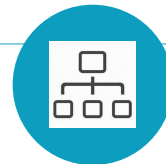
Create a
System

Use meaningful names:

- Project/experiment name or acronym
- Location/spatial coordinates
- Researcher name/initials
- Date or date range of experiment
- Type of data
- Version number file

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



Create a
System

Documentation and metadata:

- Data documentation (metadata) helps you understand data in detail
- Helps other researchers find, use, properly cite your data

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



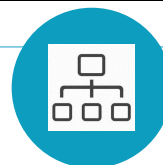
Create a
System

Dublin Core: Example

Title	Soda pop vending machine injuries
Author	Cisio, Michael
Subject	Wounds and Injuries
Description	Chart review and interviews with patients who were injured when soda machines toppled onto them.
Date	19881111
.....AND MORE	

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



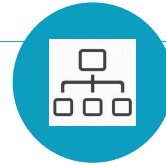
Create a
System

Documentation and metadata:

- Title
- Creator
- Dates
- Subject
- Funders
- Rights
- Language
- Location
- Methodology
- etc.....

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



Create a
System

Documentation and metadata:

- Many standards for specific research disciplines:

<http://www.dcc.ac.uk/resources/metadata-standards>

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 2



Create a
System

ACTIVITY: Handout

Take 3 minutes to list:

2. Documentation and Metadata

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE. ALL YOU NEED IS 4 KEY THINGS



Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



**Secure
Your Data**

CAMBRIDGE NEWS WED 6 °C Cloudy Site Directory Search...

News Sport Business Lifestyle What's On Jobs Property Cars Buy & Sell Directory Notices Dating ALL

Cambridge | Camboorne | Ely | Haverhill | Huntingdon, St Ives & St Neots | Newmarket | Cambridge Live | Royston | Saffron Walden | Your Community | Community Awards | Education | International Student Guide | University News | Cycling | Pictures | Jobs news | Traffic and travel | Fenbrooke's | Help Our Hospice | Obituaries | Involved | E-Edition | RSS

Cambridge student 'devastated' as burglar steals the only copy of his PHD work
By Cambridge News | Posted: January 30, 2016 By Josh Thomas

house hunting? we have hundreds of homes in your area click here

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

3-2-1 Rule:

- 3 copies of your data
- 2 different locations
- More than 1 type of storage hardware

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

Back-up options:

- Personal computer hard drive
- External hard drives
- Departmental servers (if available)

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

Example:

1. On computer
2. Automatic back-up at St. Michael's Hospital
3. Back-up manually (external hard drive): weekly reminder

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

UToronto:

1. Sensitive data must be stored in one of the following ways:
 - a) **Encrypted** if stored **outside** a secure server environment
 - b) **Encryption optional** if stored **inside** secure server environment

Source: Research and Innovation Office and ITS (Information Technology Services)
Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

UToronto:

2. Data in transit to/from the secure server environment must be **encrypted**

- encrypt.utoronto.ca

Source: Research and Innovation Office and ITS (Information Technology Services)

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

Data security plan:

- Who has access
- Strategies to limit access
- Who is responsible
- Keep plan up to date

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 3



Secure
Your Data

ACTIVITY: Handout

Take 5 minutes to list:

3. Security, Storage, Management, and Back-up of Data

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE.
ALL YOU NEED
IS 4 KEY THINGS



Make
a Plan



Create a
System



Secure
Your Data



Shareable
Formats

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data needs to be:

- Readable
- Accessible
- Understandable

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



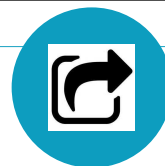
Shareable
Formats

Data needs to be:

- Readable
- Accessible
- Understandable

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data needs to be **Readable**

- Use non-proprietary file formats

Yes	No
.txt	.docx (Word)
.csv	.xlsx (Excel)
.m4a (MPEG-4)	.mov (Quicktime)
.tif	.gif or .jpg (images)
XML	RDBMS

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data needs to be **Accessible**

- Move data to new media
- Average life span ~3-5 years
- If no open file format: Preserve software

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



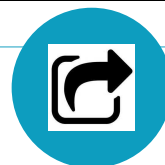
Shareable
Formats

Data needs to be Understandable

- Data must include notes
- Include details
- Others should be able to understand it

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data needs to be Understandable Example: Data dictionary

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
Urinary iodine	UI	Numeric	4.1	0.0 – 1000.0

RESEARCH DATA MANAGEMENT PART 4



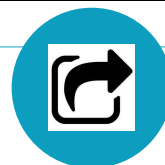
Shareable
Formats

Data repositories

- Secure, long-term place for research data
- Often can impose appropriate access restrictions and/or embargos

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data repositories at UToronto

- Dataverse
<http://dataverse.scholarsportal.info/dvn>
- TSpace (used more for text)
<https://tspace.library.utoronto.ca>
- Collections UofT (beta)
<https://collections.library.utoronto.ca>

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

Data repositories

- Subject-specific:
 - Registry of Research Data Repositories
- www.re3data.org

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

RESEARCH DATA MANAGEMENT PART 4



Shareable
Formats

ACTIVITY: Handout

Take 5 minutes to list:

4. Archiving and Preservation
5. Sharing and Re-Use

e.ggtimer.com/5minutes

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

IT'S SIMPLE. ALL YOU NEED IS 4 KEY THINGS



**Make
a Plan**

**Create a
System**

**Secure
Your Data**

**Shareable
Formats**

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

FINALLY



Sample Data Management Plans

- Generic examples: may need to alter your own plan to fit specific requirements

www.irss.unc.edu/odum/contentSubpage.jsp?nodeid=570

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016

FINALLY



UToronto

onesearch.library.utoronto.ca/researchdata

UNIVERSITY OF TORONTO LIBRARY

Research Data Management

Welcome to the University of Toronto Libraries' Research Data Management resource! Good data management is an important part of the research process and is increasingly emphasized by institutions and funding agencies. University of Toronto Libraries has data management expertise and can assist you in the following areas:

- Organizing your data
- Sharing your data
- Preserving your data
- Contact us

Share this page

Ask a Librarian Chat

Material in support of a verbal presentation, not for interpretation as a stand-alone document: February 2016

FINALLY



Coursera: <https://www.coursera.org/learn/research-data-management-and-sharing>

coursera

Research Data Management and Sharing

The University of North Carolina at Chapel Hill & The University of Edinburgh

About this Course

Upcoming session: March 1 - April 11

Enrollment ends March 6

Material in support of a verbal presentation, not for interpretation as a stand-alone document: February 2016

FINALLY



ACTIVITY: Tweets

In groups of 2 or 3:

Create 2 tweets about the most important things you learned today

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016



QUESTIONS

rdm@utoronto.ca

Laure Perrier
University of Toronto Libraries

Material in support of a verbal presentation,
not for interpretation as a stand-alone document: February 2016