June 2019

Special TLR:
*U.S. DOT Public Access and Data Management Review*
Leighton Christiansen & Charles Ducker

[https://transportation.libguides.com/TLR](https://transportation.libguides.com/TLR)
U.S. DOT Public Access and Data Management Review

Leighton L Christiansen https://orcid.org/0000-0002-0543-4268
Data Curator, National Transportation Library
leighton.christiansen@dot.gov

Charles Ducker, Jr. https://orcid.org/0000-0002-7535-5258
Sr. Intellectual Property Counsel, Office of General Counsel
charles.ducker@dot.gov

Persistent link to this webinar:
https://doi.org/10.21949/1503909
Topics to Cover

• Opening U.S. Government-Funded Research Data
• U.S. DOT Public Access Review
• Submitting Final Reports and Final Datasets
• Benefits of Data Management
• Writing Data Management Plans
• Implementing Data Management Plans
Opening U.S. Government-Funded Research Data
Opening U.S. Government-Funded Research Data
U.S. DOT Public Access Review
U.S. DOT Public Access Plan Guidance Website

http://ntl.bts.gov/publicaccess/
U.S. DOT Public Access Policy

Plan to Increase Public Access to the Results of Federally-Funded Scientific Research (2015-12-15)


- Affirm and enhance DOT’s commitment to Public Access to Scientific Research results, including digitally formatted scientific data without charge to the maximum extent possible.
- Support governance of and best practices for managing Public Access to peer-reviewed Publications and Digital Data Sets across DOT.
- Ensure continuous access to and reliable preservation of DOT-funded Publications and Digital Data Sets for research, development and education purposes, within available resources.
- Preserve and increase the use of Scientific Research results to enhance scientific discovery and deployment of research results.
- Enhance the use of Scientific Research results to promote innovation and economic competitiveness.
- Affirm DOT’s support for the reproducibility of Scientific Research results.
- Make DOT’s research portfolio available to the public at the project level.
U.S. DOT Public Access Policy from 30,000 feet

- DOT-Funded Scientific Research
  - Not goods, services, or construction
- Three Components
  - Publications
  - Digital Datasets
  - Research Project Record
- Links to Final Digital Datasets
- Data Management Plan (DMP)
  - Explain case for long-term preservation and sharing; OR,
  - Justify case against log-term preservation and sharing
- Include preservation costs in proposals
Repository Characteristics Review

1. Explicit mission of digital data archiving
2. Protect privacy rights and maintain the confidentiality of research subjects
3. Enables the users to discover and use the data, and refer to them with a persistent identifier
4. Ensures the integrity and authenticity of the data

https://ntl.bts.gov/public-access/guidelines-evaluating-repositories
Submitting Final Reports and Final Datasets
Submitting Final Reports and Final Datasets

https://ntl.bts.gov/public-access/how-comply

10. Send one email to Research Hub, NTL, and TRB:
   • Final Report URL(s) or PDFs;
   • URL(s) to final datasets and descriptive metadata;
   • Funding agreement number;
   • The Research Hub Display ID;
   • ORCIDs (unique researcher IDs) for all author(s); and,
   • Any other documented project outputs or outcomes.

**Dataset Differences**

- Final datasets used to draw conclusions in report, NOT all raw data gathered;
- URLs to datasets, NOT electronic copies
- Dataset & documentation
Documenting Data

Data Package Guidelines

- **Dataset**
  - .csv or other open format
- **Readme.txt**
  - Includes Data Dictionary
  - Notes standards used
  - Defining Zero, Null, & Unknown
  - FAQs and other notes
- **Metadata file** in Project Open Data .json and/or other schema
- **Data Management Plan (DMP)**
- **Code or scripts** used in data analysis
- **Supporting files, tables, etc.**

(Bold = Required; Italic = Optional, or Required if Applicable)

A “Data Package” is the dataset, the DMP, and all other documentation needed to contextualize the dataset for any and all users.

American Travel Survey (ATS) 1995

[https://doi.org/10.21949/1503648](https://doi.org/10.21949/1503648)
Benefits of Data Management
Data Management Definitions

“In the context of research and scholarship, ‘Data Management’ refers to the storage, access and preservation of data produced from a given investigation. Data management practices cover the entire lifecycle of the data, from planning the investigation to conducting it, and from backing up data as it is created and used; to long term preservation of data deliverables after the research investigation has concluded.”
- University Library, Texas A&M  http://guides.library.tamu.edu/DataManagement

“Data management is the compilation of many small practices that make your data easier to find, easier to understand, less likely to be lost, and more likely to be usable during a project or ten years later.”
- Kristin Briney. 2015. Data management for researchers: organize, maintain and share your data for research success.
Benefits of Managing Data

• Extend data’s useful life
• Plan for software and hardware
• Plan for storage size & cost
• Save time
  – Documented file storage paths
  – Documented file naming convention and version control
  – Documented data roles
  – Data access levels

• Find data for follow-on research
• Backup and disaster recovery plan
  – 3-2-1 Strategy
• Increase funder confidence
• Improve data sharing culture

Writing Data Management Plans
DOT DMP Resources

A data management plan (DMP) describes how researchers will handle digital data both during and after a research project. DMPs will describe how the research proposal conforms to DOT policy on the dissemination and sharing of research results. Each plan should include a 2-3 page narrative description covering:

- The final research data to be produced in the course of the project;
- The standards to be used for data and metadata format and content;
- Policies for access and sharing the final research data, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, and other rights or requirements;
- Policies and provisions for re-use, re-distribution, and the production of derivatives; and
- Plans for archiving final research data and other research products, and for preservation of access to them.

DOT-funded research projects are expected to be conducted pursuant to the approved DMP. A DMP may evolve as the research project evolves and should be reviewed for possible revision whenever a data management procedure is changed.

DOT DMP Sections

1. Data Description
2. Standards Used
3. Access Policies
4. Re-Use, Redistribution, and Derivative Products Policies
5. Archiving and Preservation Plans

DOT DMP Section: Data Description

Section Description

• Provide a description of the data that you will be gathering in the course of your project.
• Address the nature, scope, and scale of the data that will be collected.
• Describe the characteristics of the data, their relationship to other data, and provide sufficient detail so that reviewers will understand any disclosure risks that may apply.
• Discuss value of the data over the long-term.

Helpful Prompts

1. Name the data, data collection project, or data producing program.
2. Describe the purpose of the research.
3. Describe the data that will be generated in terms of nature and scale (e.g., numerical data, image data, text sequences, video, audio, database, modeling data, source code, etc.).
4. Describe methods for creating the data (e.g., simulated; observed; experimental; software; physical collections; sensors; satellite; enforcement activities; researcher-generated databases, tables, and/or spreadsheets; instrument generated digital data output such as images and video; etc.).
5. Discuss the period of time data will be collected and frequency of update.
6. If using existing data, describe the relationship between the data you are collecting and existing data.
7. List potential users of the data.
8. Discuss the potential value of the data have over the long-term for not only your institution, but also for the public.
9. If you request permission not to make data publicly accessible, explain rationale for lack of public access.
10. Indicate the party responsible for managing the data.
11. Describe how you will check for adherence to this data management plan.

Submitted U.S. DOT Public Access Data Management Plans

https://rosap.ntl.bts.gov/collection_pa_dmp
Lessons from Submitted DMPS

• Author/Creator missing from sections
• DMPs can stand alone
  • Solution: Add organizational information; branding
• Single research organization; many DMPs; lots of duplicated information
  • Solutions: Master DMP for common responses
  • Project level DMP for project-specific info, link to Master
• DOT Narrative DMP not as detailed as needs to be for researchers
  • Solution: Robust project DMPs

https://rosap.ntl.bts.gov/view/dot/36367
https://c2smart.engineering.nyu.edu/c2smartpublications/#1534358159826-cb5e9a6e-27fc
Implementing Data Management Plans
Implementing DMPs

- DMPs are Living Documents
  - Quarterly reviews
  - Record changes
    - Changes in research plan are expected
  - Version Control
  - Submitted updated DMPs to DOT
- Plan to manage and share data from the beginning
  - Many small practices
  - Data more useful to you
  - Longer shelf life for data
  - Data more easily shared
Topics Covered

• Opening U.S. Government-Funded Research Data
• U.S. DOT Public Access Review
• Submitting Final Reports and Final Datasets
• Benefits of Data Management
• Writing Data Management Plans
• Implementing Data Management Plans

Questions or feedback? Email: public.access@dot.gov
• Briney, Kristin. 2015. Data Management for Researchers: Organize, Maintain and Share your Data for Research Success.
Thank you for watching!

To provide feedback or pose questions related to this TLR, please email: public.access@dot.gov

For this and other webinars, visit the TLR Archive at: https://rosap.ntl.bts.gov/collection_tlr