

[Insert Project Title]

Data Management Plan Instructions

Draft Report — XX 2019

Publication Number: FHWA-JPO-XX-XXX



Data Management Plan Instructions

Produced by **name of agency**
U.S. Department of Transportation
Intelligent Transportation Systems (ITS) Joint Program Office

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.

Data Management Plan Instructions

Technical Report Documentation Page

1. Report No. FHWA-JPO-XX-XXX		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle				5. Report Date XX 2019	
				6. Performing Organization Code	
7. Author(s)				8. Performing Organization Report No.	
9. Performing Organization Name and Address				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address ITS-Joint Program Office 1200 New Jersey Avenue, S.E., Washington, DC 20590				13. Type of Report and Period Covered Draft Report	
				14. Sponsoring Agency Code HOIT-1	
15. Supplementary Notes Work performed for: Insert name of Program Manager, ITS JPO					
16. Abstract Please add abstract here					
17. Key Words			18. Distribution Statement		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages XX	
				22. Price	

Form DOT F 1700.7 (8-72)

Reproduction of completed page authorized

Contents

- Introduction 5**
- 1. Project Overview 5**
 - 1.1 Change Control5
 - 1.2 Relevant Documents5
- 2. Data Overview 5**
- 3. Data Stewardship 6**
 - 3.1 Data Owner and Steward6
 - 3.2 Access Level.....6
 - 3.2.1 Can all of the data from this project be shared with the public, or is controlled-access required for at least some of the data?7
 - 3.2.2 Datasets Requiring Controlled-Access7
 - 3.2.3 Informed Consent7
 - 3.2.4 Access Requests7
 - 3.2.5 Relevant Privacy and/or Security Agreements7
 - 3.3 Re-Use, Redistribution, and Derivative Products Policies.....8
 - 3.4 Data Storage and Retention8
 - 3.4.1 Storage Systems8
 - 3.4.2 Data Storage System Description.....9
 - 3.4.3 Cybersecurity Policies9
 - 3.4.4 Data Security Policies and Procedures.....9
 - 3.4.5 Back-up and Recovery Policies and Procedures.....9
- 4. Data Standards 9**
 - 4.1 Collection Format Standards10
 - 4.2 Versioning10
 - 4.3 Metadata and Data Dictionary10
- 5. Glossary 10**

Data Management Plan Instructions

Introduction

This document provides guidelines for creating a Data Management Plan (DMP) to make it easier for projects to prepare their 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 U.S. DOT's expectations for DMPs. These instructions inform applicants of elements the U.S. DOT would like to see in a DMP, but do not serve as an approved government form or template. Applicants should use their best judgement in determining what information to include and if additional information should be incorporated into the DMP.

The Data Management Plan (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 made accessible, how the data will be stored, and in what standard(s) the data will use. .

A preliminary DMP with Project Overview, Data Overview, and Data Stewardship section details must be created for the proposal phase of a research project. It is important to note that projects applying and receiving funding from ITS JPO are required to submit DMPs when applying for funding (a preliminary DMP) and after award (a post-award DMP). The preliminary DMP will evolve into the post-award DMP, and should be updated regularly to capture pertinent changes to a project's data management procedure and/or data, including schema changes.

1. Project Overview

In this section 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 in this project 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 in project lifecycle as of the date of the DMP or revision thereof. Phases include: planning, pre-award, award, post award, and closeout.

1.1 Change Control

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

1.2 Relevant Documents

List 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 called for in this template is available in other documents, those documents may be referenced to avoid duplication of content and maintenance of redundant documentation.

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.

Data Management Plan Instructions

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

- **ID:** Provide all 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 existing data. C) Discuss the potential value the dataset will have over the long-term 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, video, audio, database, 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.

3. Data Stewardship

Please provide details around data stewardship. Data stewardship involves proper management of data throughout the data lifecycle. This includes, but is not limited to, maintaining data quality and safeguarding the data.

3.1 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 data and/or datasets that are assigned the designated data owner and/or data steward. You may break out data and/or datasets that have separate data owners and/or data stewards.
- **Data Owner:** List who the data owner of the dataset is. The data owner is the person or organization that has 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 for U.S. DOT-funded research projects.
- **Data Steward:** List who the data steward for the dataset is. The data steward is, at the direction of the data owner, 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).

3.2 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

Data Management Plan Instructions

have controlled-access. Please detail which data require controlled-access, reasons for controlling access to data, and methods taken to safeguard these data.

Controlled-access is defined as restricting access to certain groups of persons due to data containing personally identifiable information (PII), information that threatens privacy of an individual or group, information that threatens confidentiality of a person or group, and/or contains confidential business information (CBI).¹

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

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

- **Public Access:** If there is no data requiring controlled-access, select the 'Public Access' checkbox.
- **Controlled-Access:** If there is data in one or more datasets requiring controlled-access (e.g. Personally Identifiable Information [PII] or Confidential Business Information [CBI]) select the 'Controlled-Access' checkbox.

3.2.2 Datasets Requiring Controlled-Access

This section is only required if "Controlled-Access" is selected for Section 4.2.1.

- 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, for example 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.

3.2.3 Informed Consent

If your 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

3.2.4 Access Requests

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

3.2.5 Relevant Privacy and/or Security Agreements

¹ <https://www.its.dot.gov/data/public-access/>

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

Data Management Plan Instructions

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.

3.3 Re-Use, Redistribution, and Derivative Products Policies

The U.S. DOT is interested in who holds the intellectual property right(s) for the research project data, whether these rights will transfer to a data archive, whether any copyrights apply to the data, and other aspects that apply to the rights to use the data. Please provide details on these elements in this section.

In accordance with the [OPEN Government Data Act](#), to the extent practicable, data must be made available through an open license that is available at no cost to the public and 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. 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) that use this license.
- **License Used:** List the license that is used for this/these dataset(s) (e.g. Creative Commons – BY version X.0) and provide a URL (or several) with more information on the license. 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 a non-open license is used for a dataset, provide the reason for using a non-open license.

3.4 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 that are used to store the research project's data, provide details on those data systems, and specify how long the data will be stored in each system.

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 personally identifiable information (PII), information that threatens privacy of an individual or group, information that threatens confidentiality of a person or group, and/or contains confidential business information (CBI).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 a system 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:

3.4.1 Storage Systems

³ <https://www.congress.gov/bill/115th-congress/house-bill/1770/text>

⁴ <https://www.its.dot.gov/data/>

Data Management Plan Instructions

- **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:** Please 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," "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. This is in accordance with the OPEN Government Data Act's requirement to maximize accessibility to federally-funded datasets.
- **Archiving and Preservation Period:** Provide the duration for which the dataset will be maintained in each data storage system.

3.4.2 Data Storage System Description

Provide a description of the data storage system, including any URLs or other means that provide additional details and documentation of the system, details on data backup of the system, and details on access and controlled-access. If using an applicant or third-party data storage system, please provide additional details that demonstrate 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 at least equal to U.S. DOT's controlled-access system, the [U.S. DOT Secure Data Commons](#)

3.4.3 Cybersecurity Policies

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

3.4.4 Data Security Policies and Procedures

Provide the data security policies and processes this system has, as well as any actions the awardee takes to augment these processes.

3.4.5 Back-up and Recovery Policies and Procedures

Provide the back-up and recovery policies and processes this data storage system has, as well as any actions the awardee takes to augment these processes.

4. Data Standards

Data Management Plan Instructions

4.1 Collection Format Standards

Please describe the data standards used, provide access to the data's metadata and data dictionary, and provide any requirements for reading the data.

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

- **Dataset Title:** Provide the dataset title.
- **Data Collection Format Standard(s):** List the name(s) of the format(s) in which the data is be collected, submitted, and/or preserved (note that these formats may be the same).
- **Data Collection Format Standard Uniform Research Identifier(s) (URIs):** List the URIs of the data collection format standard(s) for the data.
- **Open or Proprietary?:** Indicate whether the data collection format standard is "Open" or "Proprietary." It is important to note that data submitted to ITS JPO should be platform-independent and non-proprietary to ensure that they will be accessible.
- **Collection Format Rationale:** Provide an explanation for using the chosen data collection format. If you are using proprietary or non-standard data formats, discuss your rationale for using those standards and formats here. If you are using standard data formats, also state that here and why the standard applies to this dataset.
- **Access Tools:** List any special tools required to read or view this data and/or its data standards, including tools needed for all types of workstations (e.g. personal computer (PC), Mac, Chromebook, etc.).

4.2 Versioning

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

4.3 Metadata and Data Dictionary

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
- 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 using the same title provided in the Data Description section.
- **Metadata Standards Used:** List the metadata standards used for the dataset. Provide any URLs or documentation on 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 this dataset, including the different contexts metadata will be provided in.

5. Glossary

Data Management Plan Instructions

This section should include a glossary of terms used.



Data Management Plan Instructions

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free "Help Line" 866-367-7487
www.its.dot.gov

FHWA-JPO-XX-XXX