Data Management... in Writing Studies?
A case study of collaboration and outreach

Alicia Hofelich Mohr, Ph.D.
University of Minnesota

Alice Motes, Ph.D.
University of Surrey
University of Minnesota

18 Colleges

12,000 Graduate Students

4,000 Faculty

> 150 departments
Previous Data Management Outreach

Data Management Workshops
2011-2015
Students & Faculty, ~800 reached
College known for 415

Data Management Course
Summer 2015
Graduate Students, 84 applied
Primarily Attracts Sciences & Health

Data Management Workshops
2011-2015
Students & Faculty, ~800 reached
College known for 415

<table>
<thead>
<tr>
<th>COLLEGE</th>
<th>% OF KNOWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science and Engineering</td>
<td>37.6</td>
</tr>
<tr>
<td>Education and Human Development</td>
<td>13.7</td>
</tr>
<tr>
<td>Academic Health Center</td>
<td>12.5</td>
</tr>
<tr>
<td>Food, Agriculture, and Natural Sciences</td>
<td>11.8</td>
</tr>
</tbody>
</table>

Liberal Arts 10.4

<table>
<thead>
<tr>
<th>DEPARTMENT AREA</th>
<th>PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Sciences</td>
<td>26.2</td>
</tr>
<tr>
<td>Natural Sciences</td>
<td>16.7</td>
</tr>
<tr>
<td>Computer Science/Engineering</td>
<td>16.7</td>
</tr>
<tr>
<td>Agriculture Sciences</td>
<td>15.5</td>
</tr>
<tr>
<td>Health Sciences</td>
<td>11.9</td>
</tr>
<tr>
<td>Humanities</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Data Management Course
Summer 2015
Graduate Students, 84 applied
Many Reasons for selective attendance

• Science & Health fields are where data sharing/management are more normative

• More grants, more requirements

• The way we advertise “Data Management” workshops may not appeal to all
Many Reasons for selective attendance

Do you work with “data” or “research materials”?

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Research Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberal Arts</td>
<td>54%</td>
<td>46%</td>
</tr>
<tr>
<td>Science &amp; Engineering</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>Food &amp; Agriculture</td>
<td>91%</td>
<td>9%</td>
</tr>
</tbody>
</table>

What we did... Initial contact

Identify courses → Targeted Emails → Content Meeting
A diverse team for diverse research

- Quantitative Data Manager (CLA)
- Qualitative Data Curator (Libraries)
- Subject Liaisons (Libraries)

Brought in others as needed (such as the Mixed Methods Research Associate)
Response from faculty - Fall 2015

• Methods loosely defined, included prosem
• Before classes began:
  • Emailed 20 classes in 10 departments
  • Result: 11 presentations in 7 departments

• Variety of disciplines (Statistics, Journalism, Writing Studies, Gender, Women, and Sexuality Studies, Curriculum & Instruction, Kinesiology, Family Social Science)

• For most classes:
  • Met instructor to discuss content
  • 1 presentation for each class
Integrating into classes

Format

• Drop-ins (20-30min)
• Full length classes

Content

• Presentations based on content meeting
• Ad-libbed faculty contributions of RDM stories

Benefits

• Fill in class time (win-win!)
• Expanded expectations about what services Libraries/IT provide
Example: Journalism Classes

Quantitative Methods

• Replication as reason to care
• Reproducible workflows for data analysis
• Protecting, organizing, and sharing data

Qualitative Methods

• Organizing & analyzing textual materials as reason to care
• Qualitative analysis tools = data management tools
• Protecting, organizing, and sharing data

Open Science Collaboration; Nosek et al., (2015)
Example: Statistics Graduate Class

Masters students, not going into academia.

- Future jobs & data sharing
- Reproducible analysis workflows
- Using R for reproducible data analysis
  - Documentation
  - Scripting best practices
  - Integrating interpretation & analysis - R Markdown/Sweave
- Protecting, organizing, and sharing data
Example: Writing studies

Humanities-esque research
Two half classes; in syllabus

First session
- RDM principles and DMPs
- Individual meetings
- Coursework: Write a DMP

Second session
- Review DMPs
- More tech-focused RDM advice
Writing studies: First session

• Covered:
  • What is RDM
  • What do we mean by “data”
  • Introduced DMPs for research
  • Mixed methods focus
  • Archiving and preservation

• Sign up sheet for individual meetings

• Coursework: Create a DMP for your project
Writing studies: Meetings

- 12 students; 45-60 minute
- DMP structure as guide
- Workflow walkthrough
- Focused on data collection, documentation

- Diverse methods and data types
  - Text mining copyrighted book
  - Content analysis of video
  - Discourse analysis of images on web
  - Sentiment analysis
  - Interviews
  - Surveys
Writing studies: Second session

In-depth knowledge guided instruction

• High level review of DMPs:
  • What they did well
  • What could be improved

• Confusion about Archive vs. Storage - clarified

• Based on meetings and DMPs:
  • Tools, technology to facilitate active data
Current progress

• Continued outreach in Spring 2016
  • Reached 7 more CLA classes, 7 departments

• Positive word of mouth
  • DMP consultations
  • Graduate group consultations
  • Invitations to other meetings in the department
  • Working into the graduate curriculum

• Closer collaboration with Library Liaisons
Methods classes: Fertile ground for RDM

• RDM is for everyone, but language matters
• 3-way collaboration benefits us and them
• Diverse backgrounds = agile offerings
• Goodwill and salience
• Is an “in” to some departments
• Ripple effect; lasting impact on service
• Time and resource intensive
• Everyone has to be on board (even on our side)
Thank you!

Questions?

Contact us:
Alice Motes  a.motes@surrey.ac.uk
Alicia Hofelich Mohr  hofelich@umn.edu