

Trustworthy Digital Repositories



Jessica Tieman, Digital Preservation Librarian, GPO

Mary Moulton, Digital Librarian, National Transportation Library (NTL), DOT



A closer look at trustworthiness

- What does it mean to be trustworthy?
- A history of requirements for digital repositories
- Why pursue digital repository certification?
- Current digital repository certification efforts



What does it mean to be trustworthy?

- A “Trustworthy Digital Repository” is...
 1. The actualization of long-term strategic values and commitments to long-term preservation
 2. The realization of repository and community transparency, reliability, and confidence
 3. A repository which can be objectively verified against internationally recognized standards and best practices



Repository Trustworthiness: An Organizational Commitment

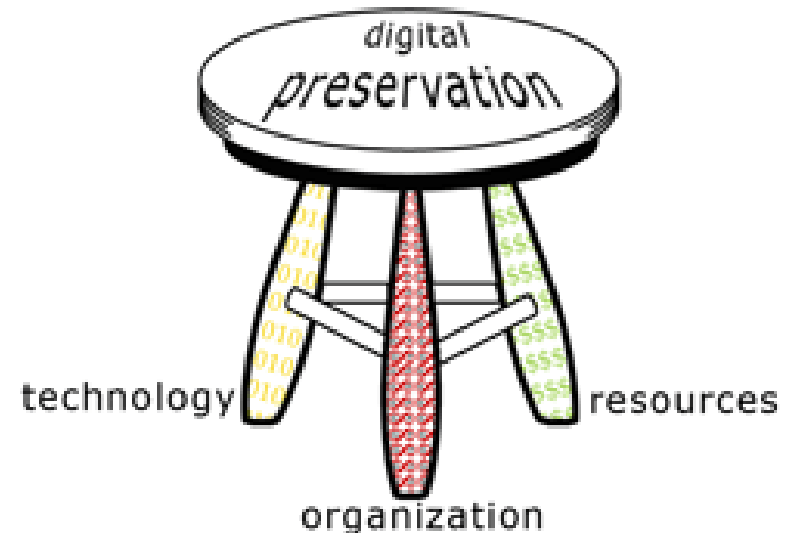
- 1998 Research Libraries Group (RLG) survey “Digital Preservation Needs and Requirements in RLG Member Institutions” Two-thirds of respondents reported not having written policies for digital preservation despite being responsible for preserving digital materials
- 2001 Digital Library Federation (DLF) Survey reported fourteen of twenty-one libraries not having digital preservation policies
- **The Five Organizational Stages of Digital Preservation**
Anne R. Kenney & Nancy Y. McGovern bridges the gap to implement effective programs for preserving digital materials

DPM Five Stages and the Three-legged Stool:

Kenny and McGovern's "The Five Organizational Stages of Digital Preservation" suggests that an organization's digital preservation program:

- Needs to fit defines needs, requirements, and resources
- Requires ongoing and iterative development
- Should reflect best practices and standards

Early steps of self-assessment:
Institutional readiness surveys regarding organizational infrastructure, technological infrastructure, and resources (contingency funding etc.)

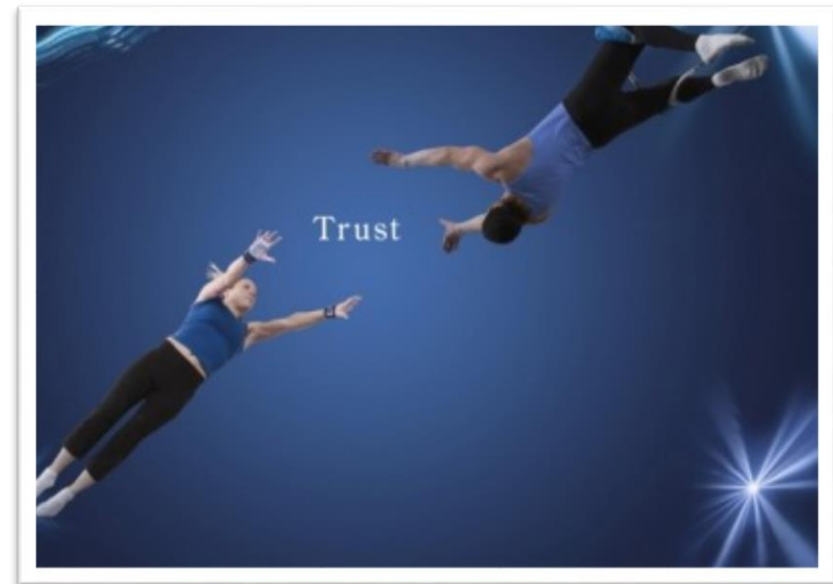


<http://www.dpworkshop.org/dpm-eng/conclusion.html>

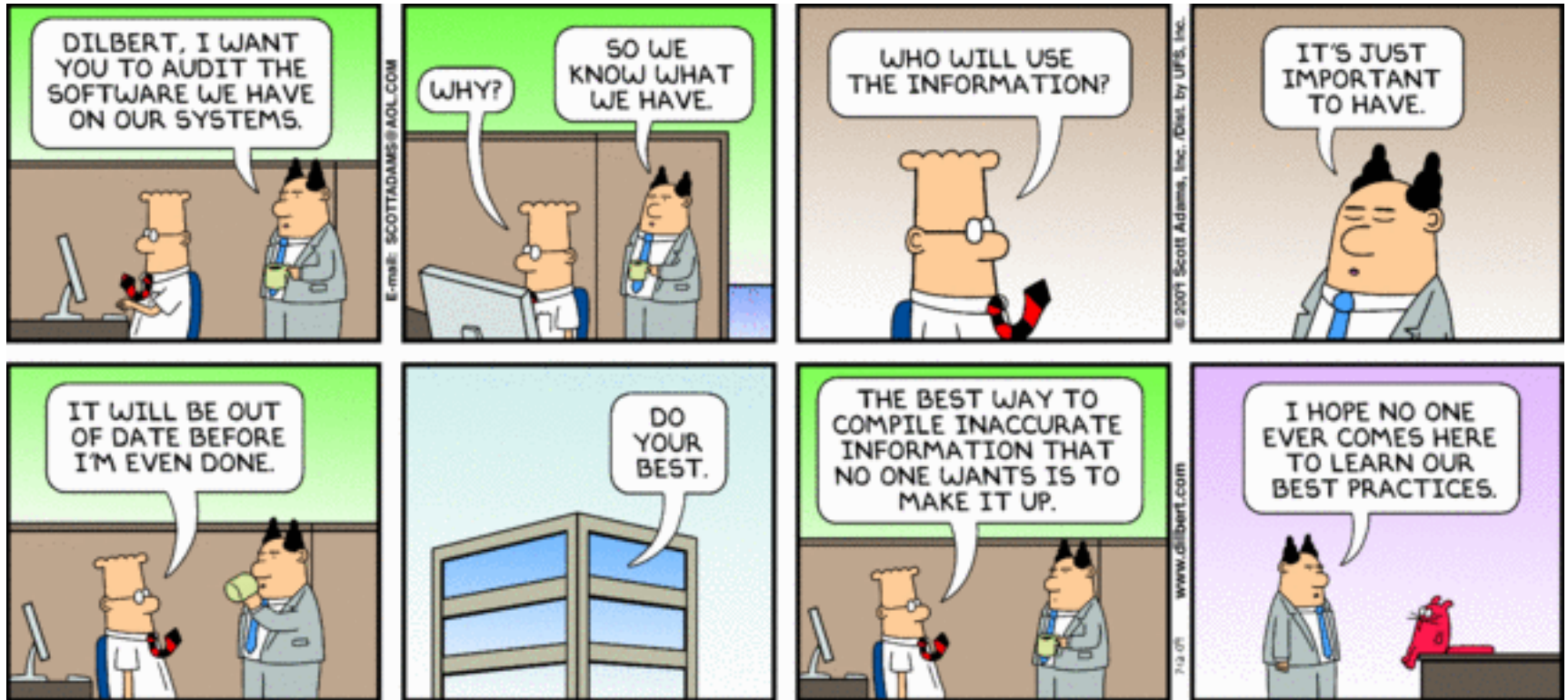
The Designated Community

TRAC A3.1 Repository has defined its designated community/communities and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation requirements will be met.

ISO 16363:2012 3.3.1 The repository shall have defined its Designated Community and associated knowledge base(s) and shall have these definitions appropriately accessible.



Third-party Certification





History of Repository Certification

1985: *Report from the Committee on Records of Government*

2002: *Digital Repositories: Attributes & Responsibilities* (Research Libraries Group & OCLC)

Recommendation: Develop a process for the certification of digital repositories

2005: Trustworthy Repositories Audit and Certification (TRAC) Criteria and Checklist (Lead by RLG-NARA Task Force)

2006: Center for Research Libraries (CRL) begins TRAC assessments

2007: TRAC presented to Consultative Committee for Space Data Systems (CCSDS) by several RLG-NARA working group members

2012: ISO 16363 Standard for Trustworthy Certification

2014: ISO 16919 for Audit requirements

2015: ISO-PTAB (Primary Trustworthy Digital Repository Authorization Body) hosts High Level Training Courses

Jan 20, 2016: ANAB (ANSI-ASQ National Accreditation Board) Announces Certification Process for Auditing Bodies

Reference: Marks, Steve, Bruce I. Ambacher, Christopher J. Prom, and Michael Shallcross.
Becoming a Trusted Digital Repository. Society of American Archivists, 2015. Print



Why Pursue Repository Certification?

- “Funders of repositories and those who entrust their valuable digitally encoded information to them need to know whether their funds and their faith is well founded. Stakeholders need to know if the repository is worthy of trust” – ISO-PTAB <http://www.iso16363.org/>
- Repository managers benefit:
 - Consolidation and management of documentation
 - Objectively convey best practices
 - Convey value and services of the digital repository to stakeholders
 - Convey credibility and boost visibility of the repository



Current Digital Repository Efforts

- Self-assessment
 - Example: University of North Texas TRAC Conformance Document
<http://www.library.unt.edu/digital-libraries/trusted-digital-repository>
- CoreTrustSeal Certification: Research Data Alliance (RDA) and ICSU World Data System (ICSU-WDS) tiered certification approach
- Data Seal of Approval: 100+ repositories have received seal
- ISO 16363:2012
 - (Jan 2018) The National Cultural Audio Visual Archives (NCAA) hosted by the Indira Gandhi National Centre for the Arts is awarded first ISO 16363 Certification
- Federal government
 - USGS Eros repository currently holds Data Seal of Approval
 - CENDI Digital Repository Working Group shares standards and initiatives



Questions?

Email: jtieman@gpo.gov

Significant recognition and thanks to the following repository managers who have provided reflections about experience with repository certification:

William Wueppelman, Canadiana.org
Steve Marks (formerly with), Scholars Portal
Amy Kirchhoff, Portico
Vicky Reich, CLOCKSS
David Minor & Sibyl Shaefer, Chronopolis
Mike Furlough, HathiTrust
Jeremy York, University of Colorado (Formerly with HathiTrust)
John Faundeen, USGS EROS Center



Trustworthy Digital Repositories: CoreTrustSeal Requirements for Self-Assessment

Mary Moulton, Digital Librarian 0000-0002-1791-068X
US DOT National Transportation Library

FDLP Academy
Washington, D.C.
July 10, 2018

About NTL

Established in 1998, we provide access to:

- Digital collections
- Data services
- Reference and research services
- Networking

We are an **open access** digital repository.

All items are in the **public domain** and available for reuse **without restriction.**

NTL Mandates

- **Transportation Equity Act for the 21st Century (1998)**
 - “establish and maintain a National Transportation Library, which shall contain a collection of statistical and other information needed for transportation decision making at the Federal, State, and local levels.”
- **MAP-21 (2012)**
 - 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 Office of Science and Technology Policy memo (2013)** requiring all Executive Departments and Agencies spending more than \$100 million/year on R&D to ensure public access to peer-reviewed publications and digital datasets arising from federally-funded scientific research

National Transportation Library

[Home](#) [Collections](#) [Recent Additions](#) [Public Access](#) [Submit Content](#)

[About ROSA P](#) ▾




All Collections ▾

Enter keyword or phrase...


Search


[Advanced Search](#)

BTS Products

 Pocket Guide to Transportation 2018

 2017 TSAR

 USDOT Intelligent Transportation Systems

 NHTSA - Behavioral Safety Research

BTS Products

USDOT Bureau of Transportation Statistics collection of statistical reports and data products

[View this collection >>](#)

[View Archive](#)


Put a ROSA P badge on your website

National Transportation Library

Recently Added

 States' Successful Practices Weigh-in

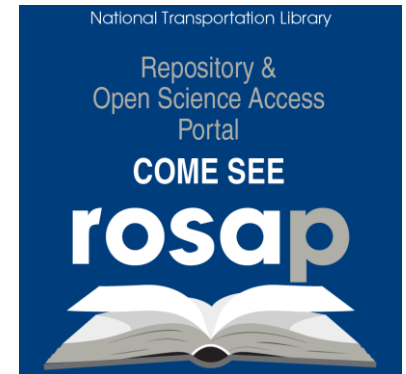
Most Popular

 Updating the school bus standard vehicle for load

NTL Digital Collections

35,180 total digital objects across 10 collections:

- BTS Products
- Federal Highway Administration
- NHTSA - Behavioral Safety Research
- US Transportation Collection
- Volpe Technical Reference Center
- Federal Motor Carrier Safety Administration
- Investigations of Aircraft Accidents (1934-1965)
- Federal Transit Administration - 50th Anniversary Document Collection
- Intelligent Transportation Systems Joint Program Office Digital Library
- BTS Data Directory



U.S. Department of Transportation

Office of the Assistant Secretary for
Research and Technology

Trustworthy Digital Repositories

- Digital repositories are not websites; require different analytics
- We needed a method to measure and communicate the value of NTL's digital repository to stakeholders
- Trustworthiness is a valid qualitative assessment
- Digital repositories should either be certified or support a repository standard



Benefits of Self-Assessment and Peer Review

- Demonstrate that we are following good practices by providing documentary evidence
- Gain independent insights on how to further increase trustworthiness
- Create a benchmark for comparison by identifying the strengths and weaknesses of a repository
- Build user, depositor and funder confidence in the repository
- Create a solid foundation for future certification



CoreTrustSeal Requirements

- 16 Core requirements, based on ISO 16363
 - Context
 - Organization infrastructure
 - Digital object management
 - Technology
 - Guidance and feedback
- Procedures
 - Self-assessment and public evidence
 - Peer review (2 reviewers)
 - Periodic review (every 3 years)



CoreTrustSeal Requirements

Context Repository type, designated community, Level of curation performed , Outsource partners, **Impact**

Organizational infrastructure

- R1. DR has **an explicit mission** to provide access to and preserve data in its domain.
- R2. DR maintains all applicable **licenses** covering data access & use & monitors compliance
- R3. DR has a **continuity plan** to ensure ongoing access to and preservation of its holdings.
- R4. DR ensures that data are created, curated, accessed, and used in compliance with **disciplinary and ethical norms**.
- R5. DR has **adequate funding** and sufficient numbers of **qualified staff** managed through a clear system of governance
- R6. repository adopts mechanism(s) to secure ongoing **expert guidance** and feedback (either in-house, or external, scientific guidance).

Digital object management

- R7. DR guarantees the **integrity and authenticity** of the data.
- R8. DR accepts data & metadata based on **defined criteria to ensure relevance and understandability** for users.
- R9. DR applies **documented processes and procedures** in managing archival storage of the data.
- R10. DR assumes responsibility for **long-term preservation** and manages this function in a planned and documented way.
- R11. DR has appropriate expertise to address **technical data and metadata quality** sufficient to make quality evaluations.
- R12. Archiving takes place according to **defined workflows** from ingest to dissemination.
- R13. DR enables users to **discover the data** and **refer to them in a persistent way** through proper citation.
- R14. DR enables reuse of the data over time, ensuring that **appropriate metadata** support the understanding and use of the data.

Technical infrastructure

- R15. DR functions on **well-supported operating systems and other core infrastructural software** and is using hardware and software technologies appropriate to the services it provides to its Designated Community.
- R16. The technical infrastructure of the repository provides for **protection** of the facility and its data, products, services, and users.



Pre-Assessment of *ROSA P*

Understanding CoreTrustSeal Requirements

- Established working groups (CENDI WG, NTL/NOAA/CDC User Group)
- Reviewed CoreTrustSeal documentation
- Updated trustworthiness guidelines on NTL website (2017-08)

Collecting Evidence

- Sources consulted:
 - NTL website
 - NTL staff
 - *ROSA P*
 - Internal documents
 - Federal legislation

Analyzing Evidence

- Created pre-assessment scorecard
- Tagged evidence with content keywords and requirements
- Identified gaps in evidence



Pre-Assessment of ROSA P

Policy Documents/Evidence for DSA Self-Assessment									
Document Name	Requirements Fulfilled	Document Location	Document Content	Document Owner	Status (Not started, Up to date, Needs)	Notes from EZ	Public or Internal?	Date Last Reviewed (Date, Initial)	Ready for Audit? (Y/N, Initials, Date)
About the National Transportation Library	R0; R1; R6	https://ntl.bts.gov/about_ntl.html	mission statement networking				Public		
NTL_Collection_Development_and_Maintenance_Policy_2016_v00	R7; R8	http://our.dot.gov/office/rita.ntl/Shared%20Documents/Collection%20Development/2016_Digital_Repository_Collection_Development/NTL_Collection_Development_and_Maintenance_Policy_%202016_v00.docx	mission statement collection development		Needs updating	Headquarters (HQ) Library Collection Criteria; still relevant? More language about checks/reviews of submissions? Clearer about what we do with non-preferred formats? Double check preferred formats?	Public		
Executive Summary – USDOT Public Access Plan	R1	https://ntl.bts.gov/publicaccess/executive_summary.html	public access government mandates			Mostly same information as DOT Public Access	Public		
DOT Public Access	R7	https://ntl.bts.gov/publicaccess/	public access government mandates submission guidelines			Mostly same information as Executive Summary	Public		

Total documents: 25



U.S. Department of Transportation

Office of the Assistant Secretary for
Research and Technology

Outcomes

- Updated NTL policies:
 - Collection Development & Maintenance
 - Metadata Policy for Digital Content
 - Metadata Inventory
 - Digital Curation
 - US DOT Public Access
- Comprehensive review of workflow and creation of Standard Operating Procedures
- Application of persistent identifiers: DOI's, ORCiD
- Forthcoming fixity checksum (CDC)
- Forthcoming documentation for technical infrastructure (CDC)

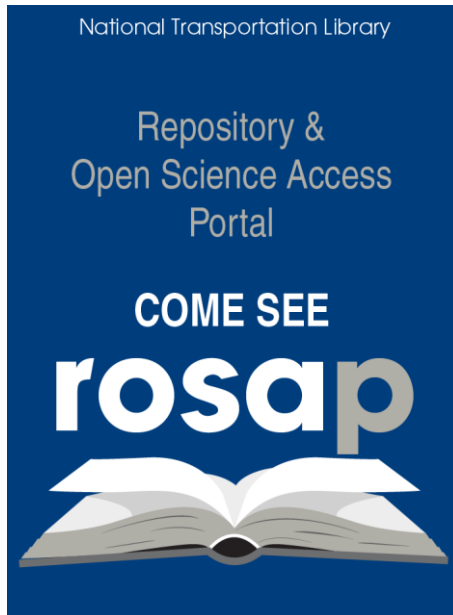


Reference List

- National Transportation Library <https://ntl.bts.gov/>
- ROSA P <https://rosap.ntl.bts.gov/>
- CoreTrustSeal <https://www.coretrustseal.org/>
- NITRD Big Data Workshop:
 - “Measuring the Impact of Digital Repositories”
<https://www.nitrd.gov/nitrdgroups/index.php?title=DigitalRepositories>
- World Data System <https://www.icsu-wds.org/>
- Data Seal of Approval <https://datasealofapproval.org/en/>
- Research Data Alliance <https://www.rd-alliance.org/>



Questions or comments?



- **Mary Moulton**
mary.moulton@dot.gov
202-366-0303
- **Ask-a-Librarian**
Answers@dot.gov
1-800-853-1351

