The Team Approach: Developing a Data Services Program at the University of Washington

University of Washington Libraries

Matthew Parsons, Malyka Ianni, Theodore Gerontakos

Agenda

* Background / Structure

* Our Approach

* Metadata

* Moving Forward

Background

Data Services Program Planning Committee, June 2009-January 2010

Outcomes from report:

- Data Services Coordinator (July 2010)
- Data Services Team (September 2010)
- Data Services CE Fund

Organizational Structure

Assoc. Dean Research & Instructional Services

Assoc. Dean for Resource Access/Description & Info. Technology Services

Head, Reference & Research Services

Data Services Coordinator

Data Services Team roster:

Data Services Coordinator, **Chair** (Stephanie)

Library Specialist, Monographic Srvcs. (Will)

U.S. Documents Librarian (Cass)

Library Specialist, Monographic Srvcs. (Heather)

Regional Tech. Coord., NN/LM PNR (Mahria)

Business Computer-Based Librarian (Corey)

Geospatial Data and Maps Librarian (Matt)

Systems Librarian (Anjanette)

Metadata Librarian (Theo)

Data Services Specialist (Malyka)

Our Approach: A Three Step Process

- Outreach
- Continuing education for Librarians
- Developing resources and services

Outreach

Outreach efforts were priority for the first year

- Institute for Health Metrics & Evaluation
- Center for Advanced Research Technology in the Arts & Humanities
- Center for the Study of Demography & Ecology
- Center for Social Science Computational Research
- Office of Research

Continuing Education for Librarian Staff

- Webinars
- Speakers
- Conferences
- Workshops
- Courses

Developing Resources and Services

- Data Services LibGuide
- Data Citation - EZID
- Email list
- Presentations to groups & classes on RDM
- Data management survey
- Data management reference & consultation
- Dataset collection

Data Management Guide

Tags: data, information management, research, research_help, research_strategies, tools, tutorials

Guide of resources related to the many aspects of research data management.

Last Updated: Jun 21, 2012 | URL: http://guides.lib.washington.edu.offcampus.lib.washington.edu/dmg | Print Guide

What is Data Management?

What is data management?

Data This guide is focused on digital research data. We use a modified definition from the DISC-UK DataShare Report (p.16):

> That which is collected, observed, or created, for purposes of analyzing to produce original research results. Research data may be created in tabular, statistical, numeric, geospatial, image, multimedia or other formats.

Data management encompasses the processes surrounding collecting, organizing, describing, sharing, and preserving data.

Why is data management important?

The most effective and efficient data management practices begin at the research planning stage. With early planning for data management you can:

- **Save time** by having a plan in place for your data from the beginning of your project.
- Comply with **legal and funder requirements**.
- **Increase the visibility and impact** of your research by making your data searchable and citable.
- **Support open access** and **foster new research** by preserving your data and making it accessible to other researchers.

How do you manage data?

Contact Us

This guide is maintained by the **UW Libraries Data Services Team**.

If you have any questions about the content or would like to arrange for instruction or a consultation in data management, please feel free to **contact us**.

Data Management Resources

There are resources on campus specifically for UW researchers to help you with various stages of the data management cycle. We've also included resources outside the UW for general guidance on data management which we've found to be the most useful.

- Center for Social Science Computing & Research (CSSCR)
  provide courses on statistical software and access to social science data
- Center for Studies in Demography & Ecology (CSDE)
  provide data management services to faculty affiliates and students conducting research involving surveys or census
- eSciences Institute
  provide consulting, tools and storage solutions

Data Management Guide

Tags: data, information management, research, research_help, research_strategies, tools, tutorials

Guide of resources related to the many aspects of research data management.

Last Updated: Jun 21, 2012 | URL: http://guides.lib.washington.edu/dmg | Print Guide

Data Ma...

What is it?

A data management plan is **a document outlining how a researcher plans to manage data during and after a research project** including how it will be organized, maintained and shared.

Why do you need one?

More and more funding agencies are now requiring researchers to submit a formal data management plan (DMP) when applying for grants. The National Institutes of Health began requiring a DMP in 2003 and the National Science Foundation formalized the requirement in 2011. See the box below on "**Other Funding Agencies**" for more information on data management requirements from other agencies.

How do you do it?

When creating a plan for NSF, NIH, NEH, IMLS or the Gordon & Betty Moore Foundation, **we recommend using the California Digital Library's DMPTool**. This tool will walk you step-by-step through the requirements for each of the above funders and upon completion will provide you an exportable data management plan. **Click here** for a **guide on how to use the DMPTool**.

If you are requesting funding from an agency not covered by the DMPTool, check the Funding Agency Requirements box (below) for assistance.

Campus Services

If you have questions about data management planning or would like to **request a data management plan consultation** with a member of the Data Services Team, please submit a request **here**.

- eScience Institute

Comments (0)

Tools & Resources

- DMPTool
  This is an online tool to assist in writing data management plans for NSF. "The DMPTool supports data management plans and funder requirements across the disciplines, including the humanities and physical, medical and social sciences."
- Guidelines for Effective Data Management Plans
  Guidance from the Inter-University Consortium for Political and Social Research on creating a data management plan.
- ICPSR Data Management Plan Resources & Examples
  List of resources from the Inter-University Consortium for Political and Social Research. Includes templates, tools, funder

Generic NSF Data Management Plan Elements

(adapted from California Digital Library DMPTool)

1. Types of data produced

Provide a brief description of the data being collected.

Consider these questions:

- What data will be generated in the research?
- What data types will you be creating or capturing?
- How will you capture or create the data?
- If you will be using existing data, state that fact and include where you got it.
- What is the relationship between the data you are collecting and the existing data?

Also note here if the data will be of a sensitive nature (confidentiality, privacy or security issues for example).

2. Data and metadata standards

Explain how you will describe your data in a way so that others can make use of it. Describe the file structure, the variables, etc.

Consider the following when contemplating your data:

- Which file formats will you use for your data, and why?
- What form will the metadata describing/documenting your data take?
- How will you create or capture these details?
- Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)
- What contextual details (metadata) are needed to make the data you capture or collect meaningful?

3. Policies for access and sharing

Describe how and when your data will be made available once your project is completed.

Consider these questions:

- How will you make the data available? (Include resources needed to

NSF Funding Requirements & Information

The National Science Foundation began requiring data management plans as of January 18th, 2011. Plans are to take the form of a **two-page supplementary document**. The data sharing policy for the foundation is linked below with specific directorate guidelines under that. We've provided basic guidance for the **generic NSF data management plan in the box to the right**. University of Virginia has developed **templates** for NSF proposals including some for most of the directorates and they are linked below. UC San Diego has made some **examples** of DMPs available and they are linked next to the appropriate directorate below. Please remember these are project specific and are only here for guidance.

- Grants.gov Application Guide (PDF)
- Data Sharing Policy
- Data Management FAQ
- Generic NSF DMP template

NSF Directorate Guidance

- Biological Sciences Directorate (PDF) - Template
- Computer & Information Sciences & Engineering Directorate - Template, Example
- Education & Human Resources Directorate (EHR) (PDF) - Template
- Engineering Directorate (PDF) - Template, Example 1, Example 2
- Geological Sciences Directorate (Directorate-wide guidance)
  - Division of Atmospheric & Geospace Sciences - Template
  - Division of Earth Sciences - Template
  - Integrated Ocean Drilling Program (PDF)
  - Division of Ocean Sciences - Example
- Mathematical and Physical Sciences Directorate Office of Polar Projects (DOC)
  - Division of Astronomical Sciences (PDF) - Template
  - Division of Chemistry (PDF) - Template
  - Division of Materials Research (PDF) - Template
  - Division of Mathematical Sciences (PDF) - Template
  - Division of Physics (PDF) - Template
- Social, Behavioral & Economic Sciences Directorate (PDF) - Template

Comments (0)

"Metadata"

Drupal
RDFa
DCMI
DC-2012
vocabularies
Web
Metadata
metadata
published
LOD
data
machine-readable
conference
special
Data
includes
sessions
Semantic
Linked
September
program
W3C
software
used
available
Lite
now
use
VIVO
syntax
markup
publish
dcmi12
using
classes
document
AgInfra
Monday
titled
planning
Malaysia
accordance
organizing
Wednesday
full-day
documentation
Agro-Know
relationship
migration
page
show
Baker
Registration
encoding
stream
structured
simplest
information
maintenance
detailed
datatypes
Gregg
Twitter
Labs
Phipps
number
come
3-4
Tom
http://dcevents.dublincore.org/index.php/IntConf/index/pages/view/specialSessions-2012.2012-06-14
variant
committee
pages
Open
Nagamori
workshop
today
joined
registries
LinkIn
systems
interest
open-source
specifically
Kellogg
ZBW
Lanyrd
MIMOS
agINFRA
session
Agrovoc
keynotes
hashtag
source
support
extended
tools
generate
modified
repository
posters
best
FAO
June
Tutorials
Barnes
event
applications
help
organizations
might
Friday
e.g
Brickley
community
platform
Vocabularies
schemes
discuss
interact
rendered
Manu
deployment
preliminary
Recommendation
human-readable
Technologies
practices.2012-06-14
browsers
Berhad
Vocbench
on-screen
plenary
bringing
together
new
provides
open.2012-06-14
Sarawak
represented
embedding
design
Mitsuharu
Project
needed
Terms
GitHub
HTML
text
Tuesday

Metadata Workers

Developing data services

Libraries: Bothell Tacoma Health Sciences All

off-campus access (log in)

150 YEARS W

UNIVERSITY LIBRARIES
UNIVERSITY of WASHINGTON
Subject and Class Guides

ask us!
email | chat | phone | text

UW WorldCat UW Catalog Site
Search UW Libraries and beyond
Advanced search | Help

Admin Sign In

Data Services

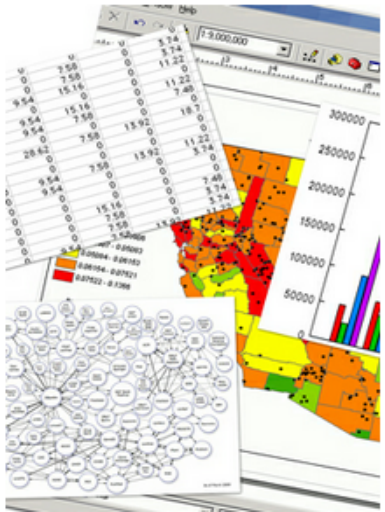
Tags: data, research, research_help

Last Updated: Apr 24, 2012 | URL: http://guides.lib.washington.edu/data | Print Guide | RSS Updates | SHARE

Home | Find Data | Visualize Data | Data Management Plans

Home Comments(0) Print Page

What's in the Guide



Welcome to the guide for **Data Services**. This guide provides links to useful physical and virtual resources for finding, analyzing, visualizing and managing data. Use the tabs to navigate through the pages of this guide.

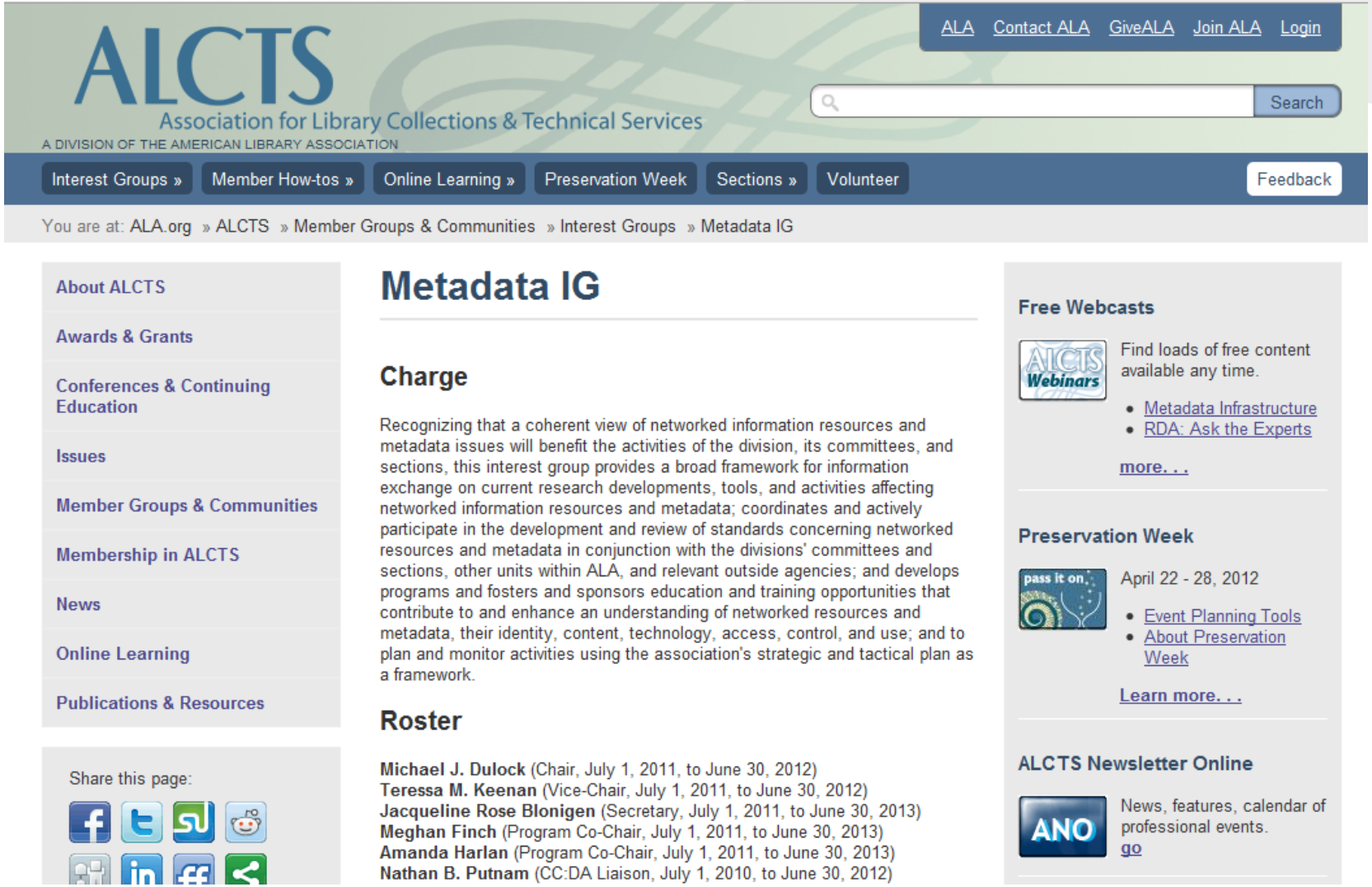
- **Find Data** - search for data in some of the major data portals or find a resource to obtain more specialized data
- **Visualize Data** - books, software and tutorials on how to turn data into information by visually presenting patterns and trends
- **Data Management Plans** - explore requirements for data management plans by funding agencies and templates for those plans, find other campus resources to help you manage your research data

Data Services Coordinator

Stephanie Wright

Contact Info
Research Commons
Allen South G086C
206.685.1540
Send Email

Metadata Interest Group



ALCTS
Association for Library Collections & Technical Services
A DIVISION OF THE AMERICAN LIBRARY ASSOCIATION

ALA Contact ALA GiveALA Join ALA Login

Search

Interest Groups » Member How-tos » Online Learning » Preservation Week Sections » Volunteer Feedback

You are at: ALA.org » ALCTS » Member Groups & Communities » Interest Groups » Metadata IG

About ALCTS
Awards & Grants
Conferences & Continuing Education
Issues
Member Groups & Communities
Membership in ALCTS
News
Online Learning
Publications & Resources

Share this page:

Metadata IG

Charge

Recognizing that a coherent view of networked information resources and metadata issues will benefit the activities of the division, its committees, and sections, this interest group provides a broad framework for information exchange on current research developments, tools, and activities affecting networked information resources and metadata; coordinates and actively participate in the development and review of standards concerning networked resources and metadata in conjunction with the divisions' committees and sections, other units within ALA, and relevant outside agencies; and develops programs and fosters and sponsors education and training opportunities that contribute to and enhance an understanding of networked resources and metadata, their identity, content, technology, access, control, and use; and to plan and monitor activities using the association's strategic and tactical plan as a framework.

Roster

Michael J. Dulock (Chair, July 1, 2011, to June 30, 2012)
Teressa M. Keenan (Vice-Chair, July 1, 2011, to June 30, 2012)
Jacqueline Rose Blonigen (Secretary, July 1, 2011, to June 30, 2013)
Meghan Finch (Program Co-Chair, July 1, 2011, to June 30, 2013)
Amanda Harlan (Program Co-Chair, July 1, 2011, to June 30, 2013)
Nathan B. Putnam (CC:DA Liaison, July 1, 2010, to June 30, 2012)

Free Webcasts

ALCTS Webinars

Find loads of free content available any time.

- Metadata Infrastructure
- RDA: Ask the Experts

more. . .

Preservation Week

pass it on.

April 22 - 28, 2012

- Event Planning Tools
- About Preservation Week

Learn more. . .

ALCTS Newsletter Online

ANO

News, features, calendar of professional events.
go

Data/Metadata

Data/metadata

Data Near Here V0.5b (Research Edition)

Please enter the following parameters:

Categories	ALL	**Quality**	ANY
SW Corner: [dec.deg]	46.207387,-123.8	**NE Corner:** [dec.deg]	46.229708,-123.8
Depth: from [m]		**Depth to:** [m]	
Start date:	2009-05-05	**End date:**	2009-05-20
with variable:	salinity	? More Delete	
	Range: -	Units: psu	

Min. Obs. Count: 1

Get 'em! Click here for Usage Notes

There were 50 results returned; all are listed, and 25 initially shown on map. Salinity was found in 50 entries.

Display	Type	Collection	Quality	Start Time	End Time	From Depth	To Depth	salinity	Observations	Data Location	Score	DNH
1	AUV	Mission 13, 2009-05-14	raw_data	2009-05-14 10:05 PDT	2009-05-14 11:05 PDT	-15.39	-0.28	0.07:9.66 m	1,995	Download	98	DNH
2	AUV	Mission 12, 2009-05-14	raw_data	2009-05-14 08:05 PDT	2009-05-14 09:05 PDT	-14.72	-0.28	0.07:10.50 m	1,632	Download	98	DNH
3	AUV	Mission 14, 2009-05-15, Segment 4	raw_data	2009-05-15 12:05 PDT	2009-05-15 12:05 PDT	-13.19	-2.43	0.20:4.98 m	1,081	Download	98	DNH

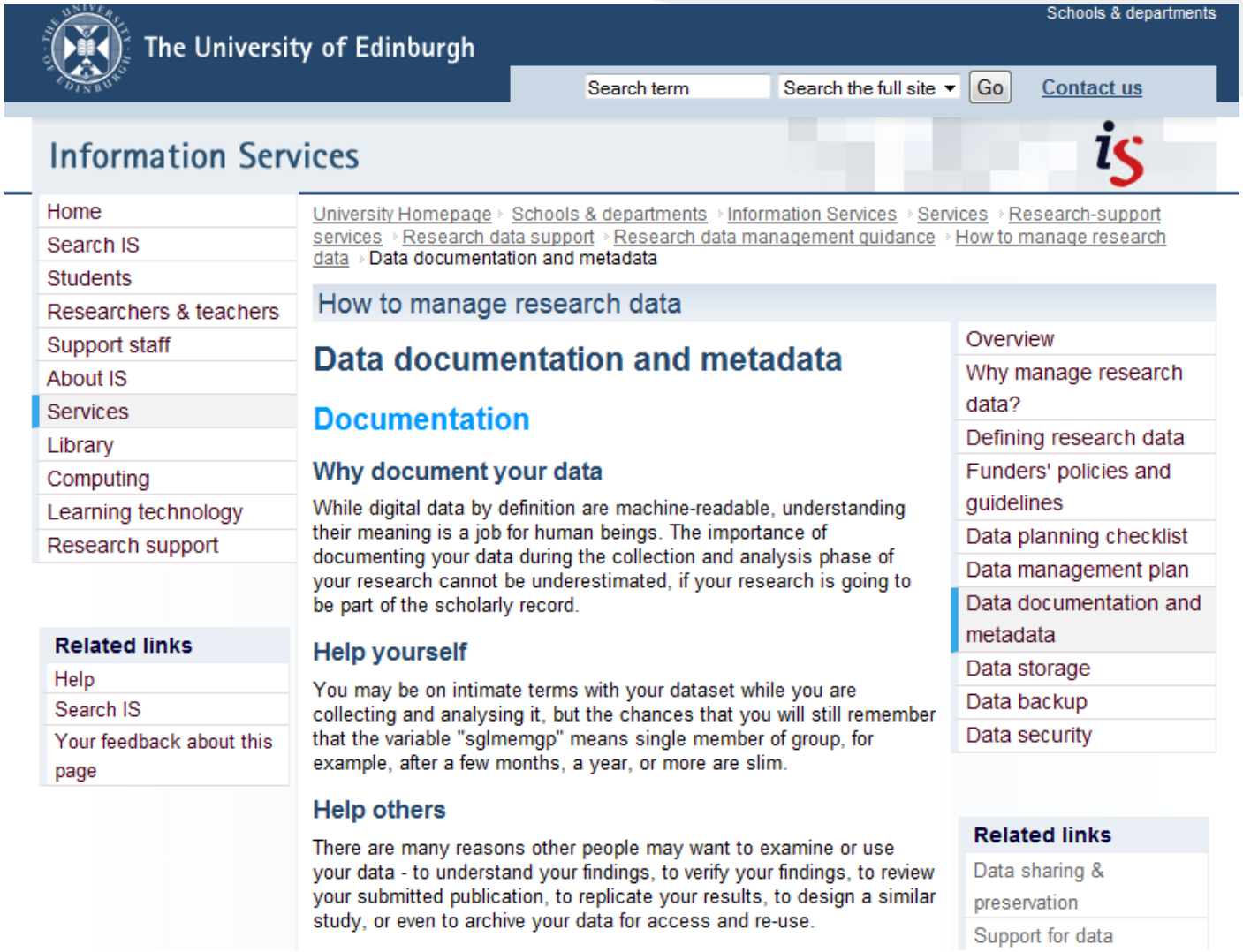
Metadata Librarian

Metadata management for the BBC's 2010 World Cup site using OWLIM

Marin Dimitrov (Ontotext)

European Semantic Technology Conference 2010

Metadata will manage datasets



The University of Edinburgh

Schools & departments

Search term | Search the full site | Go | Contact us

Information Services

- Home
- Search IS
- Students
- Researchers & teachers
- Support staff
- About IS
- Services
- Library
- Computing
- Learning technology
- Research support

Related links

- Help
- Search IS
- Your feedback about this page

University Homepage › Schools & departments › Information Services › Services › Research-support services › Research data support › Research data management guidance › How to manage research data › Data documentation and metadata

How to manage research data

Data documentation and metadata

Documentation

Why document your data

While digital data by definition are machine-readable, understanding their meaning is a job for human beings. The importance of documenting your data during the collection and analysis phase of your research cannot be underestimated, if your research is going to be part of the scholarly record.

Help yourself

You may be on intimate terms with your dataset while you are collecting and analysing it, but the chances that you will still remember that the variable "sglmemgp" means single member of group, for example, after a few months, a year, or more are slim.

Help others

There are many reasons other people may want to examine or use your data - to understand your findings, to verify your findings, to review your submitted publication, to replicate your results, to design a similar study, or even to archive your data for access and re-use.

- Overview
- Why manage research data?
- Defining research data
- Funders' policies and guidelines
- Data planning checklist
- Data management plan
- Data documentation and metadata
- Data storage
- Data backup
- Data security

Related links

- Data sharing & preservation
- Support for data

Metadata is essential to data management

"**If the data is to be analyzed by generic tools, the tools need to "understand" the data.** You cannot just present a bundle-of-bytes to a tool and expect the tool to intuit where the data values are and what they mean. The tool will want to know the metadata."

--Jim Gray, David T. Liu, Maria Nieto-Santisteban, Alex Szalay, David J. DeWitt, Gerd Heber, "Scientific Data Management in the Coming Decade," in ACM SIGMOD Record, Vol. 34, No. 4, Dec. 2005, p. 34-41

Metadata improvements driving new tools and services at a NASA data center

Moroni, D. F.; Hausman, J.; Foti, G.; Armstrong, E. M.

American Geophysical Union, Fall Meeting 2011, abstract #IN41B-1408

The NASA Physical Oceanography DAAC (PO.DAAC) is responsible for distributing and maintaining satellite derived oceanographic data from a number of NASA and non-NASA missions for the physical disciplines of ocean winds, sea surface temperature, ocean topography and gravity. Currently its holdings consist of over 600 datasets with a data archive in excess of 200 Terrabytes. The PO.DAAC has recently embarked on a metadata quality and completeness project to migrate, update and improve metadata records for over 300 public datasets. An interactive database management tool has been developed to allow data scientists to enter, update and maintain metadata records. This tool communicates directly with PO.DAAC's Data Management and Archiving System (DMAS), which serves as the new archival and distribution backbone as well as a permanent repository of dataset and granule-level metadata. Although we will briefly discuss the tool, more important ramifications are the ability to now expose, propagate and leverage the metadata in a number of ways. First, the metadata are exposed directly through a faceted and free text search interface directly from drupal-based PO.DAAC web pages allowing for quick browsing and data discovery especially by "drilling" through the various facet levels that organize datasets by time/space resolution, processing level, sensor, measurement type etc. Furthermore, the metadata can now be exposed through web services to produce metadata records in a number of different formats such as FGDC and ISO 19115, or potentially propagated to visualization and subsetting tools, and other discovery interfaces. The fundamental concept is that the metadata forms the essential bridge between the user, and the tool or discovery mechanism for a broad range of ocean earth science data records.

Keywords: [1936] INFORMATICS / Interoperability, [1946] INFORMATICS / Metadata, [1948] INFORMATICS / Metadata: Provenance, [1950] INFORMATICS / Metadata: Quality

Featured Database Articles

SQL ETC

March 24, 2010

Database Management: Metadata is more important than you think!

By Denise Rogers

> Whether it's data warehousing, MDM or business intelligence, metadata is added to the project plan, is downgraded and eventually dropped from the project plan. The impacts of not including metadata and metadata management as part of the project have far-reaching and costly repercussions throughout the organization. Read on to learn more...

We are in the midst of a data architecture revolution! Whether it's data warehousing, MDM, business intelligence or data modernization, there is some sort of related project, program or initiative.

However, as we all sit around the planning table, in work session after work session, the theme is always the same. That is, metadata is added to the project plan as an integral component but is downgraded and eventually dropped off the implementation part of the project plan. The reasons given in many instances are always related to time constraints, re-definition of scope or the metadata component is too large and should be treated as a separate project, which almost always never happens.

UNIVERSITY LIBRARIES
UNIVERSITY of WASHINGTON

Data Services Program Planning Committee

Final Report

January 27, 2010

Submitted by:
Stephanie Lamson, Co-Chair
Matthew Parsons, Co-Chair
Eleanor Chase
Ian Dotson
Ann Ferguson
Theo Gerontakos
Corey Murata
Stephanie Wright

From an earlier slide:

Data Services Team Roster:

Data Services Coordinator, Chair (Stephanie)
Business Computer-Based Librarian (Corey)
Library Specialist, Monographic Srvcs. (Will)
Geospatial Data and Maps Librarian (Matt)
U.S. Documents Librarian (Cass)
Systems Librarian (Anjanette)
Library Specialist, Monographic Srvcs. (Heather)
Metadata Librarian (Theo)
Regional Tech. Coord., NN/LM PNR (Mahria)
Data Services Specialist (Malyka)

One Metadata Librarian

Lots of metadata workers

CATALOGING AND METADATA SERVICES
- About Us
  - Organization Chart
  - Language Expertise
  - Annual Reports For Monographic Services
- Cataloging-Related Information
- Purchase Request Form
- Contact Us

General Information

Cataloging and Metadata Services

Personnel

Cataloging and Metadata Services has 33 members, including 14 librarians and 19 classified staff.

Organization

Cataloging and Metadata Services was formed in July, 2012, from sections of the former Monographic Services and Serials Services divisions.

The division is comprised of five sections:

- **Database Management**
- **International Studies Cataloging**
- **Monographic Cataloging**
- **Serials Cataloging**
- **Special Materials Cataloging**

The Monographic, Special Materials and International Studies Cataloging Sections perform complex copy and original cataloging. The Monographic Cataloging Section covers most of the humanities, social sciences and sciences in Western languages. The International Studies Cataloging Section handles materials in Near Eastern, Slavic, South Asian and Southeast Asian languages. The Special Materials Cataloging Section handles non-book materials, as well as many books in a variety of languages and subjects. Staff from the three sections also provide subject cataloging for digital projects. The Monographic Cataloging, International Studies Cataloging, and Special Materials Sections are self-managing.

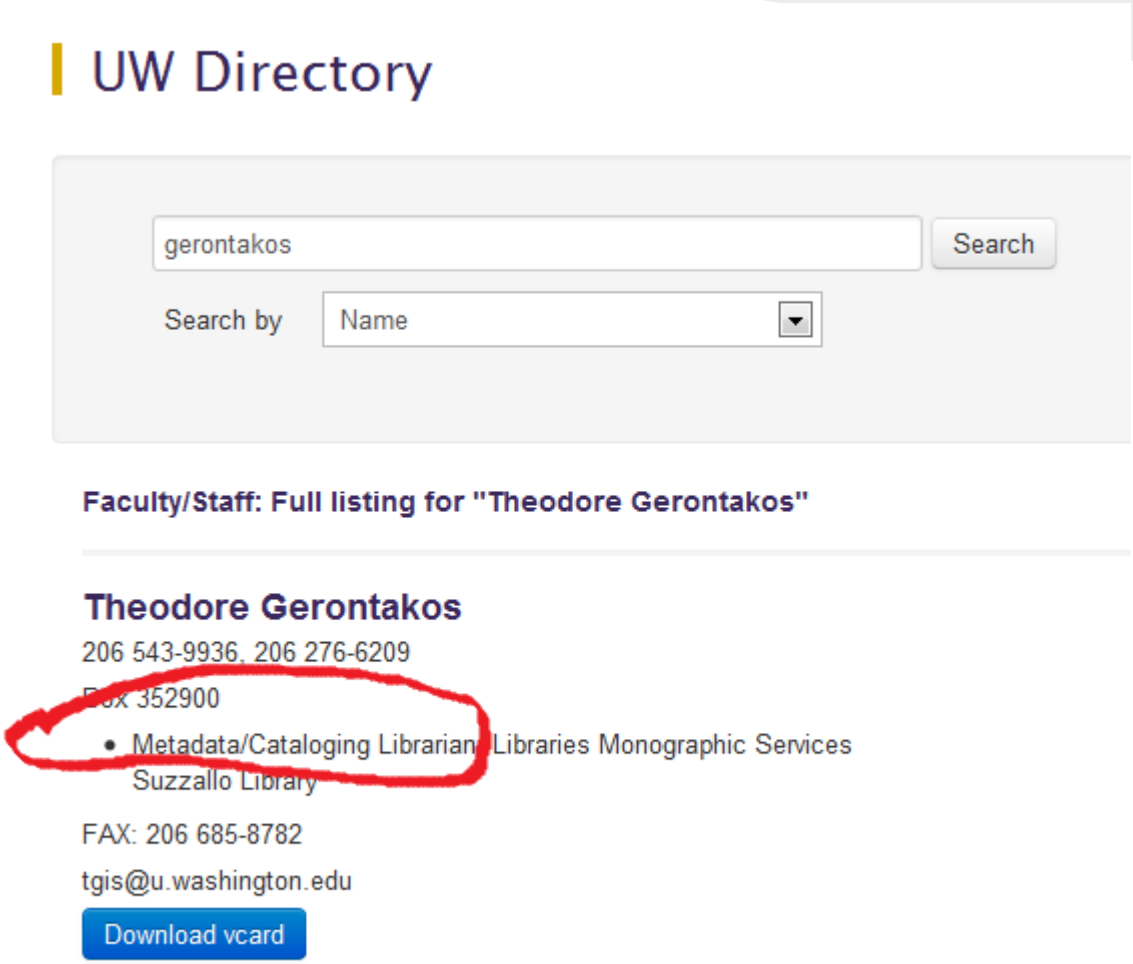
The Serials Cataloging Section performs complex copy and original cataloging of serials in print and electronic formats in all languages.

The Database Management Section is responsible for on-going maintenance activities such as corrections, including those resulting from authority processing and daily OCLC loading. Other activities include retrospective cataloging of the pre-76 U.S. Documents collection in the Government Publications Division.

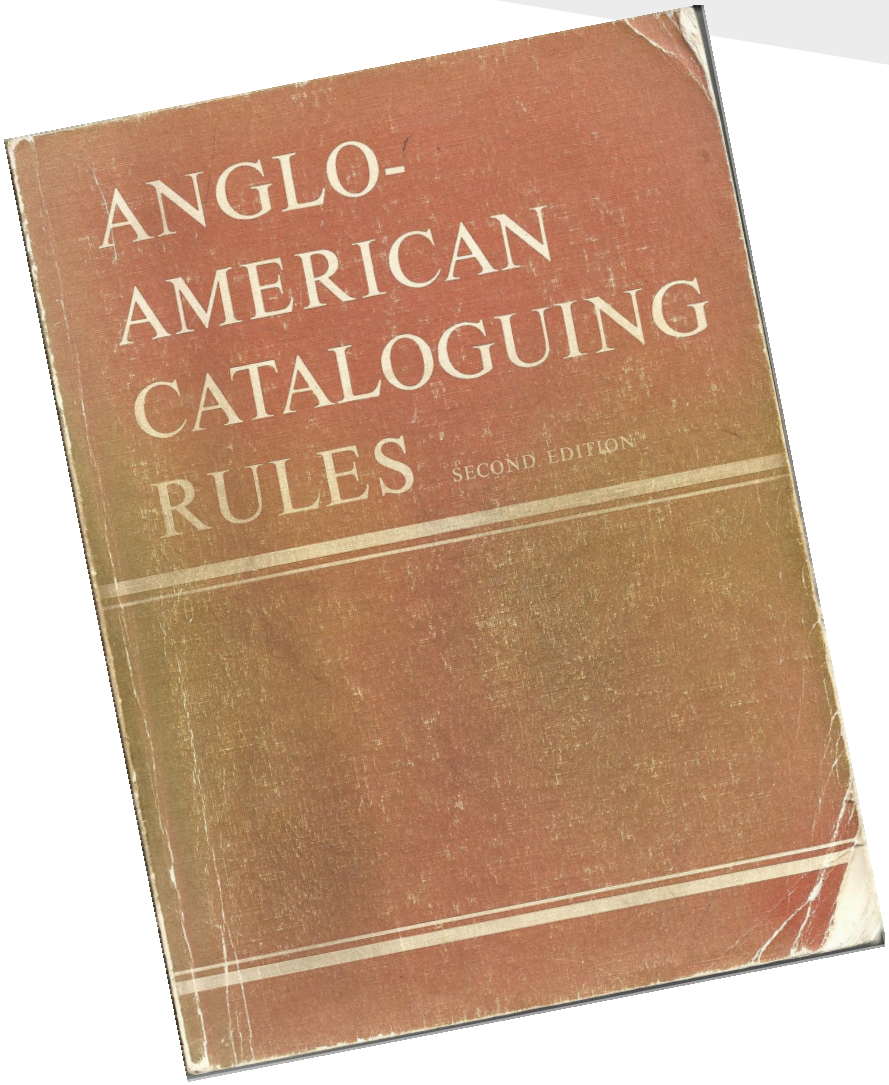
Standards

The division catalogs according to *Anglo-American Cataloging Rules*, and *Library of Congress Rule Interpretations*. Implementation of *Resource Description and Access* (RDA)

Metadata/Cataloging Librarian



Traditional Cataloging



Common standards

01338nam a22003730
450000100120000000200090001200500170002100800410003804000130007906600060009206600070009809000200010509900190012509900190014413000440016324500890020726000460029630000460034250000510038870000300043970000380046974000360050788000430054388000790058688000490066588000460071488000320076088000430079288000310083594100150086691000140088194100150089598700540091000cm2272378401542953200110181358590901121s1667 xx a 000 0 chi d aWAUcWAU c1 c$1 aDS707b.S4 1667 aDS707 .S4 1667 aDS707 .S4 16670 6880-01aShan hai jing (Chinese classic)106880-02aShan hai jingb[18 juan, tu 5 juan,cGuo Pu zhuan] Wu Zhiyi [Wu Renchen] zhu 6880-03a[n.p.bn.p.]cKangxi 6 (1667) xu] a6 v. (double leaves) in 1.billus.c26 cm 6880-04aCaption title: Shan hai jing guang zhu1 6880-10aGuo, Pu,d276-3241 6880-11aWu, Renchen,d1628?-1689?0 6880-05aShan hai jing guang zhu006130-01/$1a山海經 (Chinese classic)106245-02/$1a山海經b[18卷, 圖5卷,c郭璞傳] 吳志伊[吳任臣]注0 6260-03/$1a[n.p.bn.p.]c康熙6 (1667) 序] 6500-04/$1aCaption title: 山海經廣注106700-10/$1a郭璞,d276-324106700-11/$1a吳任臣,d1628?-1689?eed016740-05/$1a山海經廣注 a.b25105383 amt-75 rev a.b25105383 aPINYINbOCoLCc20011014dre09.30.2001f130 REV /01086nam a22003370 450000100120000000200090001200500170002100800410003804000130007906600060009206600070009809000160010509900150012109900150013624500500015126000720020130000410027350000380031465100310035270000380038374000300042188000450045188000660049688000380056288000370060088000280063794100150066591000100068094100150069098700430070 5ocm22963346015541732011018140158.0910115q15731620xx 000 0 chi d aWAUcWAU c1 c$1 aDS736b.I15 aDS736 .I15 aDS736 .I15006880-01aYi zhou shub[10 juan]cKong Chao zhu 6880-02a[n.p.,bMing Jiang shichang kan banc [between 1573 to 1620] a2 v. (double leaves) in case.c29 cm 6880-03aSong Jiading 15 (1222) ba 0aChinaxHistoryxPhilosophy1 6880-04aKong, Chao,d3rd/4th cent0 6880-05aJi zhong Zhou shu006245-01/$1a 逸周書b[10卷]c孔晁注0 6260-02/$1a[n.p.,b明姜士昌刊本c[between 1573 to 1620] 6500-03/$1a宋嘉定15 (1222)跋]106700-04/$1a孔晁,d3rd/4th cent016740-05/$1a汲冢周書 a.b25217008 acl-71 a.b25217008 aPINYINbOCoLCc20011014dce09.30.200101261nam a22003610…

Metadata work in progress

Metadata Implementation Group

Metadata in the UW Libraries

- **Metadata Guidelines**: for people who want to mount digital collections using CONTENTdm software.
- **Data Dictionaries** (a.k.a. schemas, metadata application profiles or MAPS): metadata elements, mappings to external schema, CONTENTdm database configuration properties, and data formatting instructions for various digital projects.
- **UW Digital Collections**: online multimedia collections built under auspices of the University of Washington Libraries' Digital Initiatives Program.
- **UW ResearchWorks**: institutional repository at UW powered by DSpace.
- **Encoded Archival Description (EAD) at UW Libraries**: the use of EAD at UW Special Collections.
- **Finding Aids for Archival Collections**: EAD encoded finding aids, UW Special Collections.
- **Local XML Schemas**: Locally created XML schemas.
- **Local Namespaces**: Locally created namespaces for use with XSLT.

Members of MIG:

- Diana Brooking
  Cataloging Librarian (Slavic)
- Theodore Gerontakos, convener
  Metadata/Cataloging Librarian
- Anne Graham
  Senior Computer Specialist, Digital Initiatives
- Joe Kiegel
  Head, Monographic Services Division
- Kris Kinsey
  Special Collections
- Helice Koffler
  Manuscripts and Special Collections Materials Cataloger
- Stephanie Lamson
  Asst. Preservation Librarian
- Laura Lins
  Special Materials Cataloger

Charge of the Metadata Implementation Group

As a committee of the University of Washington Libraries, reporting to the Head of the Monographic Services Division, the Metadata Implementation Group (MIG) develops and promotes the use of metadata standards to ensure reliable resource discovery within and across digital library projects. The Group will identify appropriate metadata and coordinate consistent application of metadata across a variety of software environments and resource types.

MIG data/metadata

Data Dictionaries

a.k.a. Schemas and Metadata Application Profiles or MAPS

Click on a project or collection name below to access a text version of the collection-specific data dictionary, metadata application profile, or schema.

1. American Indians of the Pacific Northwest (Graphical) *[Created September, 1998]*
2. American Indians of the Pacific Northwest (Textual) *[Created January, 1999]*
3. Ancient Near East Image Archive *[Last update July, 1999]*
4. Architecture Collection Data *[Created February, 2005]*
5. Central Eurasian Information Resource--Image Database *[Created March, 2002]*
6. Central Eurasian Information Resource--Text Database *[Created December, 2000]*
7. Cloud Seeding Collection *[Created July, 2009]*
8. Decorated and Decorative Paper Collection *[Created March, 2006]*

Data Dictionaries

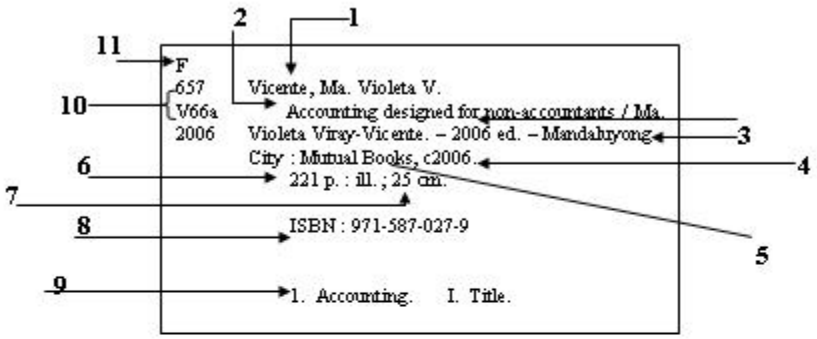
Dublin Core	Photograph Collections	Description
DC.Title: searchable, public/staff field; *required field*	Title	A brief descriptive phrase that includes these elements, generally in this order: who, what, where, when. The **Title** is a concise sentence which will be used to label the thumbnails; it will also be the first line of the descriptive information. The **Title** may be the caption assigned by the photographer or the collector, or it may be created by Special Collections staff from the given caption and from additional sources. In the absence of any caption, the **Title** will be provided by Special Collections staff. If the caption provided by the photographer is adequate, place it in the **Title** field. Then, the **Notes** field should include this phrase: *Title taken from photograph*. If the caption provided by the photographer is inadequate, a created title should be placed in the **Title** field; the **Notes** field should then read: *Caption on image: "Image title."* If there is a series caption, it should be added, in the **Notes** field, in sequence after the title on the image: *Caption on image: "Image title"; "Series title."* There should also be a separate notation in the **Notes** field in the form: *Part of Hegg series entitled: "Series title."*
DC.Creator: searchable, public/staff field	Photographer	The name of the photographer or firm associated with the creation of the image in hand. It should represent as closely as possible the creator's name or the company's name at the time of the creation of the image. Can be the same entry as in the **Studio Name** field. Includes both personal and corporate names. Input *Lastname, Firstname* for all personal names. Use LC Authority File for form of name, if available.
DC.Creator: searchable, public/staff field	Architect	The name of the architect or firm who created the architectural work depicted in the image. It should represent as closely as possible the creator's name or the company's name at the time of the creation of the image. Includes both personal and corporate names. Input *Lastname, Firstname* for all personal names. Use LC Authority File for form of name, if available.
DC.Date: Non-searchable, public/staff field; *required field*	Date	A non-searchable text field. The date the original photograph was taken. In cases where a painting, engraving, or architectural drawing was photographed, the date the artwork was created. If the date when the image in hand was produced is known, and is different from that of the original image, so note in the **Notes** field. The date should be a specific year. If the date is unknown, an attempt should be made to assign an approximate date, using the form "ca." (circa). Example ***Date:*** *ca. 1904* This field is used in combination with the **Dates** field to enable searching (see that entry for details). Specific dates (e.g., September 12, 1933; June 1912) are to be noted in the **Notes** field, and should also be incorporated in the **Title** field.
DC.Date: searchable, staff-only field; *required field*	Dates	Reflects the **Date** field. If the date is a single year, it is the same in both the **Date** and the **Dates** fields. If the date in the **Date** field is a "circa" date (e.g. ca. 1895), the **Dates** field contains the expanded version so that searching will find dates covered by the concept of a "circa" date. Five years on either side of the "ca." date is the preferred form. Example If the **Date** field reads "ca. 1910,": ***Dates:*** *1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915*

Schemas, transforms, XML technologies

```
<?xml version="1.0" encoding="UTF-8"?>
    <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:local="http://example.
    org/local">
      <xs:annotation>
        <xs:documentation>A schema to validate Dublin Core Qualified records as defined by the
          DSpace simple archive format. See the DSpace manual, ver. 1.6.1, section
          8.3.1.</xs:documentation>
      </xs:annotation>
      <xs:element name="dublin_core" type="dublincoreType">
        <xs:annotation>
          <xs:documentation> Top level container for the individual record </xs:documentation>
        </xs:annotation>
      </xs:element>
      <xs:complexType name="dublincoreType">
        <xs:sequence maxOccurs="unbounded">
          <xs:element ref="dcvalue"/>
        </xs:sequence>
      </xs:complexType>
      <xs:element name="dcvalue" type="dcvalueType">
        <xs:annotation>
          <xs:documentation>The element that contains data values, with attributes for DC element
            name, qualifier, and optional language.</xs:documentation>
        </xs:annotation>
```

Catalog card = data?

Author Catalog Card



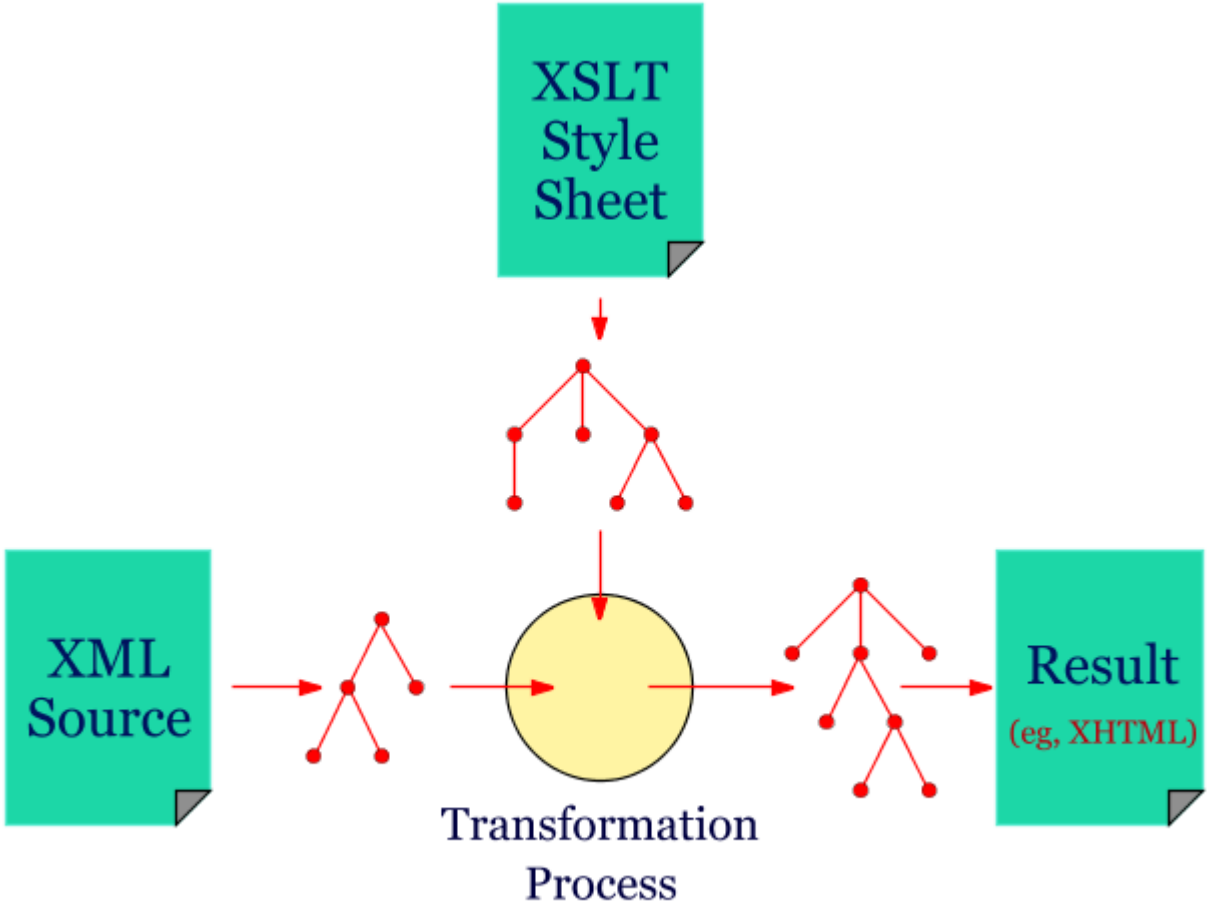
LEGEND

1 - Author
2 - Title
3 - Place of Publication
4 - Date of Publication
5 - Publisher
6 - Number of Pages
7 - Size in Height
8 - International Standard Book Number
9 - Subjects
10 - Call Number (Class Number & Author Number)
11 – Sublocation
F - Filipiniana
M – Main Library
R – Reserve
N - Nursing
G – Graduate
HS – High School
Ref – Reference
E – Engineering

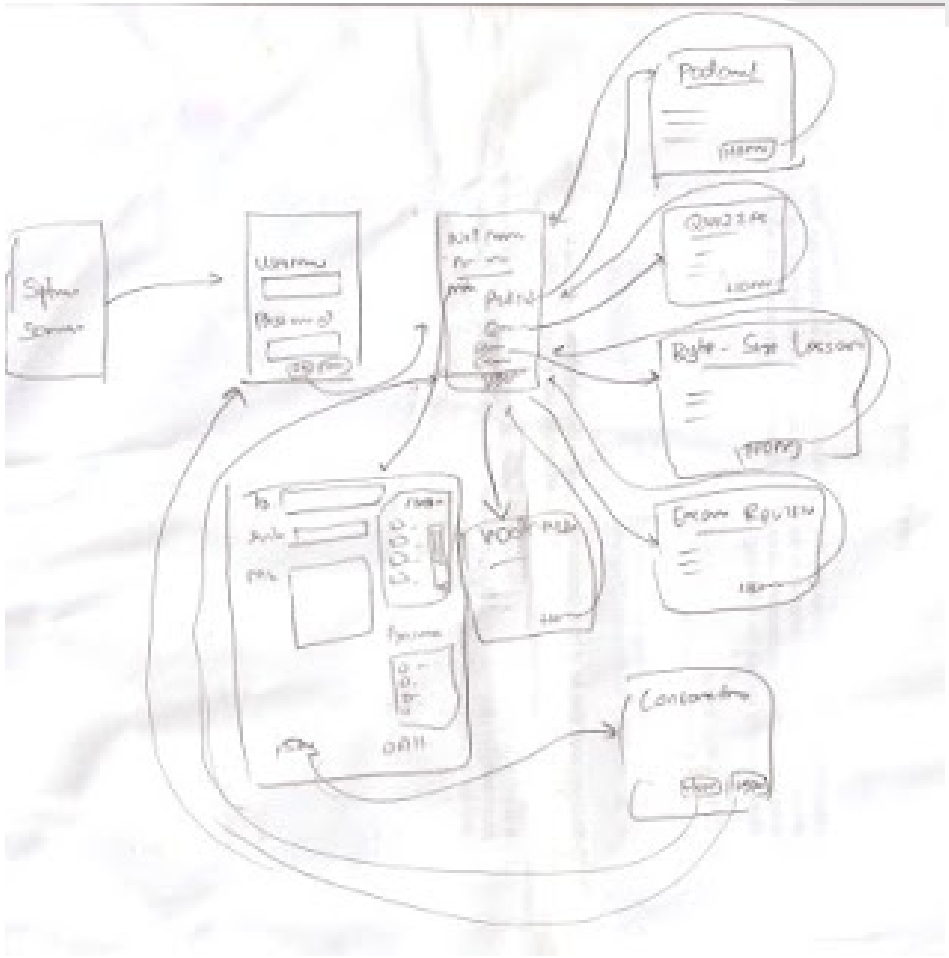
Books are also color coded:

Circulation	-	Green
Reserve	-	Apple green
Filipiniana	-	Pink
Reference	-	Fuschia
Architecture	-	Orange
Optometry	-	White
Nursing	-	White

XSLT processing diagram

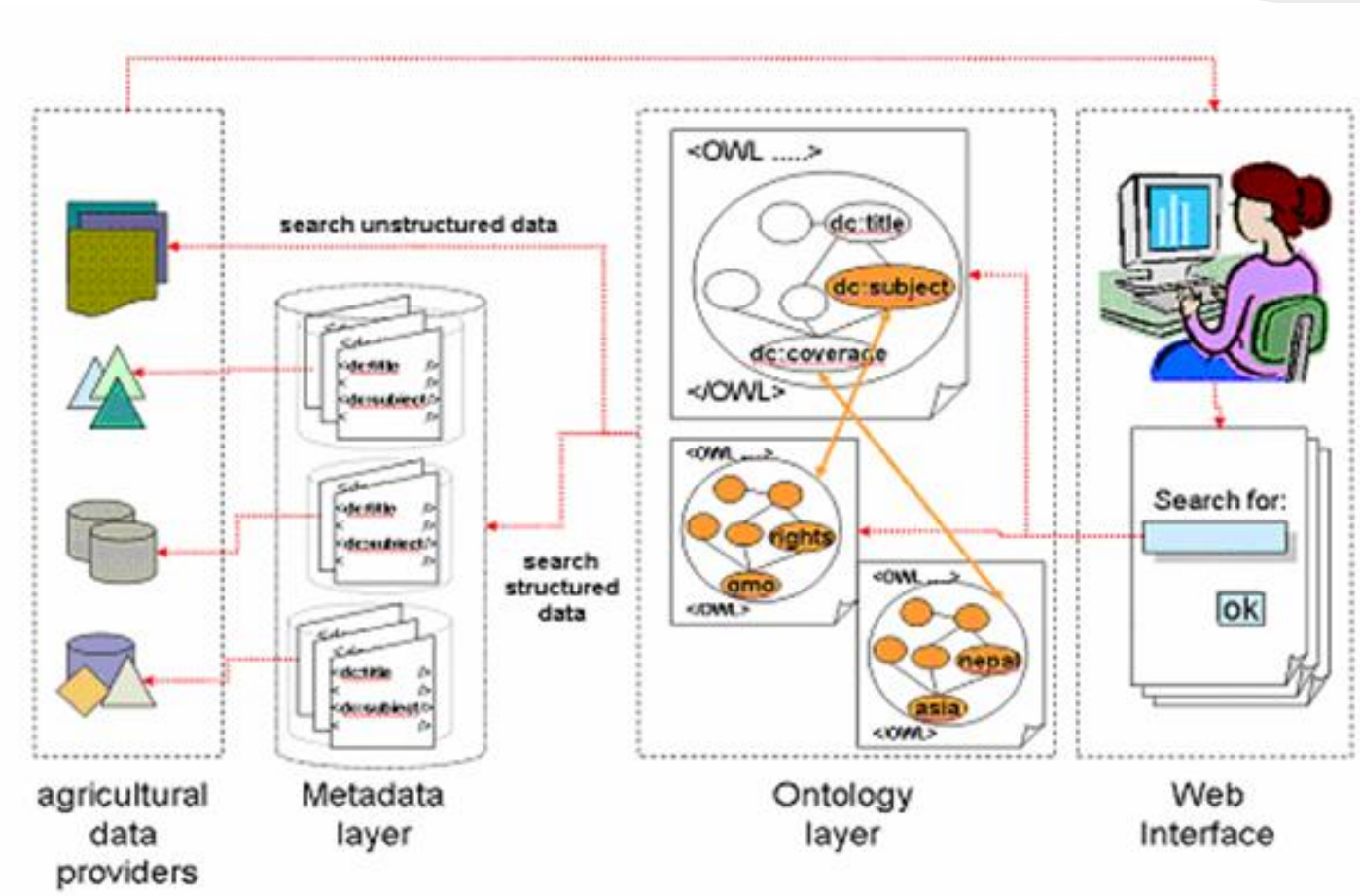


Advanced Data Model



Data Service Initiators

Data Repository



Lots to keep up with

Courtesy of http://howtosplitanatom.com/

Who has the time?

Dallas Museum of Art **Metadata Standards Crosswalk for Digital Asset Management (Images)**

DMA Metadata Standards Crosswalk for Digital Asset Management (Images)

CDWA	DMA DAMS	Data Entry Examples	TMS table	CCO Core Chapter References	VRA 4.0 CORE ELEMENTS	Dublin Core
CURRENT LOCATION Repository Name/Geographic Location (core)	**OBJECT OWNER** (enter and drop down menu)	Default: Dallas Museum of Art	n/a	Ch.5	LOCATION name	contributor
[local]	**ON VIEW**	on view; not on view; gallery location?	Objects.OnView	[local]	[local]	[local]
[local]	**TMS RECORD ID**	4171219 (system generated; cut/paste)	Objects.ObjectID	[local]	[local]	identifier
CURRENT LOCATION Repository Numbers (core)	**ACCESSION NUMBER** (import)	1984.57.A-B (assigned)	Objects.ObjectNumber	Ch. 5	LOCATION refid	identifier
[local]	**CURATORIAL DEPARTMENT**	African	Departments.Department	[local]	[local]	[local]
CLASSIFICATION Term (core)	**CLASSIFICATION**	Paintings, Prints, Sculpture	Classifications.Classification	Ch. 7	SUBJECT	subject [Controlled Vocabulary Recommended]
TITLES or NAMES Text (core)	**OBJECT TITLE**	Icebergs	OBJTitles.Title	Ch. 1	TITLE	title
CREATION Creator Description - Attribution	**ATTRIBUTION QUALIFIER**	Attributed to	Makers.Prefix [to include display name, prefix and display name would have to be concatenated]	Ch. 2	AGENT attribution	n/a - *NEED DISPLAY NAME FOR CREATOR*
CREATION Creator Description - Identity	**ARTIST**	Gerald Murphy	Makers.DisplayName	Ch. 2	AGENT name	creator
CREATION Place/Original Location	**GEOGRAPHY**	Africa	need formula (similar to eMuseum formula)	Ch. 5	LOCATION name	subject OR coverage.spatial [TGN Recommended]
CREATION Culture	**CULTURE**	Yoruba	OBJContext.Culture (note: do not duplicate in Geography Field)	Ch. 4	CULTURE (work record)	subject

COMPLEXITY


DDI
DATA
ALLIANCE
VERSION
DEVELOPMENT
STRUCTURE
INCLUDE
INSTITUTE
TECHNICAL
SCIENCE
GROUP
METADATA
MEETING
BEST
TYPE
ASSOCIATION
SERVICE
DIRECTOR
XMLSCHEMA
STATISTICAL
MODEL
PRACTICES
USERS
GENERAL
MODULE
ATTACHMENT
DOWNLOAD
STANDARD
CODEBOOK
MEMBERSHIP
COMMITTEE
CONTROLLED
INITIATIVE
VOCABULARIES
DOCUMENTATION
INFORMATION
LEVEL
COMMUNITY
TIME
RESOURCES
SPECIFICATION
WORK
ORGANIZATION
ELEMENTS
WEB
MEMBER
SCHEMA
HOME
PUBLIC
SURVEY
SOCIAL
NEED
SUPPORT
ARCHIVE
PRESENTATIONS
UNIVERSITY
FORMAT
RECOMMENDED
ATTRIBUTE
PROVIDE
PROJECT
XML
NEW
TOOLS
IMPLEMENTATION
EXPERT
INTERNATIONAL
LINKED
EXAMPLE
VARIABLE
ICPSR
RESEARCH
CREATE
USE
REPRESENTATIVE
FILES
LIFECYCLE
DESCRIPTION
PRODUCTS
REVIEW
INSTANCE
PROCESS
PROPOSAL
STEERING
COMPARATIVE
STUDY
DISCUSSION
QUESTION
SOFTWARE
CHANGES
www.tagxedo.com

This is our moment!

[insert quote here on the importance of knowing everything everything]

Some tasks performed by the Metadata Librarian (1)

Routine tasks as a member of the Data Services Team:

- Regular meetings
- Planning report
- Data management guide
- Survey: create/distribute
- Group meetings with UW faculty/staff

Some tasks performed by the Metadata Librarian (2)

Professional development:

- UK Data Archive: "How to Run a Data Service"
- DDI tutorial
- TEI workshop
- Science Commons Symposium at Microsoft
- XML technologies study group (UW Libraries)
- Linked data project group (UW Libraries)
- Geospatial metadata workshop

Some tasks performed by the Metadata Librarian (3)

Digital humanities:

- Guest lectures on Digital Humanities
- Survey: personally invited eresearchers to participate
- Text markup (TEI) project

Some tasks performed by the Metadata Librarian (4)

Other stuff:

- ARL/DLF E-Science Institute
- Data Curation profiles workshop
- Analyzed ISO 191xx and FGDC CSDGM metadata and schemas
- Dialogue with UW DDI Alliance representative based in the Center for Studies in Demography and Ecology

Credits (1)

slide 12: Image generated by Wordle at http://www.wordle.net/ using an RSS feed from the DCMI Home page http://dublincore.org/ on June 20, 2012.

slide 13: cartoon by Betsy Elswit, accessed at http://www.library.cornell.edu/staffweb/kaleidoscope/volume19/april2011.html.

slide 16: Photo by psd, Paul Downey, accessed at his Flickr photostream at http://www.flickr.com/photos/psd/1428129861/.

slide 17: Image of "Data near Here" taken from http://web.cecs.pdx.edu/~vmegler/. This was part of a research project by Veronika M. Megler, a Ph. D. student in CS at the Maseeh College of Engineering & Computer Science, Portland State University, in Portland Oregon.

Slide 18: cover of Marin Dimitrov's Metadata management for the BBCs 2010 World Cup site using OWLIM, European Semantic Technology Conference 2010, found at http://pdfcast.org/pdf/metadata-management-for-the-bbc-rsquo-s-2010-world-cup-site-using-owlim.

Slide 19: University of Edinburgh web site, Information Services page on Data Documentation and Metadata. http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt/data-mgmt/data-documentation.

Slide 20: Jim Gray, David T. Liu, Maria Nieto-Santisteban, Alex Szalay, David J. DeWitt, Gerd Heber, "Scientific Data Management in the Coming Decade," in ACM SIGMOD Record, Vol. 34, No. 4, Dec. 2005, p. 34-41.

Credits (2)

Slide 21: Abstract accessed at http://adsabs.harvard.edu/abs/2011AGUFMIN41B1408M on June 22, 2012.

Slide 22: Denise Rogers, "Database management: Metadata is more important than you think!" in Database Journal, March 24, 2010. Accessed June 24, 2012 at http://www.databasejournal.com/sqletc/article.

php/3870756/Database-Management-Metadata-is-more-important-than-you-think.htm.

Slide 25: Taken from http://sharilopatin.com/2011/09/15/im-tired-of-writing/, Shar Lopatin's blog, Shari Lopatin: Rogue Writer. No image credits cited at that site.

Slide 34: Library Card taken from http://www.ccp.edu.ph/main/index.php/how-to-guides/locate-a-book at Central College of the Phillippines web site.

Slide 35: XSLT processing model taken from http://www.w3.org/Consortium/Offices/Presentations/XSLT_XPATH/#(1), "Overview of XSLT and Xpath," published by the W3C, 2005.

Slide 36: Handwritten model taken from http://dinesh-logbook.blogspot.com/2008/12/design-and-implement-data-model.html, the Dinesh LogBook, probably by Dinesh Kumar.

Slide 37: "Terminators" taken from http://warhammer40k.wikia.com/wiki/File:Librarian_Terminator_Blood_Angels.jpg which is an image page in http://warhammer40k.wikia.com/wiki/Warhammer_40k_Wiki, a wiki to the game Warhammer 40,000 maintained by Lead Administrator Montonius.

Credits (3)

Slide 38: Data repository diagram taken from http://www.w3.org/2001/sw/sweo/public/UseCases/FAO/ "Semantic Web Use Cases and Case Studies" by Gauri Salokhe, Margherita Sini, Johannes Keizer, Food and Agriculture Organization of the United Nations, Italy, and published by W3C, 2007.

Slide 39: Found at http://afeatheradrift.wordpress.com/2009/10/03/learning-stuff-i-dont-wanna-know/, "A Feather Adrift," Sherry Peyton's blog, but she attributes it to howtosplitanatom.com, Steve Spalding's blog, where the image appears at http://howtosplitanatom.com/news/how-to-improve-productivity/ with no attribution.

Slide 40: Dallas Museum of Art Metadata Standards Crosswalk available at http://www.museumsandtheweb.com/mw2007/papers/gutierrez/gutierrez.fig.11.pdf

Slide 41: DDI Word Cloud found at http://www.ddialliance.org/what and seems to have been generated at www.tagxedo.com.

IV. Moving Forward

In the upcoming year:

- Bring e-research services under a single umbrella = create a brand = increase visibility
- Repository platforms -- evaluate
- Digital humanities -- emphasize
- UW President launching campus-wide initiative on Research Data Management
- Planning document: Strategic Agenda -- implement

Strategic Agenda (1)

- Big Picture -- multi-year
- Evolved out of our involvement in the ARL/DLF E-Science Institute
- Still in draft status

Theo

Strategic Agenda--issues (1)

- Data storage, publication and citation
- Outreach and support, for example:
  - launch data management plan consultation services
  - develop data literacy curriculum for for students
- Data literacy training for librarians.

Strategic Agenda--issues (2)

- Institution-based research data policy committee
- Partnerships and collaboration
- Make use of SciVal to explore research and researchers

Strategic Agenda--issues (3)

- Identify facilities needed and request as appropriate; for example:
  - additions to the GIS lab; e.g. all machines in the labs should have data visualization tools
  - media production facilities
- Enter a partnership to support an already-existing data evaluation lab.

Conclusion

We have discussed:

- Background / Structure
- Our Approach
- Metadata
- Moving Forward

Key components for success:

1. Administrative Support
2. Outreach
3. Collaboration

Thank you!

Stephanie Wright, Data Services Coordinator: swright@uw.edu
Theo Gerontakos, Metadata Librarian: tgis@uw.edu
Matt Parsons, Geospatial Data/Maps Librarian: parsonsm@uw.edu

UNIVERSITY *of* WASHINGTON