2013-10-22

DataCite Implementation Recommendations: A Report of the DataCite Task Force

Alvarez, Barbara; Campbell, Emily; Colman, Jason; Grochowski, Paul; Knott, Martin; MacEachern, Mark; Martin, Scott; Oehrli, Angela; Price, Rebecca; Sears, JoAnn; Sferdean, Fe; Turkel, Susan

https://hdl.handle.net/2027.42/100249

DataCite Implementation Recommendations

A Report of the DataCite Task Force

October 22, 2013

Data Cite Task Force Members:
Barbara Alvarez
Emily Campbell
Jason Colman
Paul Grochowski
Martin Knott (Chair)
Mark MacEachern
Scott Martin
Angela Oehrli
Rebecca Price
JoAnn Sears
Fe Sferdean
Susan Turkel

email: DataCiteTaskforce@umich.edu

University of Michigan Library DataCite Task Force

The DataCite Task Force's purpose is to investigate and implement DataCite's data citation services and standards throughout appropriate University of Michigan Library systems. The task force has discussed the workflows and processes for implementing DataCite services and outreach strategies to educate and build awareness of data citation on campus. This report presents the task force's recommendations for the Library.

CONTENTS

EXECUTIVE SUMMARY

DataCite (http://datacite.org/) is a non-profit, international consortium whose members collaboratively address challenges of making data visible and accessible. In the United States, DataCite is represented by three organizations: the California Digital Library (CDL), the Office of Scientific and Technical Information (OSTI), and the Purdue University Libraries.

As an addition to the growing network of services around data in the library, the University of Michigan Library has joined DataCite as a client through the Member or Allocator Agency, Purdue University Libraries. Purdue uses California Digital Library's EZID service (http://n2t.net/ezid) to offer DataCite Digital Object Identifiers (DOIs).

The DataCite Task Force of the University of Michigan Library recommends the following:

Recommendations

1. Assigned Responsibilities

To ensure the success of the data citation service and other aspects of data services, the task force recommends that a current member of the Research Data Services group be designated as the primary contact for subject specialists requiring additional support on data citation and other aspects of research data services.

2.1 Outreach: Branding & Marketing

The task force recommends that a MLibrary-branded submission form be made available to make it clear that the Library is providing the service. However, the form should also carry the EZID brand and an acknowledgement that the service is provided through the California Digital Library.

2.2 Outreach: Integration with Data Literacy

Instruction on data citation and DataCite should be integrated into workshops on data literacy rather than being taught as separate or standalone sessions.

3. Workflow

The taskforce recommends the establishment of three workflows:

a. EZID Form
b. Integration into Deep Blue Data Deposit
c. Integration of the EZID API into departmental data portals on campus to enable researchers to continue to use existing local storage while also gaining the benefits of a DOI and DataCite Metadata Store registration.

4. Dataset Discovery

a. The task force recommends that the library promote the DataCite Metadata Search service. Promotion of this data set discovery service should be targeted to the disciplines on campus that value the reuse of existing data to validate research.
b. An additional issue the task force recommends the Library leadership to discuss is the nature of local dataset discovery.

5. ORCID and ODIN

The DataCite Implementation Team should collaborate with the ORCID Implementation Team to monitor developments in ODIN's efforts and identify opportunities to develop local interoperability between the two registries.

6. DataCite Implementation Team

Charge an implementation team with furthering the previous five recommendations.

I. BACKGROUND

1. What is DataCite?

DataCite (http://datacite.org/) is a non-profit, international consortium founded in London in December 1, 2009 by the British Library, the Technical Information Center of Denmark, Delft University of Technology Library, the National Research Council's Canada Institute for Scientific and Technical Information (NRC-CISTI), California Digital Library, Purdue University, and the German National Library of Science and Technology. Several more institutions around the world have joined DataCite since 2009. To see the organizational structure of DataCite, see Appendix A.

DataCite aims to:

- Establish easier access to research data on the Internet
- Increase acceptance of research data as legitimate, citable contributions to the scholarly record
- Support data archiving that will permit results to be verified and re-purposed for future study

The DataCite structure consists of a Managing Agent, which is currently the German National Library of Science and Technology, with Members and Associate Members around the world. (See Appendix A: *The DataCite Structure and Membership*)

Every six months, members of DataCite collaboratively address challenges of making data visible and accessible. DataCite works to support:

- Researchers by helping them find, identify, and cite research datasets with confidence
- Data centers, libraries or data publishers by providing persistent identifiers for datasets, workflows and standards for data publication
- Publishers by enabling research articles to be linked to the underlying data

In the US, DataCite is represented by CDL, OSTI, and Purdue University Libraries.

2. How Does DataCite Work?

DataCite provides an infrastructure for assigning Digital Object Identifiers (DOIs) to datasets and registration of associate metadata. It runs a central open metadata repository of scientific records (http://search.datacite.org/ui). DataCite Members enable data centers and libraries to assign DOIs to their data. At present, DataCite leverages DOIs commonly

used for identifying journal articles, but other identifiers, such as Archival Resource Keys (ARKS) are also used by Member organizations. MLibrary has decided to undergo a DataCite test phase as a client through the Member or Allocator Agency, Purdue University Libraries. Purdue provides the EZID service (http://n2t.net/ezid) offering DataCite DOIs as well as ARKs to educational, non-profit, and commercial clients. (See diagram, Page 14: The *DataCite, Member and Client Responsibilities*.) For a fuller illustration of how a researcher might interact with DataCite, see Appendix A.

3. Why is DataCite Important?

Data are starting to be recognized as a primary and vital output of the research process rather than a by-product. Even though data has often been shared in the past, they were seldom cited. Now there is a growing need to reference data in the same routine way as a researcher provides a bibliographic reference to a journal article or book. DataCite allows proper data citation, which validates data in scholarly communications and powers the sharing and re-use of data. Data citation has many short-term and long-term benefits:

Short-Term Benefits	Long-Term Benefits
Ease of locating the data	Creates a publishing structure that ensures data are available long into the future
Re-using and verifying data and building on results	Data citation makes data discovery easier
Gives credit to the data producer when data are re-used	Credit, proof of completion of grant and a possible justification for funding bodies to advance awards
Promotes reproduction of research results	Impact of datasets and of data producers can be measured
If dataset links to a paper, this allows the data to be readily put into context	Increased citations leads to increased professional recognition

Investing in a data citation initiative, such as DataCite, is one of the ways MLibrary can support the University researchers along the Research and Data Lifecycles, particularly through the stages of: archive and preservation, giving access to data, and the discovery and re-use of data. To support researchers in the rapidly changing culture towards data, establishing a dataset identifier assignment and dataset discovery service would provide University researchers with the valuable short-term and long-term benefits of data citation that lead to greater research productivity and discovery.

II. TASK FORCE CHARGE, RECOMMENDATIONS, AND DISCUSSION

1. Task Force Charge

The purpose of the DataCite Taskforce was to test, evaluate, investigate, and implement DataCite services and standards throughout appropriate library systems. The taskforce was charged to:

- Become a team of DataCite experts
    - Through reading and discussing what DataCite services are and why they are important
    - Define what it means to have and implement DataCite services
    - Test and evaluate DataCite Services
- Identify and prioritize library systems that are good candidates for DataCite services implementation
- Establish and document workflows and processes that will effectively run and sustain DataCite services
- Partner with Research Working group to build awareness about DataCite services
- Make recommendations for how the Library can identify partners from outside of the library to test/pilot DataCite services in research environments

The taskforce was educated through readings such as "Introduction to DataCite," exploring the DataCite website, and discussing the DataCite organizational structure, membership, and client and member responsibilities. For information and resources about DataCite please see Appendices A and B. We discussed the importance of data citation practice for researchers and compiled learning resources and new developments in data citation and DataCite. The Taskforce discussed how the Library would implement DataCite services and tested the services as provided by DataCite Full Member, Purdue University Libraries.

We identified the Deep Blue repository as the Library's primary system for implementing the DataCite services. The group also discussed human and infrastructure level workflows and processes that would oversee and operate the DataCite services for the Library, including ways to build awareness of these services on campus. We also discussed how the Library could partner with departmental research data portals to pilot and test the DataCite services. These will be discussed in the following section.

2. Recommendations and Discussion

1. Assigned Responsibilities

To ensure the success of the data citation service and other aspects of data services, the task force recommends that a current member of the Research Data Services group be designated as the primary contact for subject specialists requiring additional support on data citation and other aspects of research data services. This recommendation grew out of the concern amongst some task force members that they do not currently know who they should contact should a researcher have a data services-related question they are unable to answer. This "data guru" could also be responsible for leading outreach and promotion of research data services. It is unclear to the task force if these responsibilities would be appropriate for the proposed data management specialist position.

2. Outreach

The task force discussed different ways to raise awareness and build a culture of data citation in the U-M campus through the DataCite services.

2.1 Branding and Marketing

The University Library is required by its DataCite services arrangement with Purdue to use the EZID provided by the California Digital Library (CDL) to generate DOIs and submit metadata to DataCite. EZID, pronounced "easy-ID," is CDL's persistent identifier service created originally for the University of California system but now available to other research centers, institutions, repositories, and publishers. The task force recommends that a MLibrary-branded submission form be made available to make it clear that the Library is providing the service. However, the form should also carry the EZID brand and an acknowledgement that the service is provided through CDL. Library administration would need to approve the wording for the acknowledgement. Marketing of the DataCite services should emphasize the DOI because it is a concept that is well known and valued by researchers. Like DOIs for electronic journals and electronic books, DataCite DOIs reliably link to data sets even if the registered data files' web addresses change. This marketing approach was strongly recommended by Paul Bracke, Associate Dean of Libraries for Assessment and Technology at Purdue University. Data citation as a general concept should be the focus with DataCite presented as the service the Library utilizes to enable data citation.

2.2 Integration with Data Literacy

Instruction on data citation and DataCite should be integrated into workshops on data literacy rather than being taught as separate, stand-alone sessions. The task

force found it difficult to envision effectively covering the topic without also addressing data management plans and data repositories. Some members of the task force have agreed to participate in the development and delivery of data literacy training during the Fall 2013 term. Rebecca Price and Scott Martin drafted a trifold handout, "DataCite @MLibrary" (see Appendix C) that educates both internal and external audiences about data citation and DataCite. The handout is based on "Introduction to DataCite" (Appendix A) drafted by Fe Sferdean, CLIR data curation fellow.

3. Workflow

Based on information from Purdue University, CDL, and other institutions utilizing EZID, the task force is recommending the establishment of three workflows. It should be noted that Michigan Publishing already generates DOIs for its published works through CrossRef. While CrossRef charges Michigan Publishing on a per-title basis, DataCite allows unlimited DOIs generation for members and their clients.

3.1 EZID Form

The EZID Form will be used for single DOI generation and DataCite registration through the CDL EZID service. Long term, this form will not be the Library's primary means for DataCite registration and DOI generation due to its standalone nature but it does allow immediate registration of individual datasets. The DSpace programmer (José Blanco) utilized CDL's existing EZID form to create a version to be MLibrary-branded and hosted. The original form allows the creation of multiple types of DOIs but at the request of the task force, the DSpace programmer customized the form to preselect DataCite and DOI options.

Use of the EZID form could be unmediated, allowing individual researchers or their support staff to complete the form. However, at this phase of service development, a mediated process that relies upon a library staff member to complete the form seems more appropriate. In either case, responsibility for ongoing maintenance must be designated. Theoretically, an institution can have the expectation that individual researchers or their support staff will take on this responsibility and update the DataCite registry via the EZID interface should the location of the dataset(s) change but this would not be practical for most researchers and is therefore unrealistic. The task force recommends that the eventual Data Management Specialist and subject specialists work closely with the Technical Services Division, leveraging Technical Services' metadata, and workflow management expertise to establish an efficient DataCite registry submission and maintenance operation.

3.2 Integration into Deep Blue Dataset Deposit

We recommend that when the dataset is deposited into Deep Blue, the submitter has the option to generate a DataCite DOI and register associated metadata to make their dataset discoverable through the DataCite Metadata Search, a catalog of all DataCite metadata at search.datacite.org. A simple, one-step process for the submitter (e.g., a check box on the Deep Blue deposit form) seems to be the ideal way to achieve this. The DSpace programmer believes it would be simple to design a process that can complete the deposit and Deep Blue handle generation step then use the EZID API to generate a DataCite DOI using the newly created handle and accompanying metadata. There may also be a need to process batches of datasets. This workflow cannot be implemented until Deep Blue is modified to accept very large files, but the EZID API does support batch processing.

3.3 Departmental Data Portals

Integration of the EZID API into departmental data portals on campus would enable researchers to continue using existing local storage while also gaining the benefits of a DOI and DataCite Metadata Store registration. Through the work of Scott Martin. Biological Sciences Librarian, the University of Michigan BioStation (UMBS) is interested in assigning DOIs to the datasets created by its researchers. Jason Tallant, the new IT Manager for the BioStation, is in the early stages of testing the integration of the EZID API into the UMBS Data Portal's existing deposit workflow to seamlessly add the ability to generate DOIs.

3.4 Workflow Implementation

Additional steps will be necessary to establish these recommended workflows. Task Force Chair Martin Knott and Research Data Services Director Jennifer Green agreed that a small team should be formed to carry out the implementation. The team will be disbanded following the establishment of the three workflows and completion of related work.

4. Dataset Discovery

The less discussed and often overlooked aspect of DataCite is the Metadata Store, which serves as the dataset discovery tool. Although dataset sharing and reuse are part of the research lifecycle and two of the DataCite initiative's goals, these two uses are not yet strong selling points for the service. The task force recommends that promotion of these aspects of DataCite and data citation be targeted to the disciplines on campus that value the reuse of existing data to validate research.

An additional issue the task force recommends the Library leadership discuss is the nature of local dataset discovery. ICPSR datasets and purchased datasets are discoverable in the library catalog, but the catalog contains only a small amount of the content in Deep Blue, where Library-hosted datasets will eventually reside. This topic could be part of a more general discussion on what should be in the Library's catalog and the relationship between our various discovery tools.

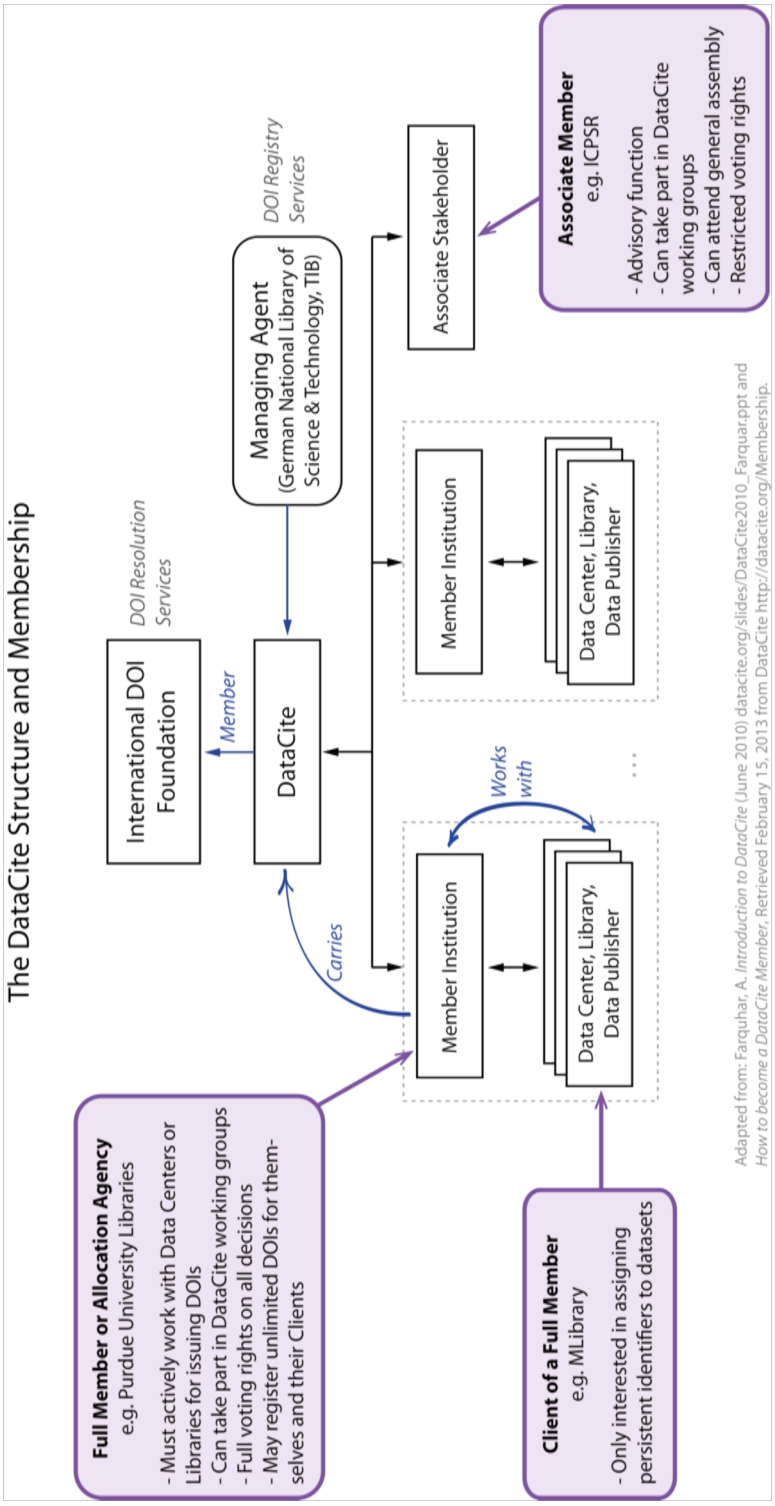
5. ORCID and ODIN

During the time this task force became active, a second Library-charged task force evaluated the recently announced Open Research and Contributor ID (ORCID). The ORCID initiative (http://orcid.org,) is an open, non-profit, community-based effort to provide both a registry of unique researcher identifiers and a method for linking research outputs to the identified researcher. Due to the obvious relationship between DataCite and ORCID, in September 2012, the European Union has commissioned the ORCID and DataCite Interoperability Network (ODIN) (http://odin-project.eu/), a two-year project to build links across scholarly communication services between DataCite and ORCID. The DataCite Implementation Team should collaborate with the ORCID Implementation Team to monitor developments in ODIN's efforts and identify opportunities to develop local interoperability between the two registries. See the ORCID Taskforce Report: http://deepblue.lib.umich.edu/handle/2027.42/98975

6. DataCite Implementation Team

The final recommendation from the Task Force is to charge an implementation team with furthering the previous five recommendations. Members of the implementation team will be charged to participate in planning data training for librarians, finalizing data citation marketing materials for the campus, identifying additional campus data portals for EZID API integration, and overseeing the integration of the EZID API into the Library-hosted data repository. The DataCite Implementation team will also work with the ORCID Implementation team to explore opportunities for local interoperability.

APPENDIX A

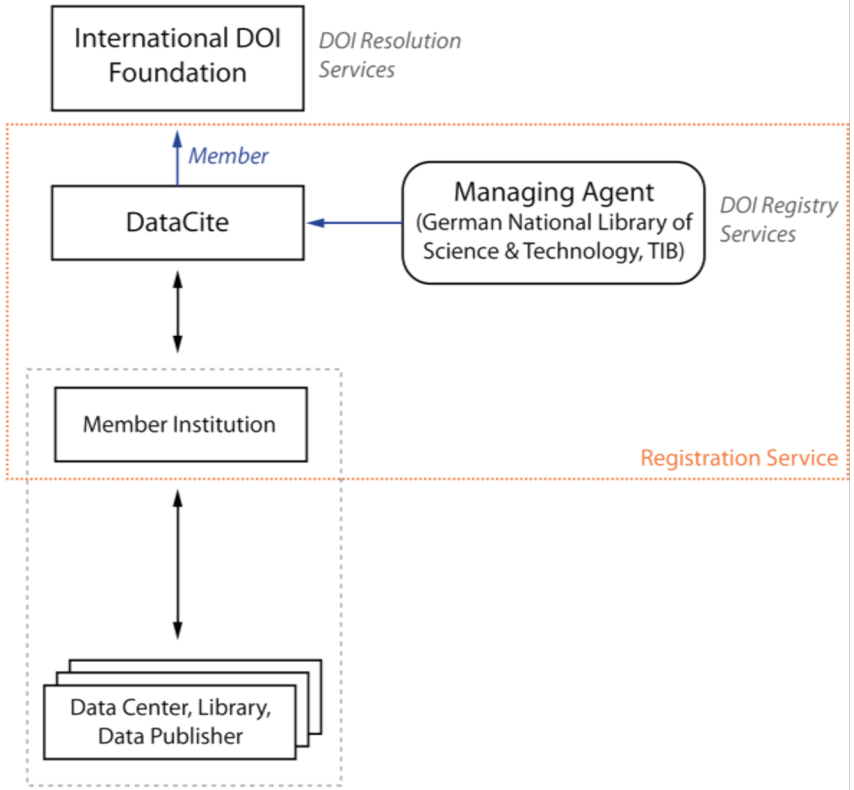

The DataCite Structure and Membership
International DOI Foundation
DOI Resolution Services
Member
DataCite
Managing Agent
(German National Library of Science & Technology, TIB)
DOI Registry Services
Associate Stakeholder
Member Institution
Data Center, Library, Data Publisher
Works with
Carries
Full Member or Allocation Agency
e.g. Purdue University Libraries
- Must actively work with Data Centers or Libraries for issuing DOIs
- Can take part in DataCite working groups
- Full voting rights on all decisions
- May register unlimited DOIs for themselves and their Clients
Client of a Full Member
e.g. MLibrary
- Only interested in assigning persistent identifiers to datasets
Associate Member
e.g. ICPSR
- Advisory function
- Can take part in DataCite working groups
- Can attend general assembly
- Restricted voting rights
Adapted from: Farquhar, A. Introduction to DataCite (June 2010) datacite.org/slides/DataCite2010_Farquar.ppt and How to become a DataCite Member, Retrieved February 15, 2013 from DataCite http://datacite.org/Membership.

The DataCite, Member and Client Responsibilities

DataCite Responsibilities:
- Maintains the resolution infrastructure
- Maintains a searchable database of metadata
- Manages identifiers over the long-term
- Establishes and shares best practice

Member or Allocator Agency Responsibilities:
- Have a contract stating responsibilities of member and client including mandatory requirements for metadata
- Develop their own services for clients
- Determine costs and fee schedules
- Provide DOI prefixes to their clients

Client Responsibilities:
- Committment to data persistence
- Create and update metadata
- Provide landing pages for all registered data
- Provide tombstone pages for data that becomes unavailable
- Define granularity of their data
- Create their own DOI syntax



Adapted from: Farquhar, A. *Introduction to DataCite* (June 2010) datacite.org/slides/DataCite2010_Farquar.ppt and *Business Models Principles* (October 2012) by Members of the Business Practices Working Group from DataCite http://datacite.org/node/74.

General Example: A Researcher Deposits a Dataset

A Researcher wants to deposit her dataset in U-M's Institutional Repository, Deep Blue, hosted by MLibrary and get an identifier for it. She follows guidelines presented on the Deep Blue website and/or obtains help from a data citation representative from MLibrary. She uses the EZID DOI generator form, which includes entering metadata:

DataCite has mandatory metadata elements:

- ***Creator**: The main researchers involved on the data, or the authors of the publication in priority order. This may be a corporate/institutional personal name. Format: Family, Given*

  Researcher types "`Smith, Jane`"

- ***Title**: A name or title by which a resource is known*

  Researcher types "`Dataset: Chemical composition of sediments from Io751 site`"

- ***Publisher**: A holder of the data (including archives) or institution, which submitted the work. (Automatically entered by Deep Blue system)*

- ***Publication Year**: Year when the data are made publicly available. If an embargo period has been in effect, use the date when the embargo period ends.*

  Researcher types "`2013`"

The Researcher then submits her dataset and metadata into the Deep Blue IR system.

MLibrary's Deep Blue technical personnel checks the metadata, creates an identifier, and sends a Registration Request to the DataCite Member organization, Purdue University Libraries. DataCite registers the DOI and associated metadata.

On the Deep Blue website, the Researcher would see an identifier on her dataset. For example:

doi: 10.5897/D3JQ8ST8G

The dataset would also have a citation in the format:
Creator (Publication Year): Title. Publisher. Identifier (DataCite recommends that DOI names are displayed as linkable, permanent URLs)

The Researcher would see the citation displayed. For example:

Smith, Jane (2013): Dataset: Chemical composition of sediments from Io751 site. Deep Blue. http://dx.doi.org/10.5897/D3JQ8ST8G

APPENDIX B

DataCite and Data Citation Resources

1. Guides

- ICPSR Guide to Data Citation: http://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/citations.jsp
- Australian National Data Service (ANDS) Data Citation Guide http://www.ands.org.au/cite-data/index.html
- IASSIST Special Interest Group Quick Guide to Data Citation: http://iassistdata.org/community/sigdc
- IASSIST is the International Association of Social Science Information Services and Technology. It is a 300-member organization of data librarians, archivists, and others.) List of resources: http://www.iassistdata.org/community/data-citation-ig/data-citation-resources

2. Readings

- *Data Citation Awareness* (May 3, 2011). Retrieved January 28, 2013, from the Australian National Data Service: http://ands.org.au/guides/data-citation-awareness.html
- Ball, A., Duke, M. (2011). *Data Citation and Linking.* DCC Briefing Papers. Edinburgh: Digital Curation Center http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/data-citation-and-linking
- Brase, J. (July 2010) *DataCite – A global registration agency for research data*. German National Library of Science and Technology (TIB). http://www.ilds2009.eu/fileadmin/user_upload/Full_text/DataCite_Brase_COINFO.pdf
- Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. (2011). Determining Data Information Literacy Needs: A Study of Students and Research Faculty. Portal: Libraries & The Academy, 11(2), 629-657. http://search.ebscohost.com/login.aspx?direct=true&db=lii&AN=503012899&site=ehost-live

3. Videos

- DataCite has a number of videos on YouTube mostly from 2012 Summer Meeting: http://www.youtube.com/user/DataCite

- Eefke Smit, STM Association on metadata and discovery: http://www.youtube.com/watch?v=HuIjwv5q900

4. Workshops & Webinars

- The British Library has been holding DataCite Workshops and have posted the Powerpoint slides on their website: http://www.bl.uk/aboutus/stratpolprog/digi/datasets/workshoparchive/archive.html#ws1
- Resources from the California Digital Library on EZID: http://n2t.net/ezid/home/outreach

APPENDIX C

Glossary

ARK- An archival resource key (ARK) is a URL-based persistent identifier that links to the identified object, its metadata, and a statement on the intended persistence of the object at that Internet location. An ARK is easily recognized by the presence of the ***ark:*** label within the URL. The basic ARK URL takes the user to the object. Appending a single question mark (?) instead retrieves the metadata record. Appending two question marks (??) returns the persistence statement. One of the principles on which ARKs are built is that persistence is achieved through reliable access service, rather than fundamental characteristics of the object or its naming syntax. The structure of the ARK allows the identifier to remain unchanged even if it is hosted by a different organization under a new name mapping authority. In the following ARK example, http://portico.org/ark:/13030/tf5p30086k, portico.org is the name mapping authority. More information is available at https://confluence.ucop.edu/display/Curation/ARK.

DOI- A digital object identifier (DOI) is a persistent, interoperable, and actionable hyperlink that serves as a digital identifier for an object. The identified object can be digital, tangible, or abstract. DOIs are commonly used for ebooks and ejournals. DataCite utilizes DOIs for datasets (e.g., http://dx.doi.org/10.5167/UZH-22097). A DOI utilizes the Handle System (http://www.handle.net/) to resolve the registered identifier to an Internet address where the object or descriptive information is located. The DOI registry, managed by the International DOI Federation (IDF), relies upon a numbering syntax, a data model, and established implementation procedures to provide consistent functionality. A federation of registration agencies assigns DOIs at the request of publishers and other content providers. Metadata describing the registered objects is required and the registering organization is required to maintain that metadata. More information on the history, functionality, and governance of DOIs can be found at http://www.doi.org/.

EZID- EZID (easy ID) is a service provided by the University of California Curation Center of the California Digital Library to enable the creation and management of long-term identifiers. DataCite DOIs and ARKs are supported. Originally created to serve the University of California system, EZID now generates identifiers for international data repositories, university libraries, and research centers. More information is available at http://n2t.net/ezid/.

APPENDIX D

Outreach Material

The following pages show a trifold handout drafted by Rebecca Price and Scott Martin called, "DataCite @MLibrary," that would serve to educate library staff and campus researchers about data citation through DataCite services at MLibrary. Once the technical infrastructure and workflows are established, more details will be described in the instructional sections of the trifold handout such as under "Getting Started with DataCite." We discussed the necessity of having a representative who would serve as the primary contact point for MLibrary's data citation services. The task force also discussed the plan to place the EZID brand on future educational and outreach handouts as well.

Data Citation @ MLibrary

MLibrary is undertaking an initiative to support data citation by researchers across the university community. The Library has chosen to use DataCite to enable the registration of and access to data.

DataCite (http://datacite.org/) is a non-profit, international consortium founded in 2009. It aims to:

Establish easier access to research data on the Internet

Increase acceptance of research data as legitimate, citable contributions to the scholarly record

Support data archiving that will permit results to be verified and re-purposed for future study

Investing in a data citation initiative, such as DataCite, is one of the ways MLibrary can support University researchers with their research and data lifecycles, particularly through the tasks of discovery and re-use of data, access to data, and archiving and preservation of data. Expectations surrounding data are changing rapidly. By assigning a dataset identifier and enabling dataset discovery, the Library will provide University researchers with valuable short and long-term benefits of data citation, leading to greater research productivity and discovery.

…more about data citation here? Or something on EZID here…

Getting Started with DataCite

Basic directions about the form/process for registering data - and a link to it…

Why is DataCite Important?

Data is a primary and vital output of the research process. There is a growing need, if not a mandate, to reference data in the same routine way as a researcher provides a bibliographic reference to a journal article or book. DataCite allows proper data citation, which validates data in scholarly communications and powers the sharing and re-use of data.

Short Term Benefits of DataCite

- *Increases ease of locating data*
- *Enables reuse and verification of data and further building on results*
- *Gives credit to the data producer when data is reused*
- *Promotes reproduction of research results*
- *Allows data to be put readily into context, when the dataset links to a paper*

Long Term Benefits of DataCite

- *Creates a publishing structure that ensures data is available long into the future*
- *Services around data citation make it easier for researchers to discover relevant*
- *Reduces danger of rival researchers failing to give credit to those researchers who cite their data openly*
- *Enables measurement of impact of datasets and of data producers*
- *Improves potential for researchers to gain professional recognition and rewards for published data.*

A Researcher Deposits a Dataset

A researcher wants to deposit her dataset in the UM Institutional Repository, Deep Blue, and get an identifier for it.

She follows the protocol for depositing her dataset into Deep Blue. MLibrary's Deep Blue personnel will check the metadata, create an identifier, and send a Registration Request to the DataCite Member organization. DataCite registers the DOI and associated metadata.

The researcher will see the citation displayed:

Smith, Jane (2013): Dataset: Chemical composition of sediments from Io751 site. Deep Blue. http://dx.doi.org/10.5897/D3JQ8ST8G

For more information about DataCite, see: <fill in name or unit when we have one>