Data Management 101
February 2017

Joanna Thielen | Research Data Librarian | jthielen@oakland.edu
Amanda Nichols Hess | Education Librarian | nichols@oakland.edu
Data are anything you perform analysis on.

A few types of data:

- research notes
- survey responses
- images (such as photographs)
- audio or video recordings
- teaching materials - lesson plans, worksheets, etc.

In Chapter 1: The Data Problem in *Data Management for Researchers* by Kristin Briney
Research data management (RDM) is...

the compilation of small practices that make your data easier to find, easier to understand, less likely to be lost, and more likely to be usable during a project or ten years later... includes data management planning, documenting your data, organizing your data, improving analysis procedures, securing sensitive data properly, having adequate storage and backups during a project, taking care of your data after a project, sharing data effectively, and finding data for reuse in a new project.

In Chapter 1: The Data Problem in *Data Management for Researchers* by Kristin Briney
Why is RDM important?

- You don’t want to lose your data!!!
- Saves valuable time and resources
- Avoids duplication of research (e.g. re-collecting data)
- Comply with federal funding agency requirements
RDM is highly individualized

What works well for your classmate won’t necessarily work for you
How do I find and re-use existing data for my own research?
Common data sources

Previous publications
● Look at article appendices and supplementary files
● Contact the corresponding author

Education sources:
● Data.gov (education section) - “Home of the US government’s open data”
● National Center for Education Statistics

General sources:
● ICPSR - Political and social research, 1960s to present
● Statistical Abstracts of the US* - Social, political and economic data, 1878-1969
● Roper iPoll* - US public opinion poll data
● American FactFinder - Demographic data from censuses and other surveys

*Note: Databases provided by OU Libraries. Must be current OU student, faculty or staff to access.
Being a good data consumer

1. Look at the data’s license - how can (and can’t) I use this data?

Creative Commons Licensing Considerations: https://creativecommons.org/share-your-work/licensing-considerations/
Being a good data consumer

2. Always cite the data you’re re-using - like you would a publication


Quick Guide to Data Citation (including citing data in APA style)
Data Documentation & Organization

How do I document and organize my data so that I can understand it next week or in three years?
Documentation

● Document the who, what, when, where, why and how of your research

● Types of documentation
  ○ Research notes
  ○ Codebooks
  ○ ReadMe.txt files

● Tools for documentation
  ○ EverNote
  ○ Word
  ○ OneNote
Documentation - Amanda’s example

What experiences have most significantly shaped how you think about library/information literacy instruction?

1. My experience in the classroom teaching. I really had no concept of my students' needs until I met with them and saw their assignments, what they responded to, etc.

2. Early on in my career, I was utterly dismayed at the lack of engagement students had with IL instruction. I made it a focus of my teaching, then, to be as engaging as possible.

3. Honestly, the most significant experience that influenced how I now teach information literacy occurred during graduate school, but not in my coursework. I was a student assistant at an academic library and had to update a mostly previously taught workshop for university faculty on using EndNote. This was my first exposure to teaching and I learned a lot from the first experience about session pacing, incorporating interactive activities, and more.

4. Interactions at the research help desk certainly influence the way that I approach my one-shot (WRT 160) instruction sessions. Students’ questions at the desk help me to decide what concepts need more explanation or what skills need more practice.

5. Readings in librarianship journals, especially critical information literacy. I find that IL instruction is all too often based on techniques, not critical thinking.

6. I have found from personal experience that if the students are not able to link the instruction to an immediate task that they need to carry out, they are far less likely to pay attention and/or retain any of the instruction. Also, I have found that even when they are paying attention, they will only retain a small number of learnable “chunks.” These ideas seem to be supported by various workshops and presentations that I have attended relating to library instruction. As a result, I hold the belief that instruction ideally needs to be planned in such a way that it supports an actual assignment that the students should be ready to use the skills being taught to carry out that assignment, and that the instruction time should be spent more deeply with fewer concepts rather than touch lightly on a broad set of concepts.

7. My coworkers, the scholarship of teaching and learning, and the students with whom I work.

8. Reading and reflection on the profession for various research projects has impacted my thinking the most, paired with excellent discussions with colleagues (both casual and professional).
Important takeaway –

BE CONSISTENT with your documentation!

Why? Because YOU are the beneficiary of good documentation!
Organization – create a file naming convention (FNC)

- FNC: standardized naming system that gives each file a unique name
  - Name describes:
    - Contents of file
    - Relation to other files
- FNC makes your files easy to understand and search

Example of a FNC

Source:
Tips for creating a FNC

- **Choose 4-7 elements**
  - Examples: date (YYMMDD), creator, location, project name
- **Avoid spaces** - Some software doesn’t recognize file names with spaces
- **Avoid special characters** - Such as ~ ! @ # $ % ^ & *
  - () {} [] ‘ ”
- **Strike a balance**
  - Too few elements = ambiguous
  - Too many elements = complicated
- **Be brief** - use meaningful abbreviations
- **Track versions using v01, v02, etc.**
  - DON’T use ‘final,’ ‘initial,’ ‘revised,’ etc.

Kennedy_2017_RevEduRes_ProfDevTeach.pdf
How do I store my data so that…

1. It’s secure (because it contains information about human subjects)?
2. It’s safe?
1. Practice the 3-2-1 rule

Example storage plan:

I will keep my data on my personal laptop and back it up on my personal external hard drive and OakShare (files.oakland.edu). My laptop and external hard drive will be stored at my house whereas OakShare is stored on servers at OU.

Best practices

2. Do regular back ups
   - General rule: back ups should occur weekly
   - Test back ups to ensure you can access backed up files

3. Don’t forget about your physical data
   - Examples: observation notes, interview protocols, paper survey responses
   - Create digital copies (photographs, scans, etc.)
Most important thing you will learn today:

Do NOT store all of your research data on a single device!
Confidential data

According to OU’s Information Security Policy, confidential research data includes information related to human subjects (such as personally identifiable information) and needs to be stored more securely than non-confidential data.

OakShare (files.oakland.edu) is the only approved storage location for confidential data (other than local storage, e.g. your laptop).
<table>
<thead>
<tr>
<th>Storage media</th>
<th>Recommended?</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal computer</td>
<td>YES</td>
<td>Prone to theft or loss</td>
</tr>
<tr>
<td>External hard drive</td>
<td>YES</td>
<td>Subject to degradation; lifetime is ~ 5 years</td>
</tr>
<tr>
<td>CD/DVD</td>
<td>YES</td>
<td>Subject to degradation due to mishandling; can be laborious to use</td>
</tr>
<tr>
<td>USB flash drive</td>
<td>NO</td>
<td>Easy to lose; very fallible</td>
</tr>
<tr>
<td>OakShare (files.oakland.edu)</td>
<td>YES</td>
<td>Initial storage limit is 200 MB; Doesn’t sync like some cloud services</td>
</tr>
<tr>
<td>Cloud service (Google Drive, Dropbox, Box, etc.)</td>
<td>NO</td>
<td><strong>Don’t use for confidential data</strong>; Terms of Service may give the company a license to use your files (including data files)</td>
</tr>
</tbody>
</table>
How to protect your confidential data

Digital data

● Don’t leave your devices unattended
● Encrypt your devices
● Password protect files
● Use a password manager (such as LastPass)

Physical data

● Store in locked building, office, filing cabinet/drawer
Platform specific encryption tools

**Windows**

- **Whole drive encryption:** BitLocker Control Panel
- **File/Folder encryption:** Right click on object: Properties>General>Advanced>Compress or Encrypt>Encrypt

**Mac OS**

- **Whole drive encryption:** FileVault - System Preferences>Security & Privacy>FileVault
- **Folder encryption:** Disk Utility app (Mac OS native) can create a folder that presents as an encrypted hard drive (volume) allowing files to be moved into, out of and edited within the encrypted disk image.
RDM resources

- **Research data in education** - guide by University of Michigan Library
- **OU Libraries Research Data Management guide**
- **Research Data Management Best Practices handouts**
- **Librarians!**
  - Joanna Thielen | Research Data Librarian | jthielen@oakland.edu
  - Amanda Nichols Hess | Education Librarian | nichols@oakland.edu
Main Takeaways

Recap of RDM practices

- Find & re-use data as a good data consumer
- Create documentation about your data
- Create a file naming convention
- Storage - back ups!, use OakShare for confidential data
- Ask for help when needed!

Ways to get started with RDM -- don’t be overwhelmed!

1. Storage - start regular back ups!
2. File naming convention
3. Documentation
4. ...
Questions?

Joanna Thielen | Research Data Librarian | jthielen@oakland.edu
Amanda Nichols Hess | Education Librarian | nichols@oakland.edu