Connecting Education Students to Research Data Management

Creating intra-library partnerships and expanding RDM instruction to the social sciences

Joanna Thielen, jthielen@oakland.edu
Amanda Nichols Hess, nichols@oakland.edu
Oakland University Libraries

MI-ALA Conference 2017
Session Goals

Highlight the value of RDM practices to students and faculty in non-STEM disciplines

Identify ways others can develop partnerships around RDM or provide RDM instruction specific groups

Explain how a research data management (RDM) librarian and education librarian partnered to provide targeted resources to doctoral students in education-focused programs

strategy by Karthik Srinivas from the Noun Project
Data is...

**materials** generated or collected during the course of conducting research.
- National Endowment for the Humanities Office of Digital Humanities

**recorded factual material** commonly accepted in the scientific community as necessary to document and support research findings
- National Institutes of Health
Data is...

Anything you perform analysis on.

- *Data Management for Researchers*
  by Kristin Briney
A few types of data

- research notes or lab notebooks
- survey responses
- software and code
- measurements from laboratory or field equipment
- images (such as photographs, films, or scans)
- audio or video recordings
- physical samples
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.
RDM includes...

- Data management planning
- Documenting your data
- Creating metadata about your data
- Organizing your data
- Improving analysis procedures
- Securing confidential data properly
- Having adequate storage and backups
- Taking care of your data after a project
- Sharing data effectively
- Finding data for reuse in a new project
Why is RDM important?

- Save time and resources
- Preserve data
- Maintain data integrity
- Meet grant requirements
- Promote new discoveries
- Support open access & open data initiatives

Most importantly: researchers don’t want to lose their research data!

RDM in academic libraries

● Approx. 2005: Academic libraries started developing data-related services
  ○ Response to forthcoming federal funding agency requirements

● Data librarians provide assistance with:
  ○ Writing data management plans
  ○ Creating documentation & organization systems
  ○ Ensuring data is stored safely and securely
  ○ etc.

● Training and outreach are important components
  ○ Prior focus on STEM disciplines
Project Context & Development
Oakland University (OU), Rochester, MI

Doctoral research institution

Fall 2016 enrollment:
● 20,012 total students
● 3,444 graduate students

Masters / Ph.D programs in:
● Humanities
● Social sciences
● Education
● STEM
● Health sciences/nursing

OU Libraries

● 13 full-time faculty librarians
● Have both liaison responsibilities and library-specific specializations
Data Management at OU Libraries

- Supports 2nd institutional goal in OU’s 2025 Strategic Plan
  - “Be recognized as a strong research and scholarly environment focused on creative endeavors and on the discovery, dissemination, and utilization of knowledge.”
- Research Data Support webpages
  - General RDM guidance
- Workshops on various RDM topics

OU Libraries Research Data Support webpages: https://library.oakland.edu/services/research-data
OU’s School of Education & Human Services (SEHS)

Five academic departments:
- Counseling
- Human Development & Child Studies (Early Childhood / Special Education)
- Organizational Leadership (Educational Leadership / Human Resource Development)
- Reading & Language Arts
- Teacher Development & Educational Studies

Total enrollment, Fall 2016: **2,094 students**
Includes undergraduate, graduate, and professional development programs

Source: Oakland University Office of Institutional Research & Assessment  https://oakland.edu/oira
OU’s Graduate Education in SEHS

Fall 2016: 1,031 graduate students in all degrees / certificates; Ph.Ds awarded in:

- Counseling (F16: 20 enrolled)
- Early Childhood (F16: 32 enrolled)
- Educational Leadership (F16: 57 enrolled)
- Reading & Language Arts (F16: 21 enrolled)

Most SEHS graduate students work full-/part-time, have family obligations, and are returning to school after some period of absence.

Average age (including Ed.S students): 42

Source: Oakland University Office of Institutional Research & Assessment  https://oakland.edu/oira
Library Collaboration around RDM

Amanda: Education Librarian and Ph.D candidate in OU’s SEHS
- Personal knowledge of Organizational Leadership Ph.D program
- Concrete / recognized need for RDM practices for students
- Existing relationships as starting point to pilot RDM in education graduate programs

Joanna: Research Data librarian
- Subject expertise - essential here
- Aligned RDM knowledge, practices, and experience with SEHS graduate educational experiences

Together, this knowledge and experience led to concrete and useful ideas.
Adapting RDM Best Practices for SEHS Students

What is important to these students?

- Ease of integration
- Usefulness
- Concrete, real-world connections
Adapting RDM Best Practices for SEHS Students

**Relevant RDM practices**
- Security of human subjects data
- Storage, documentation
- Organization

**Irrelevant RDM practices**
- External grant requirements
Connecting RDM with doctoral programs in education
Piloting with **One Program**

Beginning with Educational Leadership doctoral students:

- Mapping support into the curriculum to support students at intentional points in their doctoral experience
- Offering standalone workshops to students farther along in the process, but still in need of RDM guidance
First step:

- Identifying students’ first experiences with research (qualitative / quantitative)
  - Introductory-level RDM support targeted to course expectations, assignments, needs, etc.

**Resulting program: Data Management 101**

Second step:

- Building on the introductory RDM support in the last formal class before students work on comprehensive exams
  - Offering more advanced RDM support
  - Providing time to discuss challenges, questions, or successes experienced with RDM so far

**Resulting program: Advanced Data Management**

footsteps by Grace Reeves from the Noun Project
Standalone Workshops Outside of Coursework

Initiative to help students who were farther along the Ph.D process:

- Monthly doctoral support group for students writing comprehensive exams / dissertation proposals / dissertations
  - Offering training following regular meeting time
  - Repeating this training before SEHS evening courses

Low attendance vs. Immediate relevance to students’ work

Idea presentation by Chris Homan and Idea by Michelle Chung from Noun Project
Samples of slides from Data Management 101
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.
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
Best practices

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.

<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>Don’t use for confidential data; Terms of Service may give the company a license to use your files (including data files)</td>
</tr>
</tbody>
</table>
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. ...
Assessing the Impact of Targeted RDM Instruction
Research questions

Do RDM interventions (workshops, training sessions, individual consultations, etc.) by academic librarians affect the RDM practices of doctoral students in education-focused programs?

If so, how do these interventions affect how nascent researchers collect, organize, store, and manage their research data?
This project has been reviewed by our institution’s Office of Research Administration and **classified as not research** under the federal definition of research (i.e. our project didn’t need to be reviewed by IRB).

Research = **systematic investigation with human subjects** that seeks to **create generalizable knowledge**

*Check with your institution to see if this is an option!*
BUT we’re still following ethical practices for human subjects research (protecting PII, etc.)...

And we can still present and publish on this project.
Two Data Sources

Quantitative feedback form, post-interaction

Sent to students after in-class instruction, a workshop, or a 1:1 research consultation

Interviews, baseline and intervention groups

- Baseline interviews about RDM with Ph.D graduates who did not receive RDM support/instruction from the library
- Interviews with intervention group after successful Ph.D defense to understand their RDM practices

Data gathered from these sources will help us to understand the effectiveness of our outreach.
Baseline Group: Common themes

$n=4$; one interview unusable (participant did not meet inclusion criteria)

Commonalities across participants’ experiences:

- All recognized the importance of keeping confidential data protected and took procedures to de-identify data
  - IRB requirements as practical guidelines for **data security, data storage**
- All kept multiple copies of data in multiple formats
- None saw future use for their data
Baseline Group: Different themes

Differences across participants’ experiences:

- Varied formats of information in:
  - Data collection - paper / electronic instruments
  - Storing / capturing information - locked cabinets, electronic folders
- Only one interviewee had a specific file naming convention
Intervention Group: Initial findings

\( n = 4 \), including one faculty member who participated as part of her course

- Hard to find commonalities or differences in this small sample size
- Different sets of skills, knowledges, and intended research approaches color participants’ experiences
- Generally, all attendees found the instructional experiences useful in some way

Going forward, we will continue to ask for feedback from both formal (i.e., class sessions) and informal (i.e., RCs) instructional interactions around RDM.
Conclusions & Key Takeaways
Lessons Learned & Next Steps

- Course-embedded: More meaningful
  - Faculty usually learn new things - useful for students to see them engaged and interested in this topic
  - Faculty can see more places to integrate the library into their courses
- Continue collecting data from baseline group
- Expand RDM instruction to all four education doctoral programs (Fall 2017)
  - Introductory-level RDM support targeted to course expectations, assignments, needs, etc.
- Follow up all interactions with survey
- Prepare for intervention group interviews - Winter 2018 at the earliest
Implementing Targeted RDM at Your Institution

Use institutional expertise!

Where are students collecting or working with data? What RDM practices would be most relevant to them?

- Not just graduate students -- Oakland University Honors College; Business Undergrads
- Not just STEM fields

Remember: data is anything you can perform analysis on!
Implementing Targeted RDM at Your Institution

Use your connections!

With whom can you partner to develop RDM outreach?

- Intra-library partnerships -- RDM librarian, or other librarians who work with students / faculty and data (e.g. Business, STEM)
- Cross-campus partnerships -- offices of institutional research or research support
- Department faculty in liaison areas
- Other ideas?
Questions?

Joanna Thielen, jthielen@oakland.edu
Amanda Nichols Hess, nichols@oakland.edu
Oakland University Libraries