Data Librarianship and Small (& Large) Collections Support

SPNHC 2016
Brian Westra
Science Data Services Librarian
University of Oregon
Survey question

Go to www.menti.com and use the code 56 57 16

https://www.menti.com/77e300

Results
Why libraries are involved in Research Data Services

Expertise and philosophy similar to other services for scholarly communications/research products:

• Organize
• Describe
• Preserve
• Provide access
Data services and libraries

• Data Management Plan (DMP) consultations
• Data information literacy
• Infrastructure for organizing, preserving and sharing data
• Metadata support
• ‘Curation-friendly’ tools and practices that facilitate data sharing and preservation
NSF Data Management Plan requirement - 2011

Some of the challenges for proposal writers:
• Data and metadata formats and standards
• Where and how to share data
• Backup and storage ≠ Preservation
## How authors propose to share their data

<table>
<thead>
<tr>
<th>Method</th>
<th>All</th>
<th>BIO</th>
<th>CISE</th>
<th>ENG</th>
<th>GEO</th>
<th>MPS</th>
<th>SBE</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal / supplement</td>
<td>36</td>
<td>27</td>
<td>23</td>
<td>45</td>
<td>35</td>
<td>54</td>
<td>18</td>
<td>80</td>
</tr>
<tr>
<td>Data center or repository</td>
<td>34</td>
<td>75</td>
<td>14</td>
<td>8</td>
<td>66</td>
<td>25</td>
<td>42</td>
<td>70</td>
</tr>
<tr>
<td>On request</td>
<td>30</td>
<td>23</td>
<td>30</td>
<td>38</td>
<td>24</td>
<td>34</td>
<td>26</td>
<td>60</td>
</tr>
<tr>
<td>Personal website</td>
<td>25</td>
<td>13</td>
<td>44</td>
<td>31</td>
<td>25</td>
<td>20</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td>Other method</td>
<td>22</td>
<td>27</td>
<td>30</td>
<td>15</td>
<td>23</td>
<td>18</td>
<td>22</td>
<td>40</td>
</tr>
<tr>
<td>Institutional repository</td>
<td>17</td>
<td>6</td>
<td>12</td>
<td>20</td>
<td>6</td>
<td>28</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>Conference / proceedings</td>
<td>13</td>
<td>8</td>
<td>11</td>
<td>23</td>
<td>8</td>
<td>13</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>Did not specify</td>
<td>8</td>
<td>0</td>
<td>18</td>
<td>9</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>Thesis / Dissertation</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Not planning to share</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Book</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

DMPs as a Research Tool (DART): [https://osf.io/kh2y6/files/](https://osf.io/kh2y6/files/)
<table>
<thead>
<tr>
<th>ID</th>
<th>Complete / detailed</th>
<th>Addr., but incomplete</th>
<th>Did not address</th>
<th>Most common metadata standards (N)</th>
<th>Yes (%)</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>19.4</td>
<td>23.7</td>
<td>57.0</td>
<td>Dublin Core (23), CSDGM/ISO 19115 (12), EML (11)</td>
<td>14.9</td>
<td>60</td>
</tr>
<tr>
<td>BIO</td>
<td>38.5</td>
<td>21.2</td>
<td>40.4</td>
<td>EML (8), Darwin Core (5), CSDGM/ISO 19115 (4), Dublin Core (3)</td>
<td>38.5</td>
<td>50</td>
</tr>
<tr>
<td>CISE</td>
<td>9.1</td>
<td>28.8</td>
<td>62.1</td>
<td>DIF (2), Dublin Core (1), EML (1)</td>
<td>6.1</td>
<td>40</td>
</tr>
<tr>
<td>ENG</td>
<td>15.1</td>
<td>30.2</td>
<td>54.7</td>
<td>Dublin Core (7)</td>
<td>9.4</td>
<td>30</td>
</tr>
<tr>
<td>GEO</td>
<td>18.1</td>
<td>18.1</td>
<td>63.9</td>
<td>CSDGM/ISO 19115 (3), EML (2), ACADIS (2)</td>
<td>12.2</td>
<td>20</td>
</tr>
<tr>
<td>MPS</td>
<td>18.8</td>
<td>22.4</td>
<td>58.8</td>
<td>Dublin Core (9), FITS (3)</td>
<td>15.3</td>
<td>10</td>
</tr>
<tr>
<td>SBE</td>
<td>26.0</td>
<td>16.0</td>
<td>58.0</td>
<td>CSDGM (5), ODM (2)Dublin Core (2)</td>
<td>16.3</td>
<td>0</td>
</tr>
</tbody>
</table>

*Content Standard for Digital Geospatial Metadata; ^Ecological Metadata Language; *Directory Interchange Format; Advanced Cooperative Arctic Data and Information Service; ^File Information Tool Set; ~Observations Data Model

DMPs as a Research Tool (DART): [https://osf.io/kh2y6/files/]
Bridging communities

DMP consultations | Other conversations

Collaboration | Instruction/training
Managing Images in an Evo Bio team

• Stickleback fish

• Wanted workflow-friendly tool to improve:
  • Organizing, describing, sharing, and analyzing images

• UO + partners at UAlaska Anchorage
OMERO – Open Microscopy Environment

• Server + Desktop Client
• Virtual folder and file structure to organize images
• Tags and other descriptive information
• ImageJ integration
• BioFormats integration
OMERO – Outcomes

• ~1000 images in the system
• Less time spent looking for an image
• Share images and views of images across multi-institutional team
• Add tags and other descriptive information
• BioFormats integration – handles proprietary formats
OMERO – Lessons learned

• Decentralization
  • Independence comes with a cost
• Champions act within trusted relationships
• Unstructured metadata – ease of use but costs later on
Managing collections data

Biodiversity research collection

• Insect - fungi - substrate interactions

• Needs:
  • Data sharing
  • Image preservation
Managing collections data

Using Specify to:

Organize collection objects
Manage structured metadata
To publish data to iDigBio
Next steps?

Archive images in repository
Managing collections data – Lessons

Transition to a new system takes time

Leverage external support

• iDigBio sponsored Specify training

• Specify team for schema/mapping, Specify Cloud (hosting)
Many other opportunities

• Curriculum-integrated data information literacy
• Follow-up on DMPs: intention > practice
• Build/share infrastructure-suite of tools and services:
  • Active data processes;
  • Storage and backup;
  • Post-project data preservation and sharing
Thank you!

Brian Westra | bwестra@uoregon.edu

http://library.uoregon.edu/datamanagement

This presentation at SPNHC thanks to support from a University of Oregon Foundation/Solari Incentive award and iDigBio

DART project funded by Institute of Museum & Library Services grant number LG-07-13-0328