11 December 2014

Trusted Repository Certification and the National Transportation Library

Mary Moulton, Digital Librarian, National Transportation Library

Trusted Repository Certification and the National Transportation Library

Mary E. Moulton
Digital Librarian, National Transportation Library

Transportation Librarians Roundtable
December 11, 2014

About the speaker: Mary Moulton

- Education
  - BS Plant Science with concentration in Plant Pathology and Entomology, University of Rhode Island
  - MS Entomology, Kansas State University
  - MLIS University of Rhode Island
- Work experience
  - Chemical Sciences Librarian (Binghamton University, Ohio State University, Lehigh University)
  - Information Architect (Air Products and Chemicals, Merck & Co.)
- Began at NTL in May, 2011. Responsible for leading the development and enhancement of digital repository services, information organization, digital curation, and tools that facilitate discovery and use of NTL resources.

What is a Trusted Digital Repository?

- A trusted digital repository is one whose mission is to provide reliable, **long-term access** to managed **digital resources** to its **designated community**, now and in the future.

  - RLG 2002, Trusted Digital Repositories: Attributes and Responsibilities

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology

What is an OAIS?

- An Open Archival Information System (OAIS) is an archive, consisting of an organization, which may be part of a larger organization, of **people** and **systems** that has accepted the responsibility to preserve information and make it available for a **designated community**.

  – CCSDS 2012, [Reference Model for an Open Archival Information System (OAIS)](#)

U.S. Department of Transportation

Office of the Assistant Secretary for Research and Technology

Timeline of Initiatives

- 1968: Electronic Records Program established at National Archives and Records Services (NARS).

- 1975: NARS Data Archives Branch issues *Catalog of Machine-Readable Records in the National Archives of the United States.*

- 1988: DoD National Computer Security Center (NCSC) issues *A Guide to Understanding Audit in Trusted Systems.*

- 1994: Commission on Preservation and Access/Research Libraries Group Task Force on Archiving Digital Information formed – defines "born digital" and "re-born digital" to distinguish information created in electronic format from materials electronically reformatted.

- 1998: Fedora reference implementation developed at Cornell and made available to researchers and the public.

- 2002: CCSDS publishes *Reference Model for an Open Archival Information System (OAIS).*

- 2002: MIT Libraries and HP Labs releases first public version of DSpace open source repository software.

- 2003: RLG-NARA Task Force on Digital Repository Certification formed.

CCSDS

- Audit and Certification of Trustworthy Digital Archives. Consultative Committee on Space Data Systems (CCSDS). September 2011. http://public.ccsds.org/publications/archive/652x0m1.pdf

- Text also available as ISO 16363:2012 Space data and information transfer systems -- Audit and certification of trustworthy digital repositories

- Based on the CCSDS work on the Open Archival Information System (OAIS) Reference Model for digital archives.

- NASA, NOAA, USGS and international space agencies involved

RLG - OCLC

- Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report. Mountain View, CA: Research Libraries Group (RLG), May 2002. http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf?urlm=161690
  - basis for ISO standard 16363
- Trustworthy Repositories Audit & Certification: Criteria and Checklist. Version 1.0. Chicago: Center for Research Libraries (CRL), February 2007. http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf
  - 84 criteria for trustworthiness
  - has been used to develop many tools and metrics that assess repositories
- Downloadable checklist is available in Excel. Four repositories have been certified by Center for Research Libraries (CRL) using this process. http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying-0
  - Hathi Trust Preservation repository and access services provided by University of Michigan for works in public domain
  - Scholars Portal Ontario Council of University Libraries shared technology infrastructure and shared collections for 21 university libraries in the province
  - Portico Digital preservation service provided by Ithaka for e-journals, e-books, and other electronic scholarly content
  - Chronopolis   Digital preservation network developed by a national consortium, initial funding from LoC

Nestor

- Nestor (Network of Expertise in Long-Term Storage of Digital Resources) Catalogue of Criteria for Trusted Digital Repositories. December 2006.
  http://www.dcc.ac.uk/resources/repository-audit-and-assessment/nestor#sthash.aYz1u2bO.dpuf

- The basic principles for applying the criteria catalogue to determine trustworthiness are: Adequacy, Measurability, Documentation and Transparency.

- Draft for public comment
  http://files.d-nb.de/nestor/materialien/nestor_mat_08-eng.pdf

DCC and DPE

- Digital Curation Centre (DCC) and Digital Preservation Europe (DPE) developed DRAMBORA (Digital Repository Audit Method Based On Risk Assessment) toolkit in 2007
  http://www.repositoryaudit.eu

- Based on repository pilot audits undertaken by the DCC, conducted at a diverse range of organizations including national libraries, scientific data centers and cultural and heritage data archives

- Intended to facilitate internal audit by providing repository administrators with a means to assess their capabilities, identify their weaknesses, and recognize their strengths

ICSU World Data Systems

- ICSU World Data Systems - created by a decision of the 29th General Assembly of the International Council for Science (ICSU). June 2011 . http://icsu-wds.org/files/wds-certification-summary-11-june-2012.pdf

- Membership in the WDS requires an evaluation and certification process for the data system which is based on NESTOR and TRAC.

Comparison

Research Data Alliance

- Repository Audit and Certification DSA–WDS Partnership Working Group – Chairs: Lesley Rickards, Mary Vardigan
  https://rd-alliance.org/group/repository-audit-and-certification-ig-dsa%E2%80%93wds-partnership-wg.html

- ISO certification not feasible for all repositories – working on basic certification that is not too onerous

- Using Data Seal of Approval and World Data Systems criteria to develop a common certification standard
  - Focus on criteria rather than process
  - Simplify assessment options
  - Stimulate more certifications
  - Increase impact on the community

Data Citation Index

- Repository Evaluation, Selection, and Coverage Policies for the Data Citation Index Within Thompson Reuters Web of Knowledge http://thomsonreuters.com/products/ip-science/04_037/dci-selection-essay.pdf

- When a repository is selected for coverage, all deposited data is included in *Data Citation Index.* Criteria for inclusion:

  - Persistence and stability - repository must demonstrate longevity, have steady flow of newly deposited data, data should be deposited with the repository rather than holding metadata and a web link to a remote/external source for the data

  - Funding statements

  - English language metadata

  - Peer review

  - Links to the research literature

Challenges for Open Access Repositories

- Discovery and reuse
  - Maintenance of persistent identifiers and urls
  - Registries of repositories
- Limited lifespans for formats, software and devices
- Interoperability
  - Standard metadata, data dictionaries
  - Ability to copy from one repository to another
- Preserve for how long? Indefinite future, 5 years, or eternity?
  - Funding commitments
  - Migration
- Changes in scholarly norms
  - Citation
  - Intellectual property rights
  - Impact factors

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology

15

Opportunities

- MAP-21 (49 USC 6304) and NTL role
  - Acquire, preserve and manage transportation information and information products and services for use by DOT, other Federal agencies, and the public
  - Central repository for DOT research results and technical publications
  - Central clearinghouse for transportation data and information of the Federal Government

- White House OSTP Memo on Public Access – *Increasing Access to the Results of Federally Funded Scientific Research*:
  - The direct results of federally funded science, including peer-reviewed publications and digital data, will be made freely accessible to the public to the greatest extent possible, consistent with law and other objectives, including the continued availability of peer review.
  - Each Federal agency with over $100M in annual research expenditures will develop a plan for public access, after consultation with stakeholders including researchers, universities, libraries, publishers, and representatives of other potential users.

NTL Digital Repository and Services

- Digital Repository
  - Purpose: serve as central clearinghouse and archive for public access transportation information
  - Special features: access to specialized collections and integration of USDOT repositories in our platform
    - Volpe Reports
    - Directory of Transportation Data Sources
    - NHTSA Behavioral Safety
    - ITS JPO Electronic Library
  - Services: digitization activities, preservation & ensuring perpetual access; permanent URLs, links incorporated into TRID records

U.S. Department of Transportation

Office of the Assistant Secretary for Research and Technology

NTL Digital Repository - Gaps

Current

- Homegrown OAI-PMH compliant repository (Open Archives Initiative Protocol for Metadata Harvesting)

- In-house hardware

- Manual analysis and application of metadata

- Repository search

- Self governance using OAIS model

Future

- Open source platform
  - Evaluating Islandora (Drupal/Fedora)
  - Instance from CDC

- Cloud hosting

- Processing enhanced with auto-indexing

- Federated search and discovery

- Compliance with guidelines for Federal Repositories

Thank you and questions

U.S. Department of Transportation

Office of the Assistant Secretary for Research and Technology