

# **Research Data Management and Libraries**

**Opportunities, Challenges, & Best Practices**

Daniel C. Tsang  
Data Librarian  
University of California, Irvine  
Powerpoint for presentation at  
PRDLA 2012  
Peking University

# Research data management

- Many academic libraries worldwide now tackling research data generated by faculty
- Research data management means that data is organized, elements identified, and retrievable
- It allows data deposit in repositories
- Facilitates subsequent data sharing

# Opportunities for Libraries & Librarians

- Challenge prevailing notions of library service & of research
- Leave reference desk and library buildings
- Work more closely with faculty
- Get involved at beginning of a research project or life cycle of a dataset
- Re-energize one's job by learning new things

# Challenges

- Faculty disinterest in collaborating with librarian or in sharing data
- Lack of funding for data management
- Lack of awareness of scholarly communication advances
- Lack of subject knowledge
- Lack of time
- Lack of training
- Need to hire new knowledgeable staff

# Recent Developments

- DMPTool, a tool for complying with funding agency mandates for data management plans
- EZID, a tool for assigning identifiers to data files
- DataUp, a tool for sharing spreadsheet data
- Merritt, a data repository
- Data Curation Profiles developed
- Colectica, commercial software for tracking the life cycle of a dataset
- Institutional policies established for research data management and sharing

# UC3



[Staff Directory](#) [Contact CDL](#) [Report a Problem](#) [System Status](#)

October 14, 2012

[About CDL](#)

[Services and Projects](#)

[Information Gateways](#)

[Committees and Groups](#)

[News and Media](#)

[CDL Home](#) > [Services and Projects](#) > [UC3: University of California Curation Center](#)

## University of California Curation Center

[Merritt](#)

[EZID](#)

[Web Archiving Service](#)

[Data management plans and  
DMPTool](#)

[Consultation Services](#)

[Curation Wiki](#)

[National and International  
Initiatives](#)

[Blogs, Publications,  
Presentations and Webinars](#)

[About UC3](#)

[Contact Us](#)

## University of California Curation Center

UC3 helps researchers and the UC libraries manage, preserve, and provide access to their important digital assets.



Merritt is a repository service that lets you manage, archive, and share valuable digital content.

[More information](#)  
[Go to Merritt](#)



EZID (easy-eye-dee) makes it easy to create and manage unique, persistent identifiers.

[Go to EZID](#)



The Web Archiving Service allows you to capture, analyze, archive and publish web sites and documents.

[Go to WAS](#)



DMPTool helps researchers create and manage data management plans.

[More information](#)  
[Go to DMPTool](#)

### News

- [Open Data Explained](#)
- [Data to Receive Recognition from NSF](#)
- [DataUp is Live!](#)
- [NSF will consider research data as important as publications](#)
- [DMPTool monthly report, Sept 2012](#)
- [DMPTool unavailable Sunday Sept 23](#)
- [UC Librarians in Print](#)
- [Invitation to Data Curation Workshop — North and South](#)
- [California Digital Library and partners launch DataUp](#)
- [DMPTool Service Update — August 2012](#)

Want more information on any of these services? [Contact us!](#)

# DMPTool

- DMP stands for "Data Management Plan"
- [DMP Tool](#) is a tool developed by California Digital Library to help researchers prepare sponsored research grant applications that require data management plans. It provides an easy to use interface that:
  - Helps users create ready-to-use data management plans for specific funding agencies
  - Meets funder requirements for data management plans
  - Provides step-by-step instructions and guidance for how to manage data
  - Provides information about resources and services available at the local institution to help fulfill the data management requirements of a grant

# EZID

EZID makes it easy to create & manage unique, long-term identifiers. create identifiers for anything: texts, data, bones, terms, etc.

- store citation metadata for identifiers in a variety of formats
- update current URL locations so citation links are never broken
- use EZID's [programming interface](#) for automated operation at scale
- choose from a variety of persistent identifiers, including ARKs and DataCite DOIs
- [try out EZID](#) anytime and [contact the UC Curation Center](#) for more information
- Many libraries have licensed EZID for the entire university so that campus researchers can use it freely



# DataUp

An open source tool to help researchers document, manage, and archive tabular data with these features:

- Check for best practices
- Create metadata
- Get credit for data: obtain an identifier
- Archive & share data

# Merritt

Merritt is a repository service from that lets the University of California community manage, archive, and share valuable digital content

- Easy interface for deposit & update
- Tools for long-term data management
- Access via persistent URLs
- Permanent storage

# Data Curation Profiles Toolkit

According to Purdue where this project is located, Data Curation Profiles can:

- provide a guide for discussing data with researchers
- give insight into areas of attention in data management
- help assess information needs related to data collections
- give insight into differences between data in various disciplines
- help identify possible data services
- create a starting point for curating a data set for archiving and preservation

# Colectica



[Blogs](#) [News](#) [Evaluate](#) [Purchase](#) [Contact](#) [Help](#)

Software

Training

Services

Standards

Colectica® is the fastest way to design, document, and publish your survey research using **open data** standards.

[Learn about Colectica](#)

[See Colectica in Action](#)



## For Individuals

- [Colectica Reader](#)
- [Colectica Express](#)

## For Departments

- [Colectica Designer](#)
- [Colectica Repository](#)
- [Colectica Portal](#)
- [Colectica Workflow Services](#)

## For Institutions

- [Colectica Fusion](#)

## Software

The Colectica Platform is an ideal solution for statistical agencies, survey research groups, public opinion research, data archivists, and other data centric collection operations that are looking to increase the expressiveness and longevity of the data collected through standards based metadata documentation.

# Policies: University of Edinburgh

The University adopts the following policy on Research Data Management. It is acknowledged that this is an aspirational policy, and that implementation will take some years.

- Research data will be managed to the highest standards throughout the research data lifecycle as part of the University's commitment to research excellence.
- Responsibility for research data management through a sound research data management plan during any research project or programme lies primarily with Principal Investigators (PIs).
- All new research proposals [from date of adoption] must include research data management plans or protocols that explicitly address data capture, management, integrity, confidentiality, retention, sharing and publication.
- Etc...

# Policies:

## University of Wisconsin (2011)

- UW-Madison must retain research data in sufficient detail and for an adequate period of time to enable appropriate responses to questions about accuracy, authenticity, primacy and compliance with laws and regulations governing the conduct of the research. It is the responsibility of the *Principal Investigator* to determine what needs to be retained under this policy.



## Research Data Management

Tags: [data deposit](#), [data management](#), [data repositories](#), [data sharing](#), [research data](#)

Managing research data: Tools & resources

Last Updated: Sep 25, 2012

URL: <http://libguides.lib.uci.edu/datamanagement>

[Print Guide](#)

[RSS Updates](#)

[SHARE](#)

[Home](#)

[EZID](#)

[DataUp](#)

[Merritt](#)

[Data Management Plans](#)

[Data Repositories: Sciences](#)

[Data Repositories: Social Sciences](#)

[Data Repositories: Humanities](#)

[For Librarians](#)

[Home](#)

[Comments\(0\)](#)

[Print Page](#)

Search:

This Guide ▾

[Search](#)

### Data Management Tools

- [Tools & Services from UC Curation Center](#)

These tools and services include the following:

EZID - a service licensed to UCI researchers for identifying research files

DMPTool - a tool to help grant applicants fill out a concise data management plan geared to specific funding agencies. See tabs on this subject guide for more information.

[Comments \(0\)](#)

### Data Management & Sharing: Key Sites

Research Data is data generated in the research process. This subject guide delineates discipline-based research data repositories to help researchers and faculty seeking more information about potential repositories where they may be able to deposit their research data.

Careful Data Management of research data not only ensures preservation of such data but also helps in making such data shareable.

Sharing of research data contributes to scientific knowledge. Such data sharing has been the norm in some disciplines and is increasingly mandated by funding agencies, such as the National Science Foundation.

This guide is compiled by UCI Libraries' bibliographers who serve on the Libraries' Data Team, whose members are listed on the right column.

Listed below are some key resources:

### Subject Guide



Daniel C. Tsang

# Best Practices

ICPSR

## Guide to Social Science Data Preparation and Archiving

Best Practice Throughout the Data Life Cycle

•

5th edition





# Best Practices...



**MANAGING AND  
SHARING DATA**

**UK • DATA  
ARCHIVE**

# Looking Ahead



## Ten recommendations for libraries to get started with research data management

Final report of the LIBER working group on E-Science / Research Data Management

# LIBER recommendations

- Offer research data management support, including data management plans for grant applications, intellectual property rights advice and information materials. Assist faculty with data management plans and the integration of data management into the curriculum.
- Engage in the development of metadata and data standards and provide metadata services for research data.
- Create Data Librarian posts and develop professional staff skills for data librarian-ship.
- Actively participate in institutional research data policy development, including re-source plans. Encourage and adopt open data policies where appropriate in the re-search data life cycle.
- Liaise and partner with researchers, research groups, data archives and data centers to foster an interoperable infrastructure for data access, discovery and data sharing.

# LIBER recommendations...

- Support the lifecycle for research data by providing services for storage, discovery and permanent access.
- Promote research data citation by applying persistent identifiers to research data.
- Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.
- Get involved in subject specific data management practice.
- Offer or mediate secure storage for dynamic and static research data in co-operation with institutional IT units...

# IASSIST: Place to network



International Association for Social Science  
Information Services & Technology

[About](#) [Conferences](#) [IASSIST Quarterly](#) [Resources](#) [Community](#) [iBlog](#) [Contact](#)



The 38th IASSIST annual conference was held in Washington, D.C., USA, June 4 - June 8, 2012, hosted by the National Opinion Research Centre (NORC). The theme was Data Science for a Connected World: Unlocking and Harnessing the Power of Information. View [latest news](#).

**iassist brochure**



Available in:  
[English](#), [French](#), [Portuguese](#), [Spanish](#)

***IASSIST is an international organization of professionals working in and with information technology and data services to support research and teaching in the social sciences.***

***Its 300 members are from a variety of workplaces, including data archives, statistical agencies, research centers, libraries, academic departments, government departments, and non-profit organizations.***

[More Info](#)

**from our blog**

- [Social Sciences Librarian](#)
- [Data-related blog posts coming out of Open Repositories 2012 conference](#)
- [- Friday posting -](#)
- [A list of upcoming Big Data conferences](#)

[more](#)

**data tweets**



**iassistdata** @SusanCalman  
congrats on the high score on QI.  
1 day 23 hours ago.



**iassistdata** RT @archivetraining: Today the local organising committee for #iassist2013 had a productive meeting. Much to do, but wheels in motion, w ...  
2 days 11 hours ago.

**iassist quarterly**

**resources**

**community**

# Resources

- University of California Curation Center: <http://www.cdlib.org/services/uc3/>
- Data Curation Profiles: <http://wiki.lib.purdue.edu/display/dcp/Data+Curation+Profiles>
- LIBER's Ten Recommendations: <http://www.libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management>
- Colectica [commercial software]: <http://www.colectica.com/>
- Data Curation Profiles Toolkit: <http://datacurationprofiles.org/>
- University of Edinburgh research data management policy: <http://www.ed.ac.uk/schools-departments/information-services/about/policies-and-regulations/research-data-policy>
- University of Wisconsin policy on data stewardship, access and retention: [http://www.grad.wisc.edu/research/policyrp/policiesresdata\\_prod.html](http://www.grad.wisc.edu/research/policyrp/policiesresdata_prod.html)
- ICPSR: Guide to Social Science Data Preparation and Archiving: <http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf>
- UK Data Archive: Managing and Sharing Data: <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>
- IASSIST: <http://www.iassistada.org>
- UCI Research Data Management Subject Guide: <http://libguides.lib.uci.edu/datamanagement>