

Data Management Plan

Creating a SMART Data Management Plan

This section has been added to this plan and formatted accordingly to assist in the plan submission to the Office of the Secretary of Transportation (OST). This section has been filled out per SMART Grant Recipient DMP Guidance provided by OST.

What project are you planning?	TTD ATMS Data Aggregation Plan
Select the primary organization	United States Department of Transportation
Select the primary funding organization	United States Department of Transportation
Which DMP template would you like to use?	SMART Grants Stage 1

Part 1 of 6: Project Details tab

This section has been added to this plan and formatted accordingly to assist in the plan submission to the Office of the Secretary of Transportation (OST). This section has been filled out per SMART Grant Recipient DMP Guidance provided by OST.

Project Title	TTD ATMS Data Aggregation Plan
Project abstract	<p>The Tahoe Transportation District (TTD) and regional partners are seeking to enhance transportation safety and mobility along key corridors in the Lake Tahoe Basin. The roadway network entering, traveling within, and leaving the Tahoe Basin lacks the infrastructure required to acquire real-time and historical traffic and congestion data. This pilot project aims to support TTD and the regional partners in collecting accurate count data at each of the seven entry/exit points of the Tahoe Basin and along the Truckee/US80/SR267/SR89 roadways.</p> <p>The project is designed to plan, prototype, test, and evaluate a limited deployment of a data collection sensor infrastructure to gather transportation and traveler-related information. The goal is to integrate this information into a single cloud-based open source or interface for reporting and management. This information will be utilized by TTD, TRPA, partners, commuters, and travelers within the Tahoe Basin and adjoining areas to provide an integrated infrastructure for collecting vehicle data. This data will be incorporated into a database for various stakeholders.</p>

TTD ATMS DATA AGGREGATION PLAN

	Furthermore, the project will establish the framework for long-term data collection across the region and integrate multiple transportation data sources for efficient use by partner agencies. It will propel the region toward real-time parking availability for motorists and improve the ease of transit use, walking, and bicycling. Other long-term uses include sharing information about weather hazards, closures, construction, or crashes.
Research domain: You may select one of the research domains from the list or leave this field blank.	Smart Sensors-based infrastructure and cloud-based AI video analytics.
Project Start	09/15/2023
Project End	3 10/31 15 /2025
Funder	United States Department of Transportation
Funding Status	Funded
Funding opportunity number	DOT-SMART-FY23-01
Grant number/url	SMARTFY22N1P1G41

Part 2 of 6: Collaborators tab

This section has been added to this plan and formatted accordingly to assist in the plan submission to the Office of the Secretary of Transportation (OST). This section has been filled out per SMART Grant Recipient DMP Guidance provided by OST.

Add a Contributor

Contributors are the lead administrator(s) on your project and/or those responsible for data management. When you press “add a contributor,” you will be asked to provide the name, email, and role of contributors on your project. Note: ORCID is not required for SMART.

Name	Email	Role
Jim Marino	jmarino@tahoetransportation.org	Project Manager
Tara Styer	tstyer@tahoetransportation.org	Project Manager

Invite Collaborators

Collaborators are specific people who can read, edit, or administer your plan. If you invite a collaborator, they will receive an email notification that they have access to this plan. To invite a collaborator, insert their email address and select which permissions to grant them.

Name	Email	Role
Josh Schmid <u>Rachael Shaw</u>	<u>rshaw</u> jschmid @trpa.org	Data Management <u>MPO Planning</u>
Kira Richardson <u>Mason Bindl</u>	<u>ksmith</u> mbindl @trpa.org	MPO Planning <u>Data Management</u>

TTD ATMS DATA AGGREGATION PLAN

Part 3 of 6: Write Plan tab

This section has been added to this plan and formatted accordingly to assist in the plan submission to the Office of the Secretary of Transportation (OST). This section has been filled out per SMART Grant Recipient DMP Guidance provided by OST.

Each section of the DMPTool template can be expanded, where award recipients will see a prompt to provide more information on a specific topic. Each section of the DMP will provide a detailed description of what should be included. The DMP template includes the following topics:

1. Dataset and Contact Information

Name of the project	TTD ATMS data aggregation plan
Grant Number	SMARTFY22N1P1G41
Name of the person submitting this DMP	Jim Marino, Deputy District Manager
OPTIONAL: ORCID of the person submitting this DMP	
Email and phone number of the person submitting this DMP	jmarino@tahoetransportation.org 775-588-5500x512
Name of the organization for which the person submitting this DMP is working	Tahoe Transportation District
Email and phone number for the organization	jallen@tahoetransportation.org
Link to organization or project website, if applicable	Link to project website will be provided in report update.
Date the DMP was written	December 12, 2023

2. Data Description

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;

The data collection process will be comprehensive and involve a range of critical metrics encompassing multimodal traffic and safety data. The data will include items such as, but not be limited to, traffic signal status, traffic-volume data, O-D data, traveler profile data, traffic delays, local weather information, road closures due to snow, avalanches, landslides, flooding, road construction events, street events, vehicular or pedestrian accidents, transit movement and delays, ride-sharing locations and availability, micro-transit locations and availability, and parking locations and availability.

There are additional data sources from other agencies, such as vehicle counts and crash data from NDOT and Caltrans, in addition to other third-party sources like location-based data or GPS (connected vehicle) data. TTD, Placer County, RTC of Washoe County, and other transit providers can contribute transit ridership and vanpool use data. Bicycle and pedestrian count information will be integrated as well.

The collected data will support providing a shared understanding of extensive visitation and address peak demand, including road user travel volumes and patterns. It aims to evaluate road user safety, particularly for vulnerable road users such as bicyclists and pedestrians, and to support Vehicle Miles Traveled (VMT) and Greenhouse Gas (GHG) goals. Additionally, it aims to

TTD ATMS DATA AGGREGATION PLAN

optimize available parking and transit use, aid in communication with the traveling public, ultimately aiming to protect Lake Tahoe's famed clarity.

Address the expected nature, scope, and scale of the data that will be collected, as best as you can at this stage;

The data collection effort aims to capture a comprehensive view of multimodal interactions within the transportation system. The scale will likely involve real-time and historical data (as available) across the entry and exit points of Lake Tahoe to gauge the effectiveness of real-time multimodal traffic data collected by smart sensors. This will be facilitated through a single cloud-based open source or interface for pertinent transportation and traveler-related information. This approach intends to offer a broad scope of insights into the functionality and safety enhancements across different locations, traffic scenarios, and conditions.

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;

Some example data characteristics that will be collected are below.

- Multimodal corridor volume: Data indicating different modes of traffic flow at specific locations or times.
- Traffic count: The number of modes that pass a certain location.
- Accident data: Data concerning crashes (vehicular or vulnerable road users).
- Transit information

The data collected will be interconnected, forming a comprehensive map of multi-modal traffic interactions and safety incidents tracked over the course of the pilot period. There are inherent disclosure risks, especially concerning individual privacy and identifiable patterns. Personal details linked to incidents could pose privacy concerns, necessitating anonymization techniques to mitigate such risks before any dissemination or analysis.

The proposed project will not collect any personal identifiable information such as license plates and facial recognition.

If data might be sensitive, please describe how you will protect privacy and security, if you know that now;

To safeguard sensitive data and ensure privacy and security the following best practices will be utilized: anonymization techniques, control data access, secure storage, and data record minimization. Additional mitigations will be added later as we solidify the privacy and security data strategy for this project.

The proposed project will not collect any personal identifiable information such as license plates and facial recognition.

You may need to update your DMP later to add more detail;

Our Data Management Plan will be updated later to add more detail.

Discuss the expected value of the data over the long-term.

The collected data will be archived and stored for future analysis and insights by a many potential users such as traffic engineers, safety engineers, road operators, etc. The data collected will drive continual improvements, inform policy and decision-making, provide predictive insights to help mitigate safety risks or traffic congestions based on historical trends, foster innovation, potentially reduce commute times, provide information transparency for users, and ultimately enhance transportation safety and mobility along key corridors in the Lake Tahoe Basin.

3. Data Format and Metadata Standards Employed

Describe the anticipated file formats of your data and related files;

Our intent is to use industry-standard data collection and storage formats to allow for interoperability with other applications. The anticipated file formats for the collected data and related files will likely be a combination of various formats, including but not limited to:

- CSV (Comma-Separated Values): Suitable for structured data such as collision incidents, violations, or pedestrian volumes, enabling easy manipulation and analysis using various software tools.
- JSON (JavaScript Object Notation): Useful for storing hierarchical data or metadata related to V2X technologies, traffic signal systems, or video analytics configurations.
- XML (extensible Markup Language): A potential format for encoding diverse data structures, especially useful for describing video analytics parameters or system configurations.
- MP4 or AVI (Audio Video Interleave): Video file formats for storing surveillance or traffic camera footage, essential for video analytics and understanding traffic patterns.
- Database Files (SQL, SQLite, etc.): Organized databases might be utilized to store and manage complex relational data, facilitating efficient querying and retrieval of information.
- GeoJSON or Shapefiles: Geospatial data formats used to represent geographic features, essential for mapping traffic incidents, pedestrian volumes, or collision locations.
- PDF (Portable Document Format): A potential format for documentation, reports, or manuals related to the systems or technologies under study.

The choice of file formats will depend on the nature of the data, its structure, and the intended analyses or applications.

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;

Utilizing platform-independent and non-proprietary formats ensures the accessibility and usability of data across different systems and software, enhancing its long-term utility. The project will adhere to industry standards for data collection and storage formats, prioritizing those widely supported and platform-independent, accessible through various tools and programming languages such as CSV or JSON. Avoiding vendor-specific formats, the project will mandate documentation for metadata in all datasets.

TTD ATMS DATA AGGREGATION PLAN

Additionally, we plan to ensure that the solutions employed incorporate a clear and simple API (application programming interface), facilitating seamless integration with other applications and solutions.

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.

N/A

Identify the metadata standards you will use to describe the data. 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 will use industry standard metadata and corresponding data schemas. This will use standard .JSON file systems or similar. For example, ASTM E2468-05/E2665-08/2259-03a establishes the metadata elements that will promote the quality and usefulness of transportation data archives.

4. Access Policies

Describe any sensitive data that may be collected or used;

Describe how you will protect PII or other sensitive data, including IRB review, application of CARE Principles guidelines, or other ethical norms and practices; 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;

Describe any access restrictions that may apply to your data;

This project will not collect or store personally identifiable information. No license plate numbers or facial recognition will be collected. Also, video clips will be deleted from the system after the traffic and safety analysis is completed.

Data access will also be restricted. Data will only be accessible by personnel linked to the project from TTD and appropriate consultants and vendors on the project team on permission basis.

If necessary, describe any division of responsibilities for stewarding and protecting the data among Principal Investigators or other project staff.

During the period of performance for Phase 1, the responsibility for stewarding and protecting the data will be established per user agreements with consultant and vendor. Data sharing agreements will be formally established between partner agencies.

5. Re-use, Redistribution, and Derivatives Products Policies

Describe who will hold the intellectual property rights for the data created or used during the project;

Data collected during the project will likely not be subject to IP rights. Data that is anticipated to be collected throughout the pilot period will be subject to public disclosure.... Data will be anonymized.

Describe whether you will transfer those rights to a data archive, if appropriate;

Identify whether any licenses apply to the data;

TTD ATMS DATA AGGREGATION PLAN

Collected data will not have any licenses restrictions. Specific vendor analytical tools (NOT DATA) that provide unique mining capabilities and insights will mostly likely be subject to IP terms and conditions

If you will be enforcing terms of use or a requirement for data citation through a license, indicate as much in your DMP;

Not anticipating enforcing any data terms through a license. This will be updated during project executions

Describe any other legal requirements that might need to be addressed.

N/A

6. Archiving and Preservation Plan

State where you intend to archive your data and why you have chosen that particular option;

The project team is in the process of scoping this. We will edit and modify this plan as we move further into the project.

Provide a link to the repository;

*You must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability; 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/>*

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;

Your plan should address how your archiving and preservation choices meet these requirements.

Once the project team completes its plan for any data archiving activities, we will ensure compliance with the US Federal government DCAT-US Metadata Schema.
