ETDplus as a Bridge to Graduate Scholarly Communications Skills

Sam Byrd, scholarly publishing librarian, sbyrd2@vcu.edu
Nina Exner, research data librarian, nexner@vcu.edu

Virginia Commonwealth University Libraries
Introduction

The VCU Environment

- 30,000 FTE students
  - 6,000 graduate students
  - ca. 425 ETDs per year
- Urban institution in Richmond, VA
- R1 institution: arts, social sciences, engineering, health, & medicine
ETD context

- ETDs published solely in repository
  - Scholar’s Compass (Digital Commons platform)
  - Over 6,000 ETDs online
  - ETDs mandatory since 2009
- Close relationship between graduate school and library around ETDs
Overview of the Educopia ETD+ program

- 6 modules
- Slides with talking points in notes
- One-page handout and longer handout
- Pre/post test

There is a Friday workshop for deeper information, so we won’t get too deep here.
Dissertation and Thesis Planning Camp

- One-day program of all ETD+ modules
- Solicited instructors from all librarians
- Each instruction pair adapted the content however they wanted to
- Used handouts (with edits), not briefs
Dissertation and Thesis Planning Camp

- **Marketing**
  - TelegRam,
  - News blog,
  - webpage / registration,
  - digital posters,
  - Emails / social media through grad school and liaisons

- **Timing**: right before Fall semester
Marketing: Digital and physical posters
Welcome and Workshop Background

- Instructors
  - NAME HERE

- Purpose
  - Provide you with resources and tools to help you address the challenges and opportunities “data organization” methods pose and provide for you as a researcher, particularly regarding your research outputs.

Welcome and Workshop Background

- Instructors
  - Julie Arendt and Nina Exner

- Purpose
  - Provide you with resources and tools to help you address the challenges and opportunities “data organization” methods pose and provide for you as a researcher, particularly regarding your research outputs.

Module customization: before and after example
Module customization: before and after example
Module customization: before and after example

Structuring your data well enables you to:

- Reproduce results
- Reuse it in the future
- Share it with others
- Gain and retain credibility
- Comply with IRB/funder requirements

Structuring your data well enables you to:

- Reproduce results
- Reuse it in the future
- Share it with others
- Gain and retain credibility
- Comply with IRB/funder requirements
- Know what it is when you look at it months or years later!
Module customization: before and after example

**Discipline-based data repositories:**

- Social Sciences: ICPSR  
  [http://www.icpsr.umich.edu/icpsrweb/deposit/index.jsp](http://www.icpsr.umich.edu/icpsrweb/deposit/index.jsp)
- Genomics: GenBank  
- Earth Sciences: NASA’s Earthdata  
  [https://earthdata.nasa.gov/](https://earthdata.nasa.gov/)
- Archaeology: tDAR  
- Oceanography: NODC  
- BioSciences: Dryad  
  [https://datadryad.org/](https://datadryad.org/)

**Discipline-based data guidance:**

- Social Sciences: ICPSR  
  [http://www.icpsr.umich.edu/icpsrweb/deposit/index.jsp](http://www.icpsr.umich.edu/icpsrweb/deposit/index.jsp)
- Genomics: GenBank  
- Arts: GloPAD Global Performing Arts database  
- Earth Sciences: DataOne  
  [https://www.dataone.org/](https://www.dataone.org/)  
  and NASA’s Earthdata  
  [https://earthdata.nasa.gov/](https://earthdata.nasa.gov/)
- Archaeology: tDAR  
- Oceanography: NODC  
Data Organization

Structuring your data well enables you to:
- Reproduce results
- Reuse it in the future
- Share it with others
- Gain and retain credibility
- Comply with IRB/funder requirements

Whether your data is organized in lists, arrays, hash sets, dictionaries, queues, trees, heaps, or relational databases, it is important to be aware of disciplinary norms, as well as both institutional and funder requirements, that will make its deposit, storage, and long-term support more likely. Increasingly, the path for long-term support involves taking steps to make sure your data is deposited alongside data collected by others in your field or discipline.

Questions to consider for any data project:
1. What are the data organization standards for your field?
2. What are the data export options for your software?
3. What forms of the data will be needed for future access?

Do:
- Consider what your NULL values are and how they are represented
- Consider what data documentation is required
- Use standard data representations (e.g., YYYYMMDD for dates)

Do Not:
- Use formatting to convey information
- Place comments in cells
- Use special characters in field names
- Use blank spaces or symbols in column names

Discipline-based data repository examples:
- Social Sciences: ICPSR
- Genomics: GenBank
- Earth Sciences: NASA’s Earthdata
- Archaeology: IDAR
- Oceanography: NODC
- BioSciences: Dryad

The DataONE Best Practices database provides individuals with recommendations on how to effectively work with their data through all stages of the data lifecycle. 
https://www.dataone.org/best-practices
Data Organization

Structuring your data well enables you to:
- Reproduce results
- Reuse it in the future
- Share it with others
- Gain and retain credibility
- Comply with IRB/funder requirements

Whether your data is organized in lists, arrays, hash sets, dictionaries, queues, trees, heaps, or relational databases, it is important to be aware of disciplinary norms, as well as both institutional and funder requirements, that will make its deposit, storage, and long-term support more likely. Increasingly, the path for long-term support involves taking steps to make sure your data is deposited alongside data collected by others in your field or discipline.

Context and Data Documentation:
Include the following in a readme text file:
1. The data’s purpose
2. A list of the files in your data package
3. Data dictionary listing and describing all variables

Data Organization Principles:
1. Use one variable per column
2. Make one observation per row
3. Use human-readable column name
4. Include one table per tab
5. Include an ID or key to indicate any relationship between tables

Questions to consider for any data project:
1. What are the data organization standards for your field?
2. What are the data export options for your software?
3. What forms of the data will be needed for future access?

The DataONE Best Practices database provides individuals with recommendations on how to effectively work with their data through all stages of the data lifecycle. [https://www.dataone.org/best-practices](https://www.dataone.org/best-practices)

Do:
- Consider what your NULL values are and how they are represented
- Consider what data documentation is required
- Use standard data representations (e.g., YYYYMMDD for dates)

Do Not:
- Use formatting to convey information
- Place comments in cells
- Use special characters in field names
- Use blank spaces or symbols in column names

Discipline-based data repository examples:
- Social Sciences: ICPSR [http://www.icpsr.umich.edu/ICPSRweb/deposit/index.jsp](http://www.icpsr.umich.edu/ICPSRweb/deposit/index.jsp)

PDF customization example: Modified version
Activity

• Choose one spreadsheet you are using for a current data-gathering project.
  o Use the “Data Organization Principles” and check to see if your file meets those requirements.
  o Create a data dictionary for the spreadsheet that describes the meaning of each column header.
Activity

• Choose one spreadsheet or other data file you are using for a current data-gathering project.
  ○ Use the “Data Organization Principles” and check to see if your file meets those requirements.
  ○ Create a data dictionary for the file:
    ■ For a spreadsheet, the data dictionary describes the study and data collection, plus defines each column header’s meaning, and explains column relationships.
    ■ For other files this dictionary may describe the study, how the data were collected, has a file list with human-readable filenames, and describes file contexts.

If you don’t have your own file, use the file at http://tinyurl.com/PlanningCampData
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sheet</td>
<td>Meme name</td>
<td>Alternative narr</td>
<td>Type</td>
<td>URL</td>
<td>Animal</td>
<td>Caption</td>
<td>Year</td>
<td>Started</td>
</tr>
<tr>
<td>2</td>
<td>Meme1</td>
<td>Grumpy Cat</td>
<td>none</td>
<td>Photo</td>
<td><a href="http://know">http://know</a> Cat</td>
<td>Varies</td>
<td>2012</td>
<td>Reddit</td>
<td>Negative</td>
</tr>
<tr>
<td>3</td>
<td>Meme2</td>
<td>Nyan cat</td>
<td>Poptart cat</td>
<td>Animated</td>
<td><a href="http://know">http://know</a> Cat</td>
<td>None</td>
<td>2011</td>
<td>LOL-COMICS</td>
<td>Positive</td>
</tr>
<tr>
<td>4</td>
<td>Yes this is dog</td>
<td>Photo</td>
<td><a href="http://know">http://know</a> Dog</td>
<td>Hello Yes, This is Dog</td>
<td>2011</td>
<td>Tumblr</td>
<td>Neutral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Haters gonna h -</td>
<td>Photo</td>
<td><a href="http://know">http://know</a> Var.</td>
<td>HATERS GONNA HATE</td>
<td>2009</td>
<td>yayhooray.y.c</td>
<td>Positive?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Honey badger</td>
<td>none</td>
<td>Video / varies</td>
<td><a href="http://know">http://know</a> Honey badger Var.</td>
<td>2011</td>
<td>Youtube</td>
<td>Neutral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Socially awkward penguin</td>
<td>Modified photo</td>
<td><a href="http://know">http://know</a> Penguin</td>
<td>SOCIALLY AWKWARD PENGUIN</td>
<td>2009</td>
<td>4chan</td>
<td>neutral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Meme7</td>
<td>Insanity wolf</td>
<td>none</td>
<td>Modified photo</td>
<td><a href="http://know">http://know</a> Wolf</td>
<td>Varies</td>
<td>2009</td>
<td>?</td>
<td>Negative</td>
</tr>
<tr>
<td>9</td>
<td>Meme8</td>
<td>Courage wolf</td>
<td>none</td>
<td>Modified photo</td>
<td><a href="http://know">http://know</a> Wolf</td>
<td>Varies</td>
<td>2006</td>
<td>4chan</td>
<td>Positive?</td>
</tr>
<tr>
<td>10</td>
<td>Meme9</td>
<td>Happy cat</td>
<td>none</td>
<td>Photo</td>
<td><a href="http://know">http://know</a> Cat</td>
<td>None</td>
<td>2003</td>
<td>Something A</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Activity content for attendees without ETDs in process
Pros of the Planning Camp

- ETD+ material flexible, adaptable
  - Works out of the box
  - Scripts/briefs useful (even if not used verbatim)
  - A great beginning: getting students thinking about what they don't know they should be thinking about
Pros of the Planning Camp

- Validation: Made grad students feel like they were valued
- Filled a gap
- Good learning experience for the teachers, too (great way to engage librarians with scholcom topics)
User Results
Round One

- Higher turnout than other trainings!
- Approval 4.1 / 5, but low response rate
- Suggested less repetition, data earlier, more content beyond scholcomm topics
- Pre/Post were all correct
  - Except for some that were blank 😞
Lessons Learned
Round One

● Framing the purpose of each session within the ETD context helps
● More deduplication needed when presenting all content
● Sessions could be shorter, or activities could be longer
Changes for Fall 2019 Camp

- Introduction to dissertation/thesis planning
- Filing your research in Scholar's Compass
  - Metadata for discoverability
  - Copyright for dissertation authors
- Accessibility for ETDs
  - File storage and data organization
- Ethics Breakout sessions – animal or human
- File versioning and file types
User results
Round Two

- Rated 3.9 / 5 overall, still low response
- Still appreciate the event and the focus on planning
- Want more integration across topics
- Hoping for technical info on formatting
- Favorites: Submission overview, storage and data, copyright
Next steps

- More integration
  - Across topics
  - Between Planning Camp and Writing Retreat
- Maybe move sharing content like Data and File Types to the Writing Retreat
- Input from subject specialists
Questions and discussion?