Using the NECDMC Case Studies to Teach Scientific Research Data Management

Donna Kafel
University of Massachusetts Medical School, Donna.Kafel@umassmed.edu

Follow this and additional works at: http://escholarship.umassmed.edu/lib_articles
Part of the Information Literacy Commons, and the Scholarly Communication Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 License.

Repository Citation
http://escholarship.umassmed.edu/lib_articles/163

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Library Publications and Presentations by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Using the NECDMC Case Studies to Teach Scientific Research Data Management

ACRL NEC Scholarly Communication Interest Group
Teaching RDM with NECDMC workshop
May 8, 2014

Donna Kafel, RN, MLIS
e-Science Project Coordinator
Lamar Soutter Library
University of Massachusetts Medical School
Overview of Presentation

• Value of Case Studies
• Tour of Research Cases section of NECDMC
  • Simplified Data Management Plan
  • Teaching guides/resources
• Ways to teach with cases
• A sampling of research cases
• Selecting cases
• Summary
• Q & A
Case Study Approach

“The case method packs more experience into every hour of learning than any other instructional approach.”

Harvard Business Publishing, Hints for Case Teaching
An Overview of NECDMC’s Research Cases page
Approaches to Teaching with Cases

- Flipped classroom
- Integrated into lectures
- Group work and breakout sessions
- Homework
- Activities in module lesson plans
- One case over series of classes
A Sampling of Cases

Habitat Selection by Grassland Birds, Their Habitat Requirements and Informing Grassland Bird Conservation

Regeneration of Functional Heart Tissue in Rats

Characterizing a Component of a Rocket Engine used to Control Satellites in Orbit

Data Preservation Case Study: The ALIEnS Project
Using above data, student creates simulations that create data models that in turn, create more data.
Grassland Birds Case
Data Formats

DMP #1: Types of Data, Instruments, software

- Hard copies entered onto computer
- Used Juno Trimble
- Excel sheets ArcGIS
- Analyzed data with SAS
- Downloaded GPS data
- Processed downloaded data with Python and SAS

DMP #2: File formats

- .doc files for data descriptions and write-ups
- .xls and .xlsx files for spreadsheets and data tables
- .shp files for geographic information with associated data tables
- .csv files for both data tables and geographic info, these are easier to read with R and Python than .xls, .xlsx, and .shp files
- py, .sas, and .r files contain scripts that processed and analyzed other data or computed statistics
Grassland Birds Case
Data Storage

• Hard copy data sheets: Binders (binders for self, and copies of binders in PI’s lab)
• Electronic files: lab’s space on institutional server and external hard drive (both backed up)
Grassland Birds
Metadata

No formal protocol

• Data guide created for some folders (guide files start with 0 are listed at top of directory)

• Not every folder has a guide file and existing guide files not kept up to date
Grassland Birds Case
Data Sharing and Citation

The student recognizes that he is benefitting from the availability of public data, which he is repurposing for his own research, and he would also like to share his data. There are no formal restrictions placed on his data for reuse, but he feels that he would like to be informed if someone else wanted to reuse his data, which he feels is the unspoken rule in his discipline.
Regeneration of Functional Heart Tissue with Stem Cell Delivery

• Biomedical lab case
• Animal model experimentation
• Data related to observations of function of heart muscle after stem cell implementation
• Managing multiple formats of data
• Ethics related to use of animals in research
Regeneration of Functional Heart Tissue with Stem Cell Delivery
Regeneration of Functional Heart Tissue with Stem Cell Delivery

<table>
<thead>
<tr>
<th>Data</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Images</td>
<td>?</td>
</tr>
<tr>
<td>Left ventricular pressure measurements</td>
<td>?</td>
</tr>
<tr>
<td>Home made software</td>
<td>MATLAB or C</td>
</tr>
<tr>
<td>Slides—file name based on stain—example .act is actinin stain</td>
<td>Slides—file name based on stain—example .act is actinin stain</td>
</tr>
<tr>
<td>Contextual</td>
<td>Paper lab notebook, animal log</td>
</tr>
<tr>
<td>Data Set</td>
<td>Storage</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Optical images taken during Surgery: (~10,000 images) 1000 images for each data set</td>
<td>Hard drive acquisition computer &gt; Drobo backup &gt; hard drive of network computer backed up by institution</td>
</tr>
<tr>
<td>Left ventricular pressures (numeric) correlating with specific images</td>
<td>Same as above</td>
</tr>
<tr>
<td>Tissue sections</td>
<td>Slide boxes—could be in any of 3 or 4 freezers</td>
</tr>
<tr>
<td>Software</td>
<td>?</td>
</tr>
<tr>
<td>Images from different stained tissue after second surgery</td>
<td>Drobo &gt; DVD backups</td>
</tr>
<tr>
<td>Contextual data</td>
<td>Paper lab notebook (lab, PI’s office), surgical log (with animal)</td>
</tr>
</tbody>
</table>
Characterizing a Component of a Rocket Engine used to Control Satellites in Orbit

An Engineering Testing Lab Case

- Security restrictions
- Legal and ethical issues
- Data sharing issues
- Homegrown analysis software

Fig. 1 50” x 72” vacuum chamber used for cathode research (WPI Higgins Laboratory)
A Tale of Two Cathodes

Cathode 1
- From a private company
- Developed with Air Force funding
- ITAR regulations

Cathode 2
- From NASA
- A laboratory use model
- No ITAR regulations
- Restrictions on publications related to cathode design

DMP #6
Policies for access and sharing
Data in the Cathode Testing Lab

- Raw data → a laptop
- Experimental conditions → laboratory notebook
- Students review data—discard some
- Student produce code using MATLAB
- Students share findings with researcher
Large interdisciplinary research collaborative interested in preserving data, documents and supporting documents for reuse

Diverse data—collection consists of multiple iterations and minimal metadata

No preservation planning conducted

PI and project manager approach library for assistance and possible submission of data and documents to IR
Points to consider when selecting a case(s):

- Specific data management issues to be addressed
- Your audience
- Type of activity: class discussion, homework, breakout groups
Time crunch? Consider case excerpts.....

• Case Excerpt of Regeneration of Functional Heart Tissue in Rats

• Module 2, Activity 1: case study of future climate changes and flooding in Mass. coastal communities

• Design your own!
Conclusion

• Advantage of teaching with research cases
• Flexible use
• Cases from diverse disciplines
• Each case has a summary of teaching points and discussion questions
• Full cases or excerpts
• Create your own