Open Access & Data Management Are Do-Able Through Partnerships

ASERL Summertime Summit
August 6, 2013
Sayeed Choudhury
Remember Open Access & Institutional Repositories?

• Did we go through the Gartner Hype Cycle?
• What happened to transforming scholarly communication?
• Or reinventing libraries?
• Lots of institutions installed institutional repositories
• But we didn’t build infrastructure, which is perhaps more about when, rather than what
Data Management

• ARL recently released Spec Kit 344 – “Research Data Management Services”

• Three broad categories
  – Consultation
  – Data management plan preparation
  – Actual data archiving, preservation, curation

• Every phase is important, but there is a sequential nature to them
Why?

- Is it required that your library provide research data management services?
- If the demand isn’t there, then perhaps not
- But it’s clear the Federal Government is paying more attention to data management
- It’s also important to be prepared in case researchers start to ask for support
<table>
<thead>
<tr>
<th>Layers</th>
<th>Characteristics</th>
<th>Implication for PI</th>
<th>Implication relative to NSF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Curation</strong></td>
<td>• Adding value throughout lifecycle</td>
<td>• Feature Extraction</td>
<td>• Competitive advantage</td>
</tr>
<tr>
<td></td>
<td>• New query capabilities</td>
<td>• New query capabilities</td>
<td>• New opportunities</td>
</tr>
<tr>
<td></td>
<td>• Cross-disciplinary</td>
<td>• Cross-disciplinary</td>
<td></td>
</tr>
<tr>
<td><strong>Preservation</strong></td>
<td>• Ensuring that data can be fully used and interpreted</td>
<td>• Ability to use own data in the future (e.g. 5 yrs)</td>
<td>• Satisfies NSF needs across directorates</td>
</tr>
<tr>
<td></td>
<td>• Data sharing</td>
<td>• Data sharing</td>
<td></td>
</tr>
<tr>
<td><strong>Archiving</strong></td>
<td>• Data protection including fixity, identifiers</td>
<td>• Provides identifiers for sharing, references, etc.</td>
<td>• Could satisfy most NSF requirements</td>
</tr>
<tr>
<td></td>
<td>• Backup and restore</td>
<td>• Provides identifiers for sharing, references, etc.</td>
<td></td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>• Bits on disk, tape, cloud, etc.</td>
<td>• Responsible for:</td>
<td>• Could be enough for now but not near-term future</td>
</tr>
<tr>
<td></td>
<td>• Restore</td>
<td>• Restore</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Sharing</td>
<td>• Sharing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Staffing</td>
<td>• Staffing</td>
<td></td>
</tr>
</tbody>
</table>
Creating data infrastructure

and

ds,

Figure from F. Millerand based on S. L. Star & K. Ruhleder (1996)
Resources

• Many libraries have developed helpful, comprehensive web resources (e.g., MIT, Penn State)
• International resources such as Digital Curation Centre in the UK and the Australian National Data Services in Australia
• JHU Data Management Services - http://dmp.data.jhu.edu
• The DMPTool (https://dmp.cdlib.org) - but think of the “reference interview” implications
• Software (e.g., HubZero, Dataverse or Data Conservancy)
Another Important Resource

• All of you!
• Each of you can tap into expertise, projects, relationships and...
• ...Partnerships that you already have on your campuses
• Everyone in this room is part of the social aspects – perhaps the most important cultural aspect – of infrastructure development
Acknowledgements

• Mark Parsons for (original) “Creating data infrastructure” image

• Carole Palmer, Karen Baker, Timothy DiLauro for the Data Management Layers Stack