Research Data Management:

Level 1: Data Preservation: The Legacy of your Research Data

JUNE 14, 2017

A. MICHELLE EDWARDS, OAC

CAROL PERRY, LIBRARY

As you finish one project and move onto another, you want to ensure that the legacy of your project's research data is preserved. What best practices should you implement to encourage proper data preservation? In today's research environment, your research data may be requested by other research labs around the world to replicate your data. How should your data be documented and preserved to facilitate the exchange of research data?

Join Carol Perry from the Library and Michelle Edwards, to learn why we need to document our research data as we collect it. How we can use Excel to facilitate this task and why it is important.

Date: Wednesday, June 14, 2017 9am-12noon

Location: ANNU Rm 102 – if you have a laptop please bring it.

Objectives

Understand the basics :

- How organized is your research project?
- How well documented is your data and your research project?
- Are you ready to share/preserve your data?

RDM Refresher

A sound strategy and best practices used to……

- Organize
- Document
- Store
- Analyze
- Secure
- Preserve

……Your data

Setting the stage

Why should I care?

Can you find the data you need?

Can you understand the data you have?

Can others understand the data you have?

Are you required to handle your data in a specific way?

Are you required to share or preserve it?

How will you share/preserve your data?

Where will you share/preserve it?

The Scenarios

Scenario 1

You have just joined a new lab and have been handed a series ofExcel files that are full of data. It is now up to you to decipher the data that was collected.

How do you tackle this?

Scenario 2

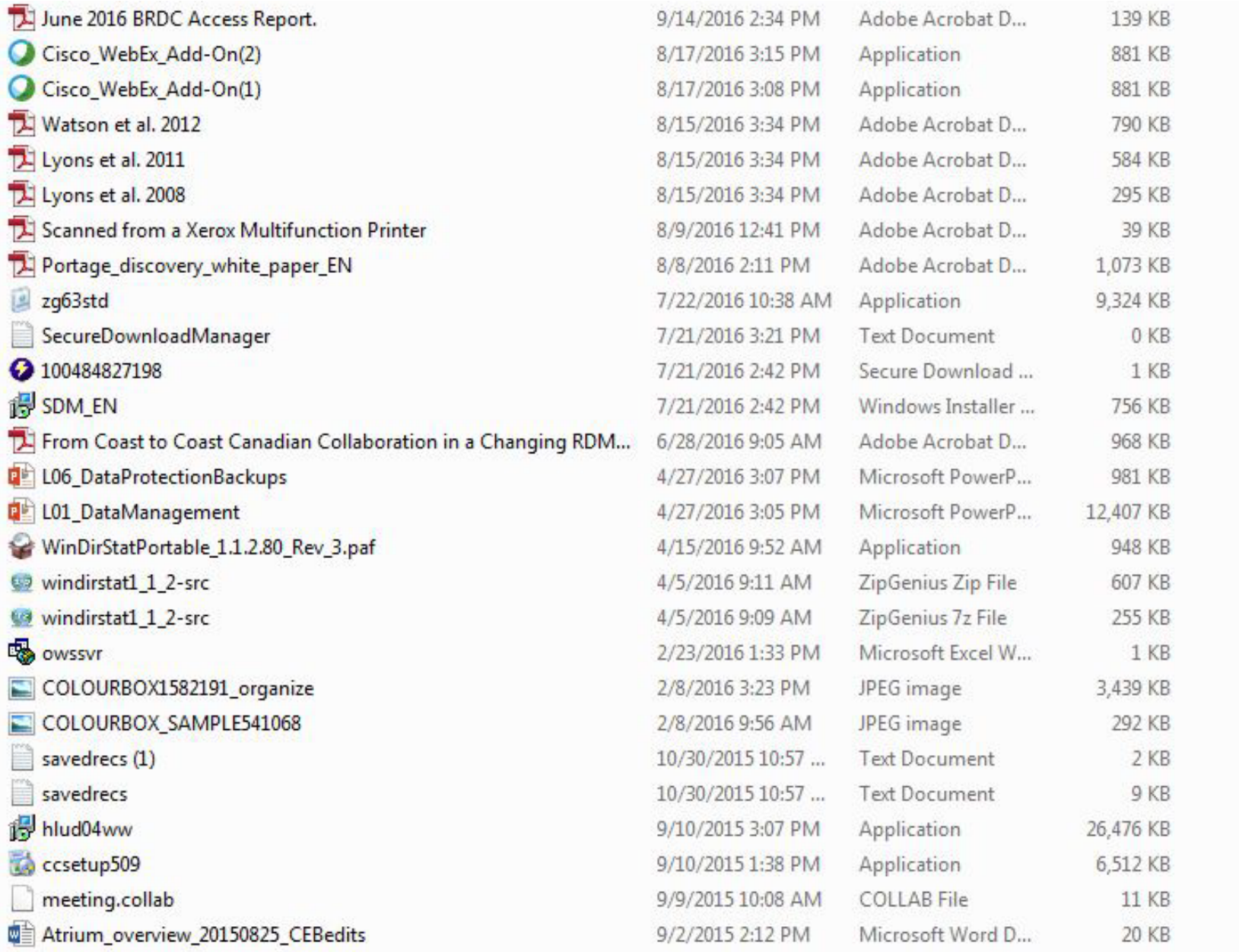
You recently published your project findings in a journal and you have been contacted by a lab in another country who wants a copy of your data so they can replicate your results.

What do you do?

Where to begin?

Is this what your directory structure looks like?

Sample directory: desktop ERS folder



June 2016 BRDC Access Report.	9/14/2016 2:34 PM	Adobe Acrobat D...	139 KB
Cisco_WebEx_Add-On(2)	8/17/2016 3:15 PM	Application	881 KB
Cisco_WebEx_Add-On(1)	8/17/2016 3:08 PM	Application	881 KB
Watson et al. 2012	8/15/2016 3:34 PM	Adobe Acrobat D...	790 KB
Lyons et al. 2011	8/15/2016 3:34 PM	Adobe Acrobat D...	584 KB
Lyons et al. 2008	8/15/2016 3:34 PM	Adobe Acrobat D...	295 KB
Scanned from a Xerox Multifunction Printer	8/9/2016 12:41 PM	Adobe Acrobat D...	39 KB
Portage_discovery_white_paper_EN	8/8/2016 2:11 PM	Adobe Acrobat D...	1,073 KB
zg63std	7/22/2016 10:38 AM	Application	9,324 KB
SecureDownloadManager	7/21/2016 3:21 PM	Text Document	0 KB
100484827198	7/21/2016 2:42 PM	Secure Download ...	1 KB
SDM_EN	7/21/2016 2:42 PM	Windows Installer ...	756 KB
From Coast to Coast Canadian Collaboration in a Changing RDM...	6/28/2016 9:05 AM	Adobe Acrobat D...	968 KB
L06_DataProtectionBackups	4/27/2016 3:07 PM	Microsoft PowerP...	981 KB
L01_DataManagement	4/27/2016 3:05 PM	Microsoft PowerP...	12,407 KB
WinDirStatPortable_1.1.2.80_Rev_3.paf	4/15/2016 9:52 AM	Application	948 KB
windirstat1_1_2-src	4/5/2016 9:11 AM	ZipGenius Zip File	607 KB
windirstat1_1_2-src	4/5/2016 9:09 AM	ZipGenius 7z File	255 KB
owssvr	2/23/2016 1:33 PM	Microsoft Excel W...	1 KB
COLOURBOX1582191_organize	2/8/2016 3:23 PM	JPEG image	3,439 KB
COLOURBOX_SAMPLE541068	2/8/2016 9:56 AM	JPEG image	292 KB
savedrecs (1)	10/30/2015 10:57 ...	Text Document	2 KB
savedrecs	10/30/2015 10:57 ...	Text Document	9 KB
hlud04ww	9/10/2015 3:07 PM	Application	26,476 KB
ccsetup509	9/10/2015 1:38 PM	Application	6,512 KB
meeting.collab	9/9/2015 10:08 AM	COLLAB File	11 KB
Atrium_overview_20150825_CEBedits	9/2/2015 2:12 PM	Microsoft Word D...	20 KB

Folder structure

ERS_weather_monthly_analysis
ERS_weather_monthly_documentation
ERS_weather_monthly_rawdata
ERS_weather_monthly_reports

OR

ers_w_m_analysis
ers_w_m_docs
ers_w_m_raw
ers_w_m_reports

Readme file – sample

Elora Research Station Meteorological Data 2016
ERS = Elora Research Station

Folders:
w = weather data
w = monthly files
raw = rawdata files

Data
Data collected 4 times per day stored in monthly files January 1 2016-December 2016.
Date = ISO date yyyy/mm/dd
UOM = units of measure
Note: equipment failure - battery problem reported on January 15 2016 and January 16 2016.
Out of range values are shown as -6999

Exercise 1

Work in groups of 2 or 3

Review

Did you encounter any problems?

Exercise 2

Review

Did you encounter any problems?

Access, sharing, reuse

Advantages to sharing

Increases impact of your research

Helps others replicate your research

Encourages further scientific enquiry

Reduces research costs by reducing duplication

Encourages transparency and accountability

Access, sharing, reuse

Review legal & ethical obligations/restrictions

- Research Ethics Board
- Funding agency obligations
- Federal/provincial regulations
  - Animal Use Data to Animal Care Services - annual reporting to the Canadian Council on Animal Care

Anonymize data where necessary

- Remove identifying information

Determine sharing criteria – understand your Intellectual Property rights

Preservation options

Institutional repository

Publish with results

Deposit in major data repository

Deposit in discipline-specific data repository

University of Guelph-based options

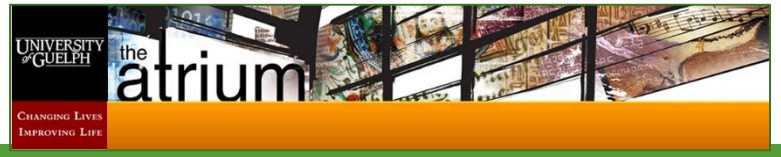
Atrium (institutional repository) - e-theses, articles, reports, videos, etc.
http://atrium.lib.uoguelph.ca/

Agri-environmental Data Repository- research data

https://dataverse.scholarsportal.info/dataverse/ugardr

University of Guelph Data Repository– research data

https://dataverse.scholarsportal.info/dataverse/ugrdr



Discipline Specific Repositories

Re3data.org – global registry of research data repositories
www.re3data.org/

Stanford University Libraries – Guide to Domain-specific Data Repositories

https://library.stanford.edu/research/data-management-services/share-and-preserve-research-data/domain-specific-data-repositories

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Metadata

Metadata = documentation that describes the project and its contents

All repositories are going to require some documentation to accompany your files

- Project documentation
- File documentation
- End user terms of use

Metadata element	Description	Fill in information as applicable
Title	Full title of the dataset	
Author (s)	Person, corporate body, or agency responsible for the work's intellectual content. Include names, affiliations, and email addresses (e.g. [NAME], [DEPARTMENT/SCHOOL], [COLLEGE], [UNIVERSITY].	
Production Date	The date the final version of the dataset being deposited in the repository was created. This is the date that will be used in the dataset citation.	
Funding agency	The source(s) of funds for the production of the work	
Time period covered	Time period covered by the data	
Date of collection	The date(s) when the data were actually collected	
Keywords	Words or phrases that describes the data collection's content (please provide at least 5 keywords)	
Description	Summary describing the purpose and scope of the data collection, and what questions the investigators attempted to answer.	

Sample accessible dataset

http://hdl.handle.net/10864/GSSJX

Contact information:

Michelle Edwards
OAC Statistics Consultant
oacstats@uoguelph.ca

http://www.uoguelph.ca/oac/stats

Carol Perry
Associate Librarian Research & Scholarship
lib.research@uoguelph.ca