It Takes a Village: The Collaborative Construction of a Data Management Infrastructure

ANNE C. ELGUINDI
AMERICAN UNIVERSITY

ALA ANNUAL, 2012
ALCTS METADATA INTEREST GROUP/LITA
Why Libraries Love and Hate Data

- The “data” format splinters quickly into many more formats
- Data is meaningless without metadata
- There is an inextricable tie to specialized software and skills
- Vendors of individual data sets are not well adapted to library licensing needs yet
Why Libraries Love and Hate Data

- You might not really know what you’re buying until you’ve already bought it
- It is an increasingly necessary component of research, but data literacy remains fairly low
- Although data is an information resource, the campus does not necessarily think of the library first
Local Factors at American University

- Private, four-year university in Washington, DC
- Around 11,000 FTE, half of which are graduate students
- Heavy emphasis on the social sciences, particularly international studies and public affairs
- Primary library systems support is through the Washington Research Library Consortium, a partnership among nine university libraries
Three Partnerships

- Hosting data
- Data workflow
- Data management plans
Partnership One: Hosting Data

- Issues of storage, access, and discoverability

- Office of Information Technology
- University Library
- Washington Research Library Consortium
- Center for Research, Teaching & Learning
Creating a Mutual Understanding

- **Storage:** Sharepoint vs. Shared campus drive vs. WRLC server
- **AU-Only Access:** Institutional login vs. Proxy
- **Discoverability:** Institutional resource vs. Library catalog and Summon
The Birth of Euripides

• A virtualized linux server with 100G storage and 2G of RAM
Cross-national Time-series Data Archive

[electronic resource].

Availability

STATUS: Available
Location: AU: Data Sets
Call Number: AU Data Sets
Number of Items: 1
Linked Resources: Click here to access CNTS Data Archive

Details

Title: Cross-national time-series data archive [electronic resource].
Description: 1 CD-ROM ; 4 3/4 in.
Subjects: Economic indicators --Databases.
Social indicators --Databases.
Political indicators --Databases.
Economic history --Statistics --Databases.
Social history --Statistics --Databases.
World politics --Statistics --Databases.
### Index of /Licensed/Cross-National_Time-Series_Data_Archive

<table>
<thead>
<tr>
<th>Name</th>
<th>Last modified</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent Directory</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Bibliography.xls</td>
<td>05-Oct-2011 12:18</td>
<td>2.4M</td>
<td></td>
</tr>
<tr>
<td>CNTSDATA.xls</td>
<td>05-Oct-2011 12:18</td>
<td>12M</td>
<td></td>
</tr>
<tr>
<td>Codebook.xls</td>
<td>05-Oct-2011 12:18</td>
<td>48K</td>
<td></td>
</tr>
<tr>
<td>Independent States Since 1815.xls</td>
<td>05-Oct-2011 12:18</td>
<td>49K</td>
<td></td>
</tr>
</tbody>
</table>
Partnership Two: Data Workflow

- Issues of licensing and workflow clarity

- WIPAR and the General Faculty
- Acquisitions
- Collection Management
- Cataloging
- Electronic Resource Management
Communication Issues

- With the faculty:
  - Data is a high demand area often related to grants and grant proposals
  - The library’s collection development policy is to purchase resources that can be accessed campus-wide – creates a conflict with typical data licenses

- Within the library:
  - Who is responsible? Differentiated by continuing vs. one-time?
  - New format brings up new description and access issues
Data Workflow

- Application of CORAL, an open source ERMS developed by the University of Notre Dame
Data Workflow: Selection
Data Workflow: Technical Services

Edit Workflow

Resource Entry Requirements

- Acquision Type: Paid
- Format: Electronic
- Type: Data set

Workflow Steps

- Name: Funding Approval Approval/Notification Group: Acq Funding Approval
- Name: Confirm Pricing, Request License Approval/Notification Group: IDS Director
- Name: Licensing Approval/Notification Group: IDS Director
- Name: Place Order with Vendor Approval/Notification Group: IDS Director
- Name: Request Invoice from Vendor Approval/Notification Group: IDS Director
- Name: Create Firm Order in Voyager Approval/Notification Group: IDS Director
- Name: Transfer Funds from BKS Approval/Notification Group: Acq Librarian
- Name: Load Data on to Euripides Server Approval/Notification Group: IDS Director
- Name: Receive Data on Purchase Order Approval/Notification Group: IDS Director
- Name: Create Invoice in Voyager Approval/Notification Group: Acq Receiving
- Name: Catalog Record and Holdings Approval/Notification Group: Cataloging
- Name: Access Approval/Notification Group: ERM Unit

Parent Step:

Add
The audata Location Code

- All data and statistics databases were moved to a new audata location code within Voyager
- This location code was shared with individually-licensed data sets

>Search the Library’s Data Holdings
Partnership Three: Data Management Plans

- Issues of legal and ethical issues, university liability, technological infrastructure

Institutional Review Board
Office of Information Technology
Risk Management
Office of Sponsored Programs
Center for Research, Teaching & Learning
Università Library
Grant-Seeking Faculty
Washington Research Library Consortium
Navigating the Politics

- Contacted campus partners to arrange a meeting
- Presentation to the Academic Technology Steering Committee
- Approval to create guide
Data Management Plans

This is a guide for researchers applying to grants that require data management plans or data sharing plans, but it also contains general information about data management practices.

Last Updated: May 31, 2012
URL: http://subjectguides.library.american.edu/dmp

What is a Data Management Plan?

Many federal agencies, including NIH and NSF, now require grant applications to include data management or sharing plans for projects that involve the collection or creation of data. These data management plans, usually expected to be around two pages or less, typically include such areas as a description of the data and metadata to be used in the project, how these data will be stored, managed, shared, and preserved, and any associated legal or ethical issues.

Here are links to the requirements of key agencies:

NIH: http://grants.nih.gov/grants/policy/data_sharing/
DMP Tool: University Affiliation with the Help of OIT

Get Started

Login

If your institution is listed below the DMP tool will provide links to local data management resources and support available to you.

If you're using the DMP tool for the first time you'll be prompted to provide some additional information

Select Your Institution

- American University
- Arizona State University
- California Polytechnic State University-San Luis Obispo
- California State University, Chico
- California State University, Fresno
The library’s ability to provide open or controlled access to data
The existing campus technological infrastructure, including servers and software
The university’s policies and procedures for ensuring privacy and confidentiality of data
Future Directions

- A Data Services Librarian position has been requested as a conversion from an existing vacancy
- The university has decided to fund a GIS lab
- Increased data cataloging skills for Cataloging Services Unit
- Marketing to the faculty about storage for and dissemination of their data
- Promoting standardization of library-friendly licensing practices for data