research data management
in five easy(ish) steps
about us

Brianna Marshall
Lead, Research Data Services
UW Libraries

Cameron Cook
Digital Curation Assistant
UW Libraries
THE DATA SUGGESTS.....

you can’t find it.

https://i.imgur.com/ntnjb.jpg
or you can’t understand it.
or it’s long gone.
combat these data tragedies!

1. Use open file formats when possible.
2. Organize + name files meaningfully.
3. Document your research process.
4. Back up your data.
5. Have a plan and stick with it.
1. USE OPEN FILE FORMATS (WHEN POSSIBLE)
use open file formats

• Some file formats are less susceptible to obsolescence than others
  – Open, non-proprietary formats
    • pick TXT over DOCX
    • CSV over XSLX
    • TIF over JPG
  – History of wide adoption, backward compatibility
(yikes!)
2. ORGANIZE + NAME FILES (IN MEANINGFUL WAYS)
organize your files

• Lots of possibilities, so consider what makes sense for your project
  – File type
  – Date
  – Type of analysis

EXAMPLE

MyDocuments\Research\Sample12.tiff

vs.

C:\\NSFGrant01234\WaterQuality\Images\LakeMendota_20141030.tiff
name your files

• Use file naming conventions any time you have related files
• Consistent
• Short yet descriptive
• Avoid spaces and special characters

EXAMPLE

File001.xls

vs.

Project_instrument_location_YYYYMMDD.xls
3. DOCUMENT YOUR RESEARCH PROCESS (RIGHT AWAY)
document your research process

Project- & folder-level

• Create a readme file. (Good example located here: http://hdl.handle.net/2022/17155)
• Document any data processing and analyses.
• Don’t forget written notes!

Item-level

• Remember the importance of file names for conveying descriptive information.
so what’s in a good readme file?

• Names + contact information for people associated with the project
• List of files, including a description of their relationship to one another
• Copyright + licensing information
• Limitations of the data
• Funding sources / institutional support

**tl;dr**
Any information necessary for someone with no knowledge of your research to understand and / or replicate your work.
4. BACK UP YOUR DATA (ALWAYS)
storage vs. backup

**storage** = working files.

The files you **access regularly and change frequently**. In general, losing your storage means losing current versions of the data.

**backup** = regular process of copying data *separate from* storage.

You don’t really need it until you lose data, but when you need to restore a file it will be the most important process you have in place.
Rule of 3

- Keep **THREE** copies of your data
  - TWO onsite
  - ONE offsite

- Example
  - One: Laptop
  - Two: External hard drive
  - Three: Cloud storage

- This ensures that your storage and backup is not all in the same place – that’s too risky!
evaluating cloud storage

• Lots of options out there – and not all are created equal
• Read the Terms of Service!
• While at UW, use your free UW Box or Google Drive accounts.
5. HAVE A PLAN (AND STICK WITH IT)
have a plan (and stick with it)

• There is no one perfect plan for managing data - HOWEVER, any plan is better than none at all
• Ideally, set aside time at the beginning of a project to plan for the basics:
  – Where stuff lives + how it is named.
  – Roles and responsibilities. Does everyone have equal access to the data? Do specific people need to do certain tasks? (As projects get bigger, this gets trickier.)
BONUS!

GET CREDIT FOR YOUR DATA (WHY NOT?)
get credit for your data

• You can publish your data!
  – MINDS@UW
  – Figshare
  – Find disciplinary repositories at re3data.org

• Data is not copyrightable; best practice is to apply a Creative Commons 0 license

• There’s even a proven citation advantage to sharing your data*

https://dx.doi.org/10.7717/peerj.175
GOOD DATA MANAGEMENT + SHARING = BETTER, FASTER SCIENCE
have we piqued your interest?

Learn more through these fall workshops:

Project Management + Productivity Tools  
SEPT 24

Crafting Your Digital Identity  
OCT 22

Research Data Management + Sharing  
NOV 19

An Introduction to Open Research  
DEC 10

Steenbock Library BioCommons  |  4-5pm
thanks for listening!

Brianna Marshall
brianna.marshall@wisc.edu

Cameron Cook
cccook3@wisc.edu

Research Data Services
researchdata.wisc.edu

@UWMadRschSvcs