Data Management Plan Instructions

U.S. Department of Transportation

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

Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The U.S. Government is not endorsing any manufacturers, products, or services cited herein and any trade name that may appear in the work has been included only because it is essential to the contents of the work.

Table of Contents

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

2

1. Introduction

This document provides instructions for creating a Data Management Plan (DMP) using a consistent format. These instructions are based off of the United States Department of Transportation's (U.S. DOT's) National Transportation Library's guidance for Creating Data Management Plans for Extramural Research, and provide additional clarification for the Intelligent Transportation Systems Joint Program Office (ITS JPO)'s expectations for DMPs. While these instructions inform applicants of DMP elements preferred by the ITS JPO, they do not serve as an approved government form or template. Project teams should use their best judgement when determining what information to include and if additional information should be incorporated into the DMP.

The DMP is a living document describing how researchers will handle digital data both during and after a research project, and how the research proposal conforms to U.S. DOT policy on the dissemination and sharing of research results. The DMP should adequately identify the data to be collected, how the data will further the goals of the U.S. DOT, how the data will be managed, how the data will be made accessible, how the data will be stored, and what data standard(s) will be used.

The post-award DMP should build on any preliminary DMP submitted during the application phase of a project and be updated regularly to capture changes to a project's data management procedures and/or data, including schema changes. It is recommended that the post-award DMP does not initially exceed the project's preliminary DMP length by more than five pages, however, project teams should refer to any solicitation or project documentation for specific page requirements. The post-award DMP should grow and evolve over the course of the project as more data is added and processes change.

Chapter 1. Project Overview

Provide summary information about the research project and its goals, as well as how the project's data helps achieve U.S. DOT's research goals.

Suggested elements for this section include, but are not limited to:

- **Project Title:** Please provide the title of the project.
- **Project Goals and Objectives:** Briefly describe the goals and objectives of this project, as well as how the data collected helps the U.S. DOT achieve its goals and objectives.
- **Project Description:** Briefly describe the research project, including project design and execution.
- **Project Lifecycle Phase:** Please provide the phase of the project lifecycle as of the date of the DMP or updated DMP. Phases include: planning, pre-award, award, post award, and closeout.

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

3

- **Project Performance Measurements:** Briefly describe any performance measurements the project will use and if appropriate, how the data collected support these measurements.

Change Control

Briefly describe the plans for modifications and updates to the DMP over time, including how those changes will be controlled. If there are planned updates, provide the estimated release dates.

Relevant Documents

List any reference documents with information relevant to data management for the project, including the name and the source of the document. If some of the information requested by this template is available in other documents, those documents may be referenced to avoid duplicating content and maintenance of redundant documentation.

Chapter 2. Data Overview

Please provide a description of the nature, scope, and scale of the data that is collected and/or produced. For each unique dataset, please provide all the elements in this section.

Suggested elements for this section include, but are not limited to:

- **ID:** Provide all U.S. DOT, state DOT or transportation-related identification numbers for this dataset. This may include, but is not limited to, ORCHID ID numbers, Transportation Research Board (TRB) RiP database numbers, contract numbers etc.
- **Dataset Title**: Provide the dataset title.
- **Description**: A) Describe the purpose of the dataset including how it relates to the broader purpose of the research project. B) If using existing data, describe the relationship between the data you are collecting and the existing data. C) Discuss the potential long-term value the dataset will have to the research project and other consumers.
- **Type/Scale**: Specify the type and scale of the data that is generated (e.g., numerical data, image data, text sequences, positional data, video data, audio data, database data, modeling data, source code, etc.).
- **Collection Method**: Describe the methods for collecting 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.).
- **Data File Format(s):** Provide the data file format(s) the datasets are anticipated to be in when they are made accessible to the U.S. DOT (e.g. .csv, .txt, json, newline json, etc.). If not yet known, explain to the best of your ability.

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

4

Chapter 3. Data Stewardship

Data Owner and Steward

Please define the data owner and data steward for this project's datasets. It is acceptable for multiple datasets to have the same data owner and/or data steward.

Suggested elements for this section include, but are not limited to:

- **Dataset Title:** Please provide the title of the data and/or datasets that are assigned to the designated data owner and/or data steward. Data and/or datasets with different data owners and/or data stewards may be listed separately.
- **Data Owner:** List the data owner of the dataset. The data owner is the person or organization with the authority, ability, and responsibility to access, create, modify, store, use, share, and protect the data. Data owners have the right to delegate these privileges and responsibilities to other parties. Please note the data owner may be the U.S. DOT Program or Project Manager for U.S. DOT-funded research projects.
- **Data Steward:** List the data steward for the dataset. The data steward, at the direction of the data owner, is the person or organization that is delegated the privileges and responsibilities to manage, control, and maintain the quality of a data asset throughout the data lifecycle. The data steward may also apply appropriate protections, restrictions, and other safeguards depending on the nature of the data, subject to the direction of the data owner. The data steward may be the recipient of U.S. DOT-funded research projects or one of the recipient's partner organizations.
- **Federal Sponsor:** Please list the U.S. DOT sponsor or project manager for this/these dataset(s).

Access Level

Providing appropriate data access to safeguard data is a key aspect of data stewardship. In accordance with the OPEN Government Data Act, datasets must be made publicly accessible unless specific concerns require the data to have controlled-access. Please detail what, if any data requires controlled-access, reasons for controlling access to data, and methods taken to safeguard.

Controlled-access is defined as restricting access to certain groups of persons due to data containing personally identifiable information (PII), confidential business information (CBI), or any other information that threatens privacy of an individual or group.[1]

Suggested elements for this section include, but are not limited to:

---

[1] https://www.its.dot.gov/data/public-access/

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

5

Can all data from this project be shared with the public or is controlled access required for at least some of the data?

Please complete this section based on what best describes the access level for the project's data:

- **All Public Access:** If there is no data requiring controlled-access, select the 'Public Access' checkbox.
- **Some/All Controlled-Access:** If there is data in one or more datasets requiring controlled-access (e.g. PII or CBI select the 'Controlled-Access' checkbox.

Datasets Requiring Controlled-Access

This section is only required if "Controlled-Access" is selected above.

- In this section, list all datasets that require controlled-access and why the datasets require controlled-access. Please provide all elements for each anticipated dataset that requires controlled-access.
- **Dataset Title:** Provide the dataset title.
- **Reasons for Controlled-Access:** For any data requiring controlled-access, please detail why the data requires controlled-access.
- **Safeguarding Methods and Processes:** *Describe the methods and processes that are taken to safeguard the data requiring controlled-access, such as redacting data before providing public access to the dataset. Please elaborate on the specific methods (e.g. blurring license plates in videos, geofencing sensor data, removing persistent IDs, etc.). Any privacy, ethical or confidentiality concerns raised due to data sharing should be described.*

Informed Consent

If this project contains human subject research, please describe how informed consent forms permit sharing with the research community and whether additional steps, such as an Institutional Review Board (IRB) are used to protect privacy and confidentiality.[2]

Access Requests

For each dataset, describe what processes this project will follow to allow full access to this data if requested, including criteria a user must pass to gain access (e.g. for a specific business purpose or if the user is a part of a certain group, like a member of the research team, etc.).

Related Tools, Software and/or Code

List any specialized tools that are needed to access, analyze or manipulate the data to support reuse or replication, and name(s) of the needed tool(s) and software. Specify how any needed tools can be

---

[2] https://ntl.bts.gov/public-access/creating-data-management-plans-extramural-research

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

6

accessed (i.e. open source and freely available, generally available for a fee in the marketplace, or available only from the research team or another source).

Relevant Privacy and/or Security Agreements

If applicable, for each dataset provide a list of relevant privacy and/or security agreements (e.g., non-disclosure agreements, rules of behavior, memoranda of understanding, confidentiality agreements, sub-contracts, etc.) that apply to these data and its users, including circumstances when these agreements would be used.

Re-Use, Redistribution, and Derivative Products Policies

In accordance with the OPEN Government Data Act, to the extent practicable, data must be made available through an open license that places with no restrictions on copying, publishing, distributing, transmitting, citing or adapting such asset.

In this section, please provide the applicable licenses for each dataset and specify if an open license will be used. If the project is U.S. DOT-funded, an open license should be used for the dataset, unless otherwise restricted by controlled access or prohibited by other intellectual property concerns. If a non-open license is used for this data, include a justification for why a non-open license is used.

Suggested elements for this section include, but are not limited to:

- **Dataset Title:** Please provide the title of the dataset(s).
- **License Used:** List the license used for this/these dataset(s) (e.g. CC0 1.0 Universal) and provide a URL linking to where the full text of the license is maintained. Specify whether the license is an open license or a non-open license. An "'open license' means a legal guarantee that a data asset is made available (A) at no cost to the public and (B) with no restrictions on copying publishing, distributing, transmitting, citing or adapting such asset.[3]" A non-open license does not adhere to all of the open license criteria.
- **Reasons for a Non-Open License:** If no license is applied or a non-open license is used for a dataset, provide the reason(s).

Data Storage and Retention

Storing and retaining the data is a key part of the data steward's responsibilities to manage, control, and maintain the quality of a data asset throughout the data lifecycle.

In this section, list all of the data storage systems used to store the project's data, providing details on those data systems, and specifying how long the data will be stored in each system.

---

[3] https://www.congress.gov/bill/115th-congress/house-bill/1770/text

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

7

A public access system provides full data access to the public. A controlled-access system restricts access to certain groups of persons due to data containing PII, CBI, or any other information that threatens the privacy of an individual or group.[4]

An applicant system is managed by the applicant, a U.S. DOT system is managed by the U.S. DOT, and a third-party system is one that is managed by a person, group or organization other than the U.S. DOT or the applicant.

Suggested elements for this section include, but are not limited to:

Storage Systems

- **Data Storage System Name:** Provide the name of the data storage system(s) the dataset(s) will be stored in. Provide URLs or other means of accessing the data storage system.
- **Data Storage System(s) Type:** For each data storage system, state whether the data storage system is:
  - "Applicant – Public System"
  - "Applicant – Controlled System"
    - If you are using an applicant-managed system, please state whether the system is currently operational, under development, or will be developed as part of this project.
  - "U.S. DOT-managed – Public System"
  - "U.S. DOT-managed – Controlled-Access System"
  - "Third-party – Public System"
  - "Third-party – Controlled-Access System"
- **Dataset Title(s):** Provide all datasets that are anticipated to be stored in each data storage system.
- **Initial Storage Date:** Provide the initial date that data will be available in each data storage system. If you are live-streaming data, this may be the initial ingest date.
- **Frequency of Update:** Provide how frequently the data will be updated in the data storage system once ingestion begins (i.e., "Continually," "Daily," "Weekly," "Monthly," "Annually," "Unknown," "As needed," "Irregular," or "None planned"). Please note the U.S. DOT prefers obtaining the data as quickly as possible after collection, preferably in near real-time and/or through automated uploads.
- **Archiving and Preservation Period:** Provide the duration for which the dataset will be maintained in each data storage system.

Data Storage System Description

Provide a description of the data storage system, including any URLs or other means that provide additional details or documentation about the system, details on the system's data backup, and details on access or controlled-access. If using an applicant or third-party data storage system, please provide additional details demonstrating that the system adheres to the U.S. DOT's policies for data storage systems, including:

- Guidelines for Evaluating Repositories for Conformance with the DOT Public Access Plan
- For controlled-access systems, describe how the system enables privacy protection, controlled-access and collaboration equal to or exceeding the U.S. DOT's controlled-access system, the U.S. DOT Secure Data Commons

---

[4] https://www.its.dot.gov/data/

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

8

Cybersecurity Policies

State any cybersecurity policies that apply to the data storage system or the entirety of this DMP.

Data Security Policies and Procedures

Provide the data security policies and procedures this system uses, as well as any actions the awardee has or will take to augment them.

Back-up and Recovery Policies and Procedures

Provide the back-up and recovery policies and procedures this data storage system uses, as well as any actions the awardee has or will take to augment them.

Chapter 4. Data Standards

Data Standards

Please provide details on the data standard(s) used for each dataset and provide access to the data's metadata and data dictionary.

Suggested elements for this section include, but are not limited to:

- **Dataset Title:** Provide the dataset title.
- **Data Standard(s):** List the name(s) of the data standard(s) in which the data are made available to the U.S. DOT.
- **Data Standard(s) Digital Object Identifier(s) (DOI[s]):** List the DOI(s) of the standard(s) for the data. If possible, provide a URL to the data standard(s).
- **Open or Proprietary?:** Indicate whether the data standard(s) is/are "Open" or "Proprietary." It is important to note that data submitted to the ITS JPO should be platform-independent and non-proprietary to ensure accessibility.
- **Data Standard(s) Rationale:** Provide an explanation for using the chosen data standard(s). If you are using a proprietary or non-standard data standard, discuss your rationale for using that standard here. If you are using a non-proprietary data standard, state that here and indicate why the standard applies to this dataset.

Versioning

Briefly outline the procedures for ensuring that different versions of the dataset(s) are identifiable, properly controlled and properly used. Include any standards used for version control in the outline.

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

9

Metadata and Data Dictionary

The ITS JPO requires a metadata schema with each dataset submission to enable search and discovery. Review the Metadata Questionnaire and provide access and metadata summary information in this section. Note that the U.S. DOT expects metadata in several contexts, including:

- At the field level, describing the elements in the data and/or dataset
- At the asset level, describing the contents of the data or dataset generally, in accordance with the Project Open Data Metadata Schema or another appropriate standard, or
- At any other level that may be appropriate for the data and/or dataset,

Suggested elements for this section include, but are not limited to:

- **Dataset Title:** Provide the dataset title.
- **Metadata Standards Used:** List the metadata standards used for the dataset. Provide any URLs or documentation for these metadata standards.
- **Metadata Discoverable:** Indicate whether metadata can be made discoverable by the ITS JPO.
- **Data Dictionary Discoverable:** Indicate whether a data dictionary for the dataset can be made discoverable by the ITS JPO.
- **Metadata and Data Dictionary Access:** Provide any URLs that provide access to the dataset's metadata and data dictionary.

Metadata Description

Describe how metadata will be provided for each dataset, as well as the different contexts metadata will be provided in (e.g. for individual datasets, a group of datasets, etc.).

Chapter 5. Glossary of Terms

U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation Systems Joint Program Office

10

U.S. Department of Transportation