USDOT Public Access Review & ORCIDs

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

2019 National Research Advisory Committee and TRB State Representatives Meeting
2019-07-22
Topics to Cover

• Opening US Government-Funded Research Data
• US DOT Public Access Review
• Submitting Final Reports and Final Datasets
• Benefits of Data Management
• Disambiguating Researchers with ORCIDs
Opening US Government-Funded Research Data
Opening US Government-Funded Research Data
US DOT Public Access Review
U.S. DOT Public Access Plan
Guidance Website

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

• 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
• Will undergo Policy Review Fall 2019

https://doi.org/10.21949/1503909

https://ntl.bts.gov/public-access/faqs
NCHRP 20-11: A Guide to Ensure Access to the Publications and Data of Federally Funded Transportation-Related Research

Project Data

<table>
<thead>
<tr>
<th>Funds:</th>
<th>$249,960</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Agency:</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>Principal Investigator:</td>
<td>Dr. Carol Flannagan</td>
</tr>
<tr>
<td>Effective Date:</td>
<td>8/5/2016</td>
</tr>
<tr>
<td>Completion Date:</td>
<td>6/30/2018</td>
</tr>
</tbody>
</table>

**STATUS:** Complete. Publication pending.

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
  - Choose a conformant data repository https://ntl.bts.gov/public-access/data-repositories-conformant-dot-public-access-plan
- Dataset & documentation
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

Disambiguating Researchers with ORCIDs
ORCIDs

Free and Easy

https://orcid.org/

Open Researcher & Contributor ID

- ISO 27729
- Distinguishes any person from every other person
- Registered
- https URI: 16-digit numerical identifier preceded by "https://orcid.org/"
  http://orcid.org/0000-0002-0543-4268
## ORCIDs

### ORCID ID

<table>
<thead>
<tr>
<th>ORCID ID</th>
<th>First/given name</th>
<th>Last/family name</th>
<th>Other names</th>
<th>Affiliations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000-0002-1990-411X</td>
<td>Xin</td>
<td>Wang</td>
<td>Xin Wang</td>
<td></td>
</tr>
<tr>
<td>0000-0002-5695-2754</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0001-8866-5223</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0002-4463-5830</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td>Beijing Institute of Technology, Beijing University of Technology</td>
</tr>
<tr>
<td>0000-0002-3607-738X</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0001-8355-3536</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td>Northwestern University, Northwestern University, Tianjin University</td>
</tr>
<tr>
<td>0000-0001-7876-1359</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0001-5436-5688</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0001-3969-1415</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0001-5192-8093</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0002-6991-9512</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000-0002-6090-0161</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td>Zhejiang University, China University of Geosciences</td>
</tr>
<tr>
<td>0000-0002-6819-0384</td>
<td>Xin</td>
<td>Wang</td>
<td></td>
<td>University of Electronic Science and Technology of China</td>
</tr>
</tbody>
</table>

### Biography

Leighton L Christiansen joined the National Transportation Library (NTL) at the U.S. Department of Transportation in May 2016. As the new Data Curator, Leighton will be part of a team directed to make research datasets publicly accessible and for seeking better data management and curation practices within NTL and the Bureau of Transportation Statistics (BTS). Leighton is a member of the Research Data Alliance (RDA). Leighton served as co-chair of the Transportation Research Board’s Library and Information Science for Transportation (LIST) Committee (ARSA) from April 2015 to April 2018. Leighton is the past chair of the Midwest Transportation Knowledge Network.

- **Education (2)**
  - University of Illinois at Urbana-Champaign: Urbana, IL, United States
  - University of Iowa: Iowa City, IA, United States

- **Employment (2)**
  - National Transportation Library: Washington, DC, United States
  - Iowa Department of Transportation Library: Ames, Iowa, United States

- **Works (11 of 11)**
  - Delivering Data Packages for Discovery, Analysis, and Preservation
  - Transportation Research Board 97th Annual Meeting
  - conference poster

### Keywords

- data curation
- library science
- information science
- data management
- open data
- public access
- digital preservation
- data transportation

### Websites

- Transportation Data Management
- Twitter
- LinkedIn

**Email**: leighton.christiansen@dot.gov

**Email**: leighton.christiansen@gmail.com
Thank you Bobbi deMontigny!
Topics Covered

• Opening US Government-Funded Research Data
• US DOT Public Access Review
• Submitting Final Reports and Final Datasets
• Benefits of Data Management
• Disambiguating Researchers with ORCIDs

Questions or feedback? Email: NTLDataCurator@dot.gov
Resources & Works Cited


• Briney, Kristin. 2015. Data Management for Researchers: Organize, Maintain and Share your Data for Research Success.


Supplemental Slides
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
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; Italics = 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
Writing Data Management Plans
DOT DMP Resources

DOT DMP Overview

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 US 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

Connected Cities for Smart Mobility towards Accessible Resilient Transportation (C2SMART)

Master Data Management Plan

Version 2.0
(Revised April 23, 2018)

Lead Institution
New York University

Partner Institutions
City College of New York
Rutgers University
University of Texas, El Paso
University of Washington, Seattle

http://c2smart.engineering.nyu.edu/c2smartpublications/#1534358159826-cb5e9a6e-27fc31

https://rosap.ntl.bts.gov/view/dot/36367
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