Plan Overview

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D1ND5W

Title: Transit Reliability Improvement and Performance System (TRIPS)

Creator: David Kobayashi - **ORCID:** 0009-0000-9232-5977

Affiliation: United States Department of Transportation (DOT) (transportation.gov)

Contributor: Bao Nguyen

Funder: United States Department of Transportation (DOT) (transportation.gov)

Funding opportunity number: EAZVJZNJBS22

Grant: SMARTFY22N1P1G11

Template: SMART Grants Stage 1 Data Management Plan (DMP)

Project abstract:

The Transit Reliability Improvement and Performance Systems project is a multi-agency centralized transit signal priority (TSP) system to provide seamless TSP service across multiple local agencies (up to 11 agencies) operating traffic signal controllers from up to four different vendors in Santa Clara County. The project plans to feed real-time transit vehicle location data from the transit operator's computer aided dispatch/automatic vehicle location system to a centralized server and the centralized server determines optimal time to transmit commands for TSP service to the traffic signal controllers in the field via the Internet. The transmission of commands would utilize standard communication protocols such as National Transportation Communication for Intelligent Transportation Systems Protocol (NTCIP) 1202 and NTCIP 1211.

The first phase of the Transit Reliability Improvement and Performance Systems (TRIPS) project provides a framework for Santa Clara Valley Transportation Authority’s (VTA) centralized transit signal priority (TSP) solution to improve speed and reliability of transit operations throughout Santa Clara County, California, with a focus on VTA’s frequent service routes. The project’s Data Management Plan (DMP) will provide the framework for the data that will be collected for the project and how it will be maintained. This DMP is a living document that will be continuously reviewed and updated throughout the project’s data lifecycle. Additionally, the project manager will submit a compliance statement. The compliance statement will include any deviations from the DMP framework set forth in subsequent sections

Start date: 09-15-2023

End date: 07-15-2025

Last modified: 01-20-2026

Transit Reliability Improvement and Performance System (TRIPS)

Dataset and Contact Information

Please provide as much of the the following information as possible:

1. **Name of the project;**
2. **Grant number;**
3. **Name of the person submitting this DMP;**
4. **ORCiD of the person submitting this DMP (need an ORCID? Register here: https://orcid.org/);**
5. **Email and phone number of the person submitting this DMP;**
6. **Name of the organization for which the person submitting this DMP is working;**
7. **Email and phone number for the organization;**
8. **Link to organization or project website, if applicable; and,**
9. **Date the DMP was written.**

Transit Reliability Improvement and Performance System (TRIPS)

SMARTFY22N1P1G11

David Kobayashi

N.A.

David.Kobayashi@vta.org 408-321-5892

Santa Clara Valley Transportation Authority (VTA)

David.Kobayashi@vta.org 408-321-5892

www.vta.org

12/15/2023

Data Description

Please provide as much information as possible:

1. **Provide a description of the data that you will be gathering in the course of your project or data from a third party that you will re-use, if any;**
    1. **If there will be no data collected or re-used from another source, state that this is case;**
2. **Address the expected nature, scope, and scale of the data that will be collected, as best as you can at this stage;**
3. **As best as you can, 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;**
    1. **If data might be sensitive, please describe how you will protect privacy and security, if you know that now;**
    2. **You may need to update your DMP later to add more detail;**
4. **Discuss the expected value of the data over the long-term.**

TRIPS will conduct data collection spanning a broad array of topics related to VTA's transit signal priority system. Third party data will not be collected nor be reused from another source at this time, but this might change over the progression of the project. Historical data will be provided by VTA when, and if necessary. The data collected will be focused on both technical and non-technical components as follows:

Systems Engineering Management Plan: this data will help establish the fundamental technical requirements needed to operate a Transit Signal Priority System that is able to effectively operate not only across multiple jurisdictions but also various signal operators each with their own technical requirements. The SEMP will include the following subcomponents:

- Concept of Operations (CONOPS) plan
    - Summary of Request for Information
    - Listing in national/region Intelligent Transportation Systems (ITS) architecture

    - Identification of key stakeholders
    - An inventory of traffic signal controller cabinet/equipment, including communication networks

- Definition of system requirements
    - User preferences and requirements
    - Development of risk management plan

- Detailed Design
    - Detailed Specifications for Procurement
    - Testing and Verification Plans

Public Outreach: the data collected will help to provide an equitable and inclusive public engagement process to inform current and future prioritization of where and how TSP should be deployed. Qualitative and quantitative data including the following will be collected and used in the evaluation:

- Community Based Organization (CBO) working with disadvantaged communities to gather data on transit needs
- Surveys of community that address potential trade offs that might occur with improved transit speed and reliability
- GIS based evaluations that produce quantitative data based on census information
- Needs based prioritization of future deployments

Multiple unknown or to be determined data sources could fall under the overarching topics of 'System Engineering Management Plan' and 'Public Outreach'. A detailed data description will be provided in the DMP upon completion of the data collection. Until then, a general description of all the collected data is provided below. Please note that these general descriptions will apply to hard data, SEMP, and not all soft data, Public Outreach. General Data Description will include but not be limited to the following:

1. Name of the data, data collection methodologies, and data producing program.
2. Description of the purpose of the data collection.
3. Description of the methods for creating the data (e.g., simulated; observed; experimental; software; physical collections; sensors; satellite; enforcement activities; stakeholder-generated databases, tables, and/or spreadsheets; instrument generated digital data output such as images and video; etc.).
4. The period of existing time data will be collected and frequency of update.
5. The description of the relationship between the data being collected and existing data.
6. List potential audiences and uses of the data.
7. The potential value of the data over the long-term for the home institution and for the public and/or research community in general.
8. Rationale for lack of public access, if appropriate, and name of person with primary responsibility for data stewardship.
9. Description of how the VTA will check for adherence to this data management plan.
10. Description of how the data will be protected from accidental or malicious modification or deletion when stored on VTA TRIPS SharePoint (see section 5). It is the stakeholders' responsibility to maintain the data until it is uploaded to VTA TRIPS SharePoint (see section 5).

Some of the data will provide long term benefits for the operations of TSP throughout Santa Clara County. The technical requirements established in the CONOPS for operating a countywide TSP system with various local agencies will establish the basis for how this system will be able to function in the future. In addition, the outreach data that is collected will help to further establish an equitable transit system to serve all communities within Santa Clara County.

Data Format and Metadata Standards Employed

Please provide as much information as you can:

1. **Describe the anticipated file formats of your data and related files;**
2. **To the maximum extent practicable, your DMP should address how you will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future;**
    1. **If you are unable to use platform-independent and non-proprietary formats, you should specify the standards and formats that will be used and the rationale for using those standards and formats.**
3. **Identify the metadata standards you will use to describe the data.**
    1. **At least one metadata file should be a DCAT-US v1.1 (https://resources.data.gov/resources/dcat-us/) .JSON file, the federal standard for data search and discovery.**

The project shall adhere to the following formatting standards:

1. Use the standard file formats for files, utilizing non-proprietary formats such as .csv for excel files, .txt for Word files, and .pdf whenever possible.

2. If using proprietary data formats, software, and applicable versions must be provided for a user to open and view the data or documentation files.
3. Describe the data process log to clarify the final version of data shared to the public.
4. Describe how alternative formats will be documented and why they are being.
5. List what documentation will be needed to make the data understandable by others.
6. In general, the data generated by the project would be an open accessible format, excluding those data sets as described under access policies section, and would follow the guidance as provided on the data.org website such as the Data Category - United States (DCAT-US) standardized format.

Access Policies

In general, data from DOT-funded projects must be made publicly accessible. Exceptions to this policy are: data that contain personally identifiable information (PII) that cannot be anonymized; confidential business information; or classified information. Protecting research participants and guarding against the disclosure of identities and/or confidential business information is an essential norm in scientific research. Your DMP should address these issues and outline the efforts you will take to provide informed consent statements to participants, the steps you will take the protect privacy and confidentiality prior to archiving your data, and any additional concerns. In general, in matters of human subject research, your DMP should describe how your informed consent forms will permit sharing with the research community and whether additional steps, such as an Institutional Review Board (IRB), may be used to protect privacy and confidentiality. Additionally, when working with, or conducting research that includes Indigenous populations or Tribal communities, researcher will adhere to the CARE Principles for Indigenous Data Governance https://www.gida-global.org/care and make an explicit statement to that effect in this portion of the DMP.

Please provide as much information as possible:

1. **Describe any sensitive data that may be collected or used;**
2. **Describe how you will protect PII or other sensitive data, including IRB review, application of CARE Principles guidelines, or other ethical norms and practices;**
    1. **If you will not be able to deidentify the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, you should describe the necessary restrictions on access and use;**
3. **Describe any access restrictions that may apply to your data;**
4. **If necessary, describe any division of responsibilities for stewarding and protecting the data among Principal Investigators or other project staff.**

TRIPS does not anticipate to collect sensitive information that includes personally identifiable information (PII) for the project. However, it is expected there could be some technical data that will be used related to traffic signal operations and equipment that may be considered sensitive. The project will work with all stakeholders to ensure the technical information used for the project will not include any PII, sensitive traffic signal data, or other information that is proprietary. Any data that is collected and includes personal information will be deidentified to protect privacy and confidentiality of individuals.

The project requires strict access restrictions, and only team members will have access to the technical data as needed.

The project will address issues and outline the efforts and steps they will take to protect privacy and confidentiality prior to archiving data, and any additional concerns (e.g., embargo periods for the data). If necessary, the project will describe any division of responsibilities for stewarding and protecting the data among other project staff. The privacy and confidentiality of these datasets would need to comply with applicable data policies from VTA and the other stakeholders. Some examples of these data policies are as follows but not limited to these:

https://www.vta.org/privacy-policy#:~:text=Data%20Security,access%2C%20disclosure%2C%20or%20use

https://www.sanjoseca.gov/your-government/departments-offices/information-technology/digital-privacy

https://saecommon.sccgov.org/countypolicy/Information-Technology-Security-Policies.pdf

https://www.santaclaraca.gov/our-city/government/governance/privacy-website-policies#:~:text=In%20the%20future%2C%20the%20City,to%20state%20or%20federal%20law.

https://www.sunnyvale.ca.gov/your-government/codes-and-policies/website-policies/privacy-and-security

If the data will not be able to deidentified in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, the project will describe the necessary restrictions on access and use of this data.

The project will be required to adhere to the following throughout the project:

- Describe what data will be shared, how data files will be shared, and how others will access them.
- Describe what, if any, privacy, ethical, or confidentiality concerns are raised due to data sharing.
- If applicable, describe how data will be deidentified before sharing. If not:
    - Identify what restrictions on access and use will be placed on the data.
    - Discuss additional steps, if any, will be used to protect privacy and confidentiality.

Re-use, Redistribution, and Derivatives Products Policies

Recipients are reminded:

1. **Data, as a collection of facts, cannot be copyrighted under US copyright law;**
2. **Projects carried out under a US DOT SMART Grants is federally funded; therefore, as stated in grant language:**
    1. **Recipients must comply with the US DOT Public Access Plan, meaning, among other requirements, project data must be shared with the public, either by the researchers or by US DOT;**
    2. **That by accepting US DOT funding through this grant, recipients have granted to US DOT a comprehensive non-exclusive, paid-up, royalty-free copyright license for all project outputs (publications, datasets, software, code, etc.). This includes all rights under copyright, including, but not limited to the rights to copy, distribute, prepare derivative works, and the right to display and/or perform a work in public; and,**
    3. **In accordance with Chapter 18 of Title 35 of the United States Code, also known as the Bayh-Dole Act, where grant recipients elect to retain title to any invention developed under this grant, US DOT retains a statutory nonexclusive, nontransferrable, irrevocable, paid-up license to practice or have practiced for or on behalf of the United States any such invention throughout the world.**

Please provide as much information as possible:

1. **Describe who will hold the intellectual property rights for the data created or used during the project;**
2. **Describe whether you will transfer those rights to a data archive, if appropriate;**
3. **Identify whether any licenses apply to the data;**
    1. **If you will be enforcing terms of use or a requirement for data citation through a license, indicate as much in your DMP;**
4. **Describe any other legal requirements that might need to be addressed.**

For this federally funded project, VTA and its stakeholders shall retain ownership of data created by the project and will comply with the US DOT Public Access Plan. The data set will be accessed using Microsoft SharePoint platform during the life of the project.

The project does not anticipate transferring rights to a data archive based on the preliminary information. However, if the data becomes unmanageable VTA may have to transfer these rights to a data archive. The project may elect to store the data at an offsite facility on a physical media (e.g., digital tape backup) or any other alternative media, if needed.

Archiving and Preservation Plan

Please provide as much information as possible:

1. **State where you intend to archive your data and why you have chosen that particular option;**
2. **Provide a link to the repository;**
3. **You must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability;**
    1. **Whatever archive option you choose, that archive should support the capture and provision of the US Federal Government DCAT-US Metadata Schema https://resources.data.gov/resources/dcat-us/**
4. **In addition, the archive you choose should support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and must provide for maintenance of those identifiers throughout the preservation lifecycle of the data;**
5. **Your plan should address how your archiving and preservation choices meet these requirements.**

At the completion of the first phase of the project, the project will transition to the close out phase and take the following actions:

Store Project data from Stage 1 on the VTA SharePoint platform. Upon project completion, all data, except those containing sensitive information, will be deposited into the VTA Open Data Portal (https://data.vta.org/). All data will be archived with

required metadata to ensure its discoverability. This will include essential details adhering to the DCAT-US Metadata Schema (https://resources.data.gov/resources/dcat-us/).

The team will deposit the final dataset into Zenodo. The repository Zenodo provide persistent identifiers to their published data and supports the capture and provision of the DCAT-US Metadata Schema.

Store the data for 10 years while adhering to VTA standard practices and policy of preserving the data for records retention. After 10 years, the data will be moved to VTA SharePoint archive folder or archived to an offsite facility on a physical media (e.g., digital tape backup).

Provide copies of the System Engineer Management Plan (SEMP) and public outreach documents to the USDOT staff that summarizes the data collected in this first phase of the project.

This approach is proposed for this project to be consistent with other VTA project file storage and archiving processes. The files will include identifiers such as file name, date, and author.

Planned Research Outputs

Text - "TRIPS FINAL IMPLEMENTATION REPORT_Revised_1.1.pdf"

The Final Implementation Report illustrates how the proof of concept or prototype performed relative to expectations and evaluates whether a full-scale implementation of the CTSP would meet the SMART Program’s goals. The report includes the System Engineering Management Plan that includes the Draft Verification Plan, Draft Memorandum of Understanding, Proof-of-Concept Route 57 Report, Public Engagement Reprot, and the Centralized Transit Signal Priority Technical Deployment Evaluation.

Text - "VTA TRIPS FINAL_ConOps_ver1.3.pdf"

The Concept of Operation Report describes how the centralized transit signal priority system will function from an operational perspective.

Dataset - "route57_april24-25_weekday.csv"

This CSV file contains stop-to-stop travel time data for VTA’s Fast and Frequent Route 57 for April 2024 and 2025. The data was pulled directly from VTA’s CAD/AVL (Computer-Aided Dispatch/Automatic Vehicle Location) system. The data package includes "README_route57_april24-25_weekday.txt", as well as a general "README.txt" that provides background information on the project and associated reports.

Text - "README_route57_april24-25_weekday.txt"

This file provides context for the route57_april_weekday.csv file and defines the data elements necessary for proper interpretation.

Text - "README"

This file provides contextual information regarding the Transit Reliability Improvement and Performance (TRIPS) Stage 1 Final Data Set.

Text - "2026-01-20_vta_trips_metadata.json"

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
TRIPS FINAL IMPLEMENTATION REPORT_Revised_1.1.pdf	Text	2026-01-19	Open	Zenodo		Creative Commons Attribution 4.0 International	DCAT-US	No	No
VTA TRIPS FINAL_ConOps_ver1.3.pdf	Text	2026-01-19	Open	Zenodo		Creative Commons Attribution 4.0 International	DCAT-US	No	No
route57_april24-25_weekday.csv	Dataset	2026-01-19	Open	Zenodo		Creative Commons Attribution 4.0 International	DCAT-US	No	No
README_route57_april24-25_weekday.txt	Text	2026-01-19	Open	Zenodo		Creative Commons Attribution 4.0 International	DCAT-US	No	No
README	Text	2026-01-19	Open	None specified		Creative Commons Attribution 4.0 International	None specified	No	No
2026-01-20_vta_trips_metadata.json	Text	2026-01-19	Open	None specified		Creative Commons Attribution 4.0 International	None specified	No	No