

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 Euripedes



WinSCP window: Cross-National Time-Series Data Archive - elguindi@euripedes.wrhc.org

Local Drive: G:\ELGUINDI

Name	Ext	Size	Type	Changed
..			Parent directory	10/5/2011
Bibliography	.xls	2,478,080	Microsoft Office...	2/27/2012
CNTS Files 2011	.zip	4,496,426	Compressed (zi...	10/5/2011
CNTSDATA	.xls	13,011,968	Microsoft Office...	2/27/2012
Codebook	.xls	49,152	Microsoft Office...	2/22/2012
Contents	.doc	25,600	Microsoft Office...	2/22/2012
Independent States Since 1815	.xls	50,176	Microsoft Office...	2/22/2012
License Agreement	.doc	31,232	Microsoft Office...	2/22/2012
User's Manual	.doc	146,944	Microsoft Office...	2/22/2012

Remote Drive: /var/www

Name	Ext	Size	Changed
..			10/26/2011 3:1...
Licensed			3/22/2012 9:56...
Free			11/30/2011 12:...
AU_Created			11/30/2011 12:...
index.html		138	3/2/2012 9:47:...

0 B of 19,814 KiB in 0 of 8

0 B of 138 B in 0 of 4

F2 Rename F4 Edit F5 Copy F6 Move F7 Create Directory F8 Delete F9 Properties F10 Quit

SFTP-3 0:01:14

- A virtualized linux server with 100G storage and 2G of RAM

Access through the Online Catalog



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](#)

Request
through
Consortium
Loan Services.

Details

Title: Cross-national time-series data archive [electronic resource].

Publisher/Date: [Binghamton, NY] : Arthur S. Banks : Distributed by Databanks International, c2005.

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.](#)

Access on the Euripedes Server



Index of /Licensed/Cross-National_Time-Series_Data_Archive

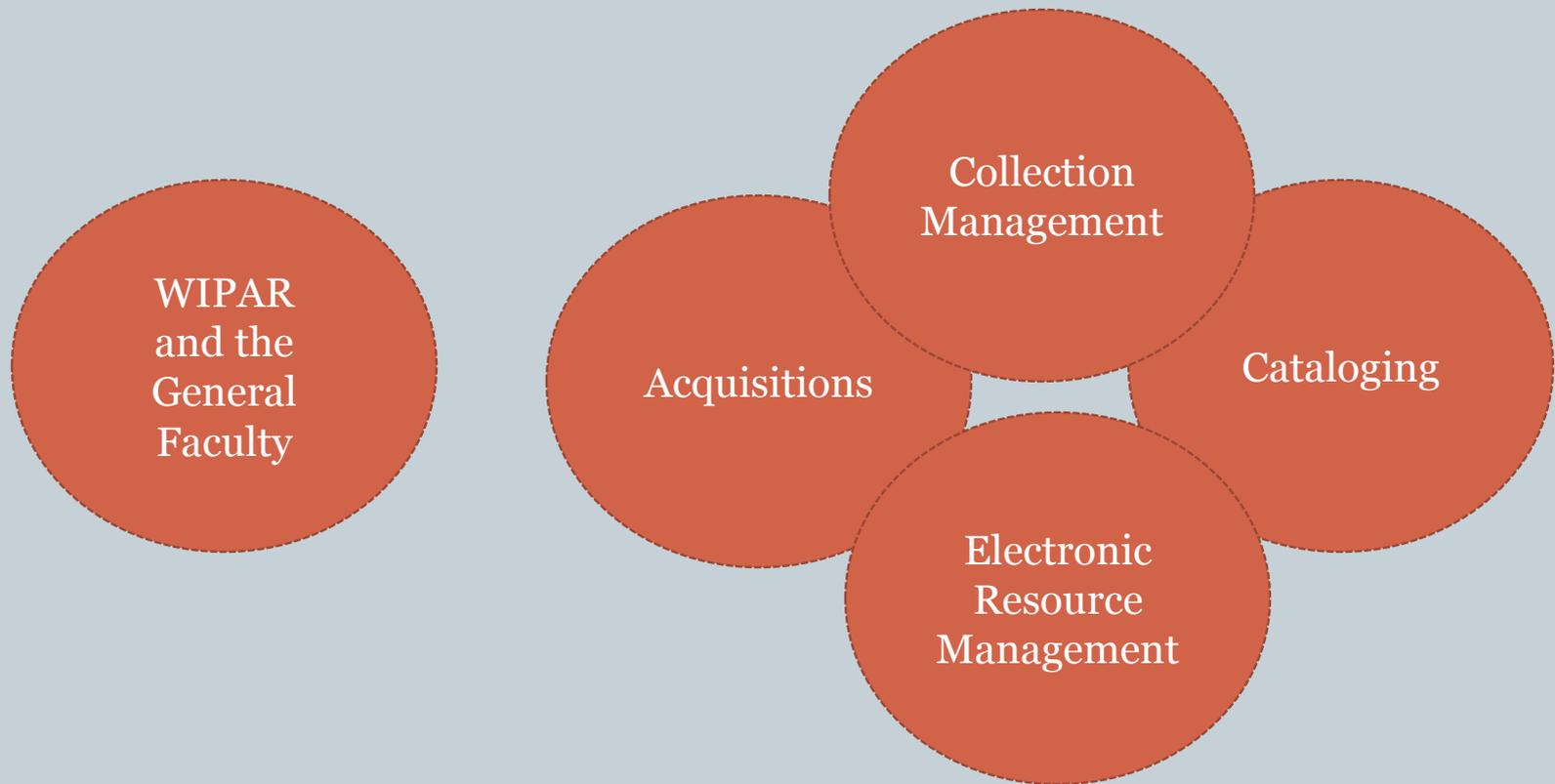
<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory		-	
 Bibliography.xls	05-Oct-2011 12:18	2.4M	
 CNTSDATA.xls	05-Oct-2011 12:18	12M	
 Codebook.xls	05-Oct-2011 12:18	48K	
 Independent States Since 1815.xls	05-Oct-2011 12:18	49K	
 User's Manual.doc	05-Oct-2011 12:18	144K	

Apache/2.2.14 (Ubuntu) Server at euripedes.wrlc.org Port 80

Partnership Two: Data Workflow



- Issues of licensing and workflow clarity



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



Add New Resource



CORAL Resources

* required fields

Product

Name: *

Description:

Provider:

URL:

Format *

Electronic CD-ROM

Acquisition Type *

Paid Free Gift Fund
 Open Access Supplemental Funds Trial

Initial Cost

Fund: Price: USD Type:

Resource Type

- Data set Database Digital Collection
- E-book - Single Title Firm Order E-book - Single Title Standing Order E-book Collection
- E-journal - Single Title E-journal Package

Notes

Include any additional information

save

submit

cancel

Data Workflow: Technical Services



Edit Workflow

Resource Entry Requirements

Acquisition Type:

Paid

Format:

Electronic

Type:

Data set

Workflow Steps

Name:	Approval/Notification Group:	Parent Step	
<input type="text"/>	Acq Funding Approval	<input type="text"/>	<input type="button" value="Add"/>
↓ Funding Approval	Acq Funding Approval	<input type="text"/>	×
↑ ↓ Confirm Pricing, Request License	IDS Director	Funding Approval	×
↑ ↓ Licensing	IDS Director	Confirm Pricing, Request Li	×
↑ ↓ Place Order with Vendor	IDS Director	Licensing	×
↑ ↓ Request Invoice from Vendor	IDS Director	Place Order with Vendor	×
↑ ↓ Create Firm Order in Voyager	IDS Director	Licensing	×
↑ ↓ Transfer Funds from BKS	Acq Librarian	Create Firm Order in Voyag	×
↑ ↓ Load Data on to Euripedes Server	IDS Director	Place Order with Vendor	×
↑ ↓ Receive Data on Purchase Order	IDS Director	Load Data on to Euripedes	×
↑ ↓ Create Invoice in Voyager	Acq Receiving	Transfer Funds from BKS	×
↑ ↓ Catalog Record and Holdings	Cataloging	Load Data on to Euripedes	×
↑ Access	ERM Unit	Catalog Record and Holdin	×

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

Center for
Research,
Teaching &
Learning

University
Library

Office of
Information
Technology

Washington
Research
Library
Consortium

Office of
Sponsored
Programs

Risk
Management

Institutional
Review Board

Grant-
Seeking
Faculty

Navigating the Politics



Approval to create
guide

Presentation to the Academic
Technology Steering
Committee

Contacted campus partners to
arrange a meeting

The Data Management Plan LibGuide



LIBRARY SUBJECT GUIDES



Library » [The Library's Subject Guides!](#) » [Data Management Plans](#)

[Admin Sign In](#)

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>

[Print Guide](#)

[RSS Updates](#)

[Email Alerts](#)

[SHARE](#)

[About DMPs and Data Management](#)

[DMP Templates](#)

[AU Campus Resources](#)

About DMPs and Data Management

[Print Page](#)

Search:

This Guide

Campus Contacts

Funder requirements: [Office of Sponsored Programs](#)

Data practices: [Social Science Research Lab](#)

Legal and ethical issues: [Institutional Review Board](#)

[Comments \(0\)](#)

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:

NSF: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>

NIH: http://grants.nih.gov/grants/policy/data_sharing/

NEH: <http://www.neh.gov/files/grants/datamanagementplans.pdf>

DMP Tool: University Affiliation with the Help of OIT



DMPTool
Guidance and Resources for your Data Management Plan

Home About DMP Tool DMP News My Plans Funder Requirements Help ▾

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 —

- Select Your Institution —
- American University
- Arizona State University
- California Polytechnic State University-San Luis Obispo
- California State University, Chico
- California State University, Fresno

AU Template: Building Standardized Language



- 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