Multimodal Projects Discretionary Grant (MPDG)- Mega

Appendix F: Draft MEGA Data Plan

Date: 5/6/2024

INNOVATE 680

IMAGINE THE POSSIBILITIES

CONTRA COSTA transportation authority

Table of Contents

INTRODUCTION

Purpose of this document

The purpose of this document is to detail the Data Collection Plan (DCP) specific to the NB680 Express Lane Completion (Phase 1), Shared Mobility Hubs (SMH), and Coordinated Adaptive Ramp Metering (CARM) projects, as required for the Multimodal Project Discretionary Grant (MPDG) Mega funding opportunity. This plan, tailored to the Contra Costa Transportation Authority's (CCTA) strategic initiatives under the *Innovate 680* program, adheres to the following:

1. MPDG NOFO's Mega Data Collection Requirements;
2. Title 49 U.S.C 6701(g) requirements[1]; and
3. USDOT National Transportation Library's Data Management Plan (DMP) guidelines for extramural research activities[2].

This document describes how CCTA will handle data during and after the period of performance and provides a description of data that will be collected during the course of the project as part of CCTA's *Innovate 680* program. The DCP describes the plan for data gathering prior to construction to develop a baseline, perform data analysis to identify the status of existing conditions, identify Key Performance Indicators (KPIs) and impacts of project implementation, monitor the accuracy of forecasted KPI measures through the use of data analysis, simulation, and periodic updates to data to verify accuracy of KPI measures, and performing before and after-study reports to identify actual outcomes. The DCP includes standards that will be used for data collection and access restrictions for the gathered data. A description of how data is stored and archived is detailed in this document.

Approach and Work Steps

CCTA's proposed approach is to perform a before and after study; and in between, perform periodic monitoring of the status of predicted project characteristics and forecasted KPI measures with the actual outcomes to ensure the projects are on the right track. In addition, CCTA will monitor the project to achieve its high-level goals and determine how it meets the Mega project's high-level requirements, such as its ability to improve economy, mobility, and safety while being cost-effective. Section 1.4 of Chapter 1 of this report identifies performance indicators related to project goals and maps them to criteria defined in the NOFO. Chapter 2 of this document describes the data types, standards, metadata, and data sources that will be used for gathering data before, during, and after the implementation of the project. Chapter 3 describes the process of data gathering and data collection. Chapter 4 describes the proposed approach for assembling and analyzing data, including baseline, KPIs, predicted project characteristics, forecasted KPIs, simulation, monitoring, and generating reports. Chapter 5 describes the access policies for access restriction or sharing. Chapter 6 describes the data storage location, policies, requirements, and tools.

[1] https://uscode.house.gov/view.xhtml;jsessionid=13F7049BF27B2AA627BAEA90117C795C?req=49&f=treesort&fq=true&num=3833&hl=true&edition=prelim&granuleId=USC-prelim-title49-section11145

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

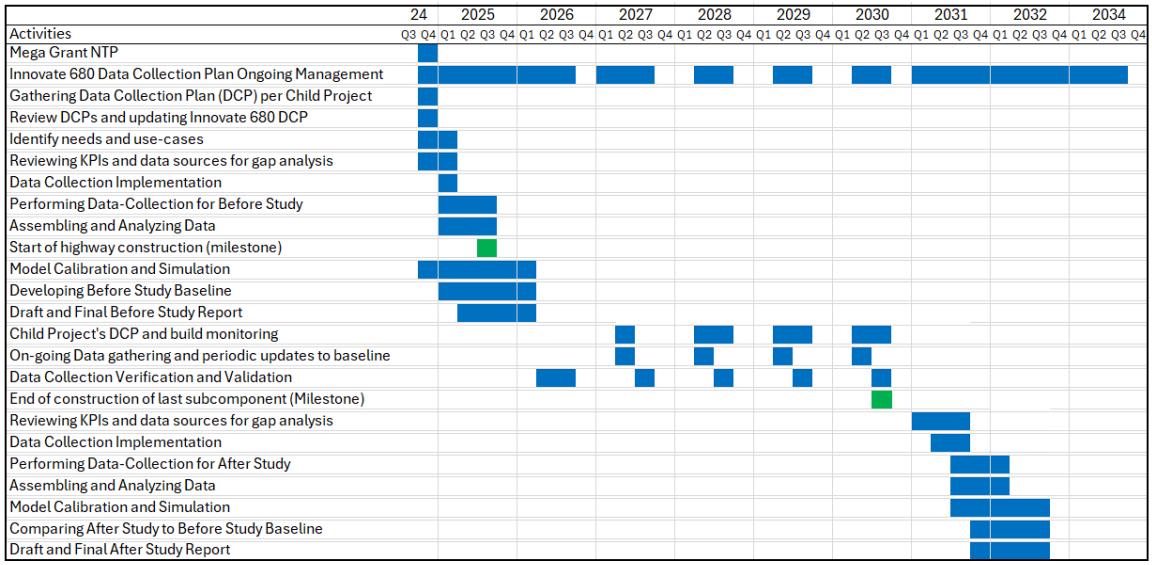
CCTA's approach includes the following work steps:

1. Update and finalize this plan within 60 days after the award and/or release of USDOT's Mega Grant Standard framework.
2. Identify KPIs, data needs, types, formats, standards, sources, and gaps.
3. Identify data collection and data gathering steps needed.
4. Conduct Data Collection prior to implementation and/or construction of the project.
5. Perform data analysis and assemble data to develop a baseline.
6. Identify predicted project characteristics and forecast methodology for KPIs.
7. Perform simulation and data analysis to verify the accuracy of the forecasted KPIs.
8. Prepare Before Study report.
9. Perform periodic data gathering, simulation, forecasting, and monitoring accuracy of KPIs.
10. Perform monitoring of possible alternatives analysis if needed.
11. Identify steps needed to update data gathered within five years after project completion.
12. Perform data gathering within five years after completion of the project.
13. Perform data analysis and assemble data to compare with the baseline.
14. Prepare an After Study report.

As part of the work steps listed above, CCTA will perform and deliver the following reports:

15. Before Study (project baseline) report: before the start of construction of a project, CCTA will submit a Before Study report providing baseline data for the purpose of analyzing the long-term impact of the project in accordance with the Mega Grant Standard framework established per Section 1.3 of this document.
16. After Study (Updated project baseline) report:
    - Not later than six years after the date of substantial completion of a project, the eligible entity carrying out the project shall submit a project outcomes report that compares the baseline data to quarterly project data for the duration of the fifth year of the project after substantial completion.

The high-level schedule of the main data collection activities is presented below; this schedule will be further expanded as described in Section 3.2 of this document. Please note that due to the varying schedules for delivery of the various project components, data collection milestones will vary:



Mega Grant Standard Framework

CCTA anticipates that the U.S. Department of Transportation (USDOT) will issue a comprehensive framework outlining the performance measures and Key Performance Indicators (KPIs) associated with this grant. This framework is expected to standardize the units of measurement for assessing the objectives and targets integral to USDOT's internal evaluations. Once these updated requirements are published, we will promptly revise our Data Collection Plan (DCP) to conform to the new standards. This update will ensure that the data collection methodologies, forecasting techniques, and data storage protocols fully align with the Mega Grant Standard framework and meet all relevant USDOT specifications. Our commitment is to seamlessly integrate these requirements to enhance the accuracy and effectiveness of our project evaluations.

Key Performance Indicators (KPIs) Mapping to Mega Grant Criteria

MPDG Mega Grant NOFO requires DCP to map KPIs to Mega Grant Project Outcome Criteria identified in Section E of the NOFO to identify the key performance measures that need to be verified in a Before and After study for accuracy. The Grant does identify the following key outcome criteria:

1. Safety
2. State of Good Repair
3. Economic impacts, Freight Movement, and Job Creation
4. Climate Change, Resiliency, and the Environment
5. Equity, Multimodal Option, and Quality of Life
6. Innovation

The following table summarizes project component to their expected benefits and KPIs:

Project Components	Expected Benefits.	KPIs
Shared Mobility Hubs (SMH)	Improving mobility and safety, GHG reduction, improving travel time reliability, mode shift, affordable and accessible transportation	Percent accident reduction, percent increased transit usage, travel time savings, percent GHG reductions, percent change in VHD, transit passengers counts
NB 680 Express Lane Completion (ELC) (Phase 1)	Improving mobility and safety, GHG reduction for Vehicles and Trucks, improving Travel time reliability, Accident and Collision prevention, improving System Operation, Economic Benefits and Freight Movement	Percent accident reduction, percent increased transit usage, travel time savings, percent GHG reductions, percent change in VHD, percent increase in freight volume, express lanes performance
Coordinated Adaptive Ramp Metering (CARM)	Improving mobility and safety, GHG reduction for Vehicles and Trucks, improving Travel time reliability, Accident and Collision prevention, improving Operation and Maintenance, Improving State of Good Repair for Ramp Meters sub-systems, Improving the freeway's throughput for commute and freight movement, improving ability to recover after a major accident and benefits to improved infrastructure resilience, reduced operating costs	Percent accident reduction, percent increased transit usage, travel time savings, percent GHG reductions, percent change in VHD

DATA DESCRIPTION AND REQUIREMENTS

As Section 1.4 of this document mentions, the *Innovate 680* program has multiple project components; some are traffic-related, and some are transit-related. For each project component, CCTA will provide a description of the detailed data that will be collected and analyzed, including the nature and scale of data, the methods that will be used to assemble and analyze data, and the methods that will be used to verify accuracy of the predicted outcome and forecasted KPIs. CCTA will gather that information into a central repository that will be used to generate the before and after study reports required for the Mega Grant, as identified in Section 1.2 of this document. The information included in this chapter, such as the type of data, data sources, formats, and methods, is the preliminary information that will be updated as the project progresses and during the finalization of this document.

Types of Data

CCTA will carry out comprehensive before and after studies as part of the Mega grant application, focusing on strategic improvements across the following projects: SMH, NB680 Express Lane Completion (Phase 1), and CARM. These studies will assess enhancements in the following areas:

- **Safety**: Through the reduction of accidents, collisions, and related injuries or fatalities. This includes detailed monitoring of incidents both on the NB680 express lanes and the areas influenced by adaptive ramp metering systems.
- **Mobility**: By measuring improvements in travel time reliability, traffic volume and density, and vehicle hours of delay (VHD), assessing how these projects enhance the fluidity of traffic and reduce congestion.
- **Economy**: Evaluating cost savings from decreased need for road closures, reduced maintenance requirements, and more efficient operation of transportation systems.
- **Environmental Impact and Health**: Quantifying the reduction in greenhouse gas emissions and enhancements in air quality, particularly through the increased use of shared mobility hubs that promote public transit and other non-vehicular modes. Additionally, examining the environmental benefits of improved stormwater management systems associated with these transportation projects.
- **State of Good Repair**: Enhancing the durability and resilience of infrastructure through better maintenance practices and new technologies that promote long-term sustainability.
- **Technology and Connectivity**: Implementing advanced systems that enable future connectivity and automation, thus improving the integration between different transportation modes and enhancing the user experience for pedestrians, bicyclists, and public transit users, including those on intercity and commuter rails.
- **Multimodal Integration**: Increasing the connectivity between various modes of transportation to facilitate smoother transitions for users moving goods and services, which is vital for economic and operational efficiency.

These studies will utilize the specific data types collected from each project, such as accident and collision data, transit ridership figures, traffic flow metrics, environmental monitoring data, and operational performance indicators, to comprehensively evaluate the impact of the CCTA's initiatives. By

focusing on these key areas, CCTA aims to demonstrate the substantial benefits these projects bring to the region, supporting the overall goals of the Mega grant to improve safety, mobility, and environmental sustainability across transportation networks.

The data that will be collected will vary by project components and their expected outcomes. The following is a preliminary list of data that will be collected:

1. SMH
   - Accident and collision data.
   - Transit ridership data.
   - Travel time reliability and savings.
   - GHG reduction for Vehicles, Transit, and Trucks.
   - Traffic Volume and Density, Vehicles Hours of Delay (VHD).
2. NB680 Express Lane Completion (ELC) (Phase 1)
   - Accident and collision data.
   - Transit ridership data.
   - Travel time reliability and savings.
   - GHG reduction for Vehicles, Transit, and Trucks.
   - Vehicles Hours of Delay (VHD).
   - Freight volume.
   - Toll tags and Express Lane usage.
   - Vehicle occupancy and high-occupancy vehicles (HOV) count and volumes.
3. CARM
   - Traffic Volume and Density, Vehicle Miles Travel (VMT).
   - GHG reduction for Vehicles and Trucks.
   - Travel time reliability.
   - Vehicles Hours of Delay (VHD).
   - Accident and collision data.
   - data related to collisions on the mainline and ramps, average number of vehicles per collision, and average number of victims per injury.
   - Operation and Maintenance cost.
   - Data related to Improving the State of Good Repair for Ramp Meters sub-systems.
   - Data related to Improving the freeway's throughput for person and freight movement.
   - Data related to the ability to recover after a major accident and benefits to improved infrastructure resilience.

In addition to specific data related to project components, there will be other data that will be collected and captured, such as weather data, video, audio, subjective survey responses, and information from media and news for specific events and/or major or minor incidents that may have short term non-recurrent impact to the before and after study results. The data types will be further expanded during the data collection process and procedures related to assembling and analyzing data, as described in chapters 4 and 5 of this document.

Data Sources and Available Data

CCTA, in collaboration with various regional and local entities, including California Department of Transportation (Caltrans) District 4, the Metropolitan Transportation Commission (MTC), and various transit agencies such as Bay Area Rapid Transit (BART), Contra Costa County Transit Authority (County Connection) and Livermore Amador Valley Transit Authority (LAVTA), will utilize a comprehensive array of data sources for the projects. These sources encompass:

- **Central Systems:** Advanced Transportation Management Systems (ATMS), vehicle detection systems, radars, and Closed-Circuit Television (CCTV) cameras.
- **Control Systems:** Traffic signal and ramp meter controllers.
- **Data Providers:** Regional Integrated Transportation Information Systems (RITIS), Caltrans Performance Measurement System (PeMS), Inrix, Google Maps data, Waze data.
- **Local Data Collection:** Subjective survey responses, local Automated Traffic Signal Performance Measures (ATSPM), Realtime General Transit Feed Specification (GTFS-RT) data, ridership and transit ticket sales, micromobility device usage including bikes and scooters, and usage data from local apps like the 511 Contra Costa bike mapper.

Additionally, specific project-related data such as carpool statistics and express lane usage will be captured using traffic detectors and tolling systems, supplemented by targeted surveys. Safety data, including accidents and collisions, will be sourced from California Highway Patrol (CHP) portals and local police departments.

Environmental impact data, notably greenhouse gas (GHG) emissions reductions, will be collected from the Bay Area Air Quality Management District (BAAQMD) and other relevant sources. Operational and maintenance cost data will be obtained from public agencies' financial reports, ensuring a comprehensive analysis of cost efficiencies.

The State of Good Repair for infrastructure assets along the project areas will be evaluated using asset management systems and maintenance logs from Caltrans, MTC, and local city and transit agencies. Economic impact measurements, such as changes in sales tax revenue and job creation, will be coordinated with the Bay Area Council to assess the broader economic benefits derived from the projects.

This diverse and robust set of data sources will enable CCTA to conduct thorough before and after studies, ensuring all aspects of project impacts are meticulously documented and analyzed.

Data Gaps and Additional Data Needs

As CCTA consolidates data from multiple sources for the SMH, NB 680 Express Lane Completion (Phase 1), and CARM projects, CCTA's initial analysis will focus on identifying any potential gaps in the information collected. Should discrepancies or deficiencies in data emerge, we will employ various strategies to address these gaps, ensuring comprehensive coverage and accuracy in our findings.

To bridge these data gaps, CCTA may deploy temporary detection systems or conduct manual data collection efforts, such as targeted field surveys and observational studies. These methods will allow the gathering of essential data that may not be captured through our regular monitoring systems.

Data Format and Metadata Standards

Types of Data | In order to organize the data sets, CCTA considers two kinds of data per USDOT ITS JPO that is generated from any extramural research. The first is the primary research data, and the second is the research results.

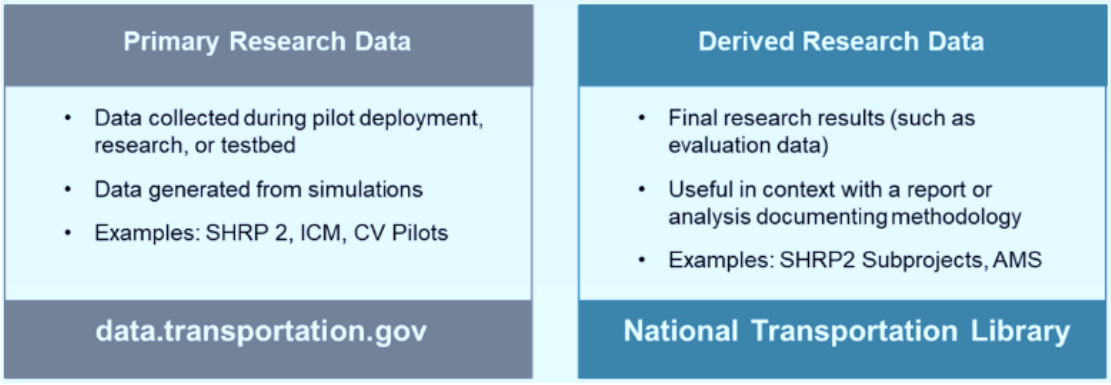


Figure 2: Mature Datasets[3]

The research results for this project are considered project reports. Chapter 3 of this document describes how reports will be gathered and maintained during the project as part of the DCP. Section 6 describes how the final publications of key deliverables can be posted to the National Transportation Library.

Scale of Data | CCTA understands that the data available through the various sources identified in Section 2.2 of this document may exceed the scale needed to perform the before and after study reports for this program. As a result, CCTA will work with each project component team members and stakeholders to collect and capture data required for the studies to support reports generated to measure the accuracy of the KPIs.

Data Formats and Metadata Standards | CCTA will utilize standard file formats, including CSV, text, Excel, Word, and PDF files, to collect and disseminate data to the extent possible. CCTA aims to use non-proprietary formats for all before and after studies to ensure accessibility and compatibility across various platforms.

Each project team will collaborate to document the data process log, detailing the data's final version released to the public. This log will also account for any alternative project builds and their impact on the studies. The teams will define the metadata schema and standards necessary to accurately measure Key Performance Indicators (KPIs) and anticipated outcomes. This will include specifying any additional tools or software needed to access, read, or interpret the data, ensuring that all stakeholders have the necessary resources to engage with the data effectively.

Quality control measures will be strictly adhered to, ensuring that all data conforms to the standards and plans outlined in this Data Collection Plan (DCP). Documentation, including any plans, drawings, or supporting materials developed for the projects, will be stored in the CCTA's Project Management Information System (PMIS), such as E-builder, as specified in Chapter 6.

[3] ITS JPO Data Access and Retention Policies, https://www.its.dot.gov/data/public-access/

Furthermore, visual data such as videos, photos, conceptual diagrams, and device configurations necessary for project demonstrations will be meticulously collected. This also includes any numerical data, text sequences, data messages, and logs obtained during the project implementation from vehicles, communications infrastructure, and related correspondence. All collected data will be processed and uploaded to CCTA's project SharePoint and E-builder systems for comprehensive review and use by project team members. This systematic approach ensures that all project data is managed effectively, maintaining integrity and facilitating ease of access for ongoing review and analysis.

Data Collection Methodology

The data collection work steps are identified in Section 1.3 of this report, and data sources are identified in Section 2.2. The data will be gathered periodically from sources available and reviewed for gap analysis, as identified in Section 2.3. CCTA will work with project teams to gather data needed for studies related to each project component and will gather them at a central repository.

Data Collection Planning | The Data Collection Plan described in this section will apply to all project components identified in Sections 1.4 and 2.1. The CCTA team will update this Experimental Data Collection Plan (DCP) to ensure the capture of data sets that achieve the project goal. In the Plan, we will identify the experiment's operational condition(s) and the data that needs to be collected to characterize those conditions with the following objectives:

1. collecting all data required to perform its evaluation,
2. ensuring the quality of the data,
3. storing data for processing, and
4. preparing reports and data for transfer to USDOT.

The DCP will also identify data sets gathered manually or automatically in real time. It will also identify if the data being gathered automatically can be checked with scripts (software codes) for quality in real time.

Each project component will have its data recording and maintaining data collection records. At a minimum, the data collection procedure will include the following:

1. Based on the operational scenarios, CCTA will start with a baseline (test run) data collection period before the regular data collection is activated. The test run will be jointly carried out with the designer and contractor for each project component. The baseline data collection is necessary to ensure the data is available and to verify the data elements are captured correctly.
2. If a disruption of data or abnormality is observed, we will work with the contractor and other agencies to capture a copy of logged or historical data from data sources identified in Section 2.2 where possible.

CCTA will also monitor the operational conditions to check if any local conditions that may change over time can lead to effects in the before or after study. For example, roadway closures or reconfigurations may cause traffic patterns to change significantly. During this period, the data will be captured, processed and reviewed, which will provide an updated status of conditions as well as an opportunity to verify the quality and volume of data elements being collected. If we detect any inadequacy or abnormality in data, the problem will be rectified expediently to minimize any disruption to the data collection process.

To maximize the quality of data to be captured, CCTA will adopt several approaches that are comprehensive:

1. Diversity in the driving environment, where we include ample examples of daily data in an urban setting, congested and free-flow, local streets and arterial (when applicable), and freeway driving environment;
2. Diversity in vehicle operation; for example, a mix of multimodal trips and personal-use vehicles;
3. Diversity in normal traffic and other road conditions.
4. Diversity in time of the day, weather, and lighting conditions.

In order to maximize the value of data and to investigate the accuracy of the KPIs, CCTA plans to select a set of strategic locations for data collection. These locations will include intersections, junctions, middle-of-road segments, and passenger pickup and drop-off areas. We will work closely with each project manager to review this strategy prior to data gathering. CCTA will also perform inspections and periodic monitoring and review of the captured data for accuracy.

Data Collection Testing and Verification | The CCTA team will work closely with the providers of the project components, along with other team members and stakeholders, to establish a unified and efficient method for executing data collection tasks. This collaborative approach will not only address the requirements for comprehensive data analysis but also pinpoint specific data elements essential for operational purposes.

To ensure the reliability and accuracy of the data collection methodologies, CCTA will implement a pre-testing phase. This initial testing stage will involve rigorous checks to validate the effectiveness and precision of the data collection processes. By verifying these methods early in the project, we aim to identify and rectify any potential issues, thereby guaranteeing that the data gathered meets our standards for quality and completeness before full-scale data collection begins. This structured verification process is critical in maintaining the integrity of our data and ensuring that it accurately reflects the project's outcomes and impacts.

Development of Data Collection Plan | CCTA will collaborate with members of the project component teams and stakeholders to refine and update the data elements outlined in the initial Data Collection Plan. This revision process will include a thorough examination of how the data elements are aligned with the analytical tasks necessary for assessing the project's impact.

CCTA will meticulously determine how these data elements correlate with performance metrics, Key Performance Indicators (KPIs), and other quantifiable outcomes expected for each project component. This careful alignment ensures that the data collection efforts are strategically positioned to capture the necessary information for evaluating the effectiveness and success of each initiative within the scope of our broader transportation projects.

Pre-Construction Data Gathering for Before Study Baseline | CCTA will implement a detailed process to gather, store, and archive the required data before construction begins. This baseline data collection is critical to establishing a comprehensive pre-construction snapshot, ensuring we have accurate benchmarks against which to measure project impacts.

Specific Data Collection for Each Project:

- **SMH:** CCTA will collect baseline data on current traffic volume, accident and collision statistics, transit ridership numbers, and existing greenhouse gas emissions levels in the areas surrounding proposed hub locations.
- **NB680 Express Lane Completion (Phase 1):** Baseline data will include current travel times, vehicle occupancy rates, toll tag usage, and frequency of accidents along the NB680 corridor. Additionally, current freight volumes and vehicle hours of delay (VHD) will be measured to assess pre-project performance.
- **CARM:** Current traffic density, VHD, and accident data related to ramp areas will be documented. Operational and maintenance costs associated with existing ramp metering systems will also be recorded.

The processes for gathering, storing, and archiving data will be rigorously verified to prevent any data loss. Should data loss occur, CCTA will promptly identify the causes and evaluate their impact to mitigate any potential effects on project integrity. This step is vital to recognize any technical limitations and ensure the efficiency and effectiveness of our data collection strategy.

Upon concluding the pre-construction data gathering phase, CCTA will refine the Data Collection Plan by specifying the precise data needs, types, formats, standards, sources, and identifying any remaining gaps. CCTA will outline methods to manually capture any additional data required to fill these gaps and detail the steps necessary for ongoing data collection. Furthermore, the team will engage in simulation and data analysis to confirm the accuracy of the forecasted KPIs. This thorough analysis will form the foundation of our Before Study report, setting the stage for effective project implementation and assessment.

Monitoring and Frequent/Periodic Updates to Data | CCTA, in collaboration with the project team members, will conduct frequent updates and rigorous monitoring of the collected data to ensure the accuracy of the KPIs and alignment with the program's objectives. CCTA will regularly refresh our simulation models with the latest data, allowing for ongoing comparisons that gauge the effectiveness of various construction alternatives. This iterative process ensures that each project component remains on target to meet our overarching goals.

Modify Data Collection Plan | Following the initial baseline data collection and subsequent periodic reviews, our project team will evaluate the efficacy of the current Data Collection Plan. This assessment will focus on whether the data gathered is of sufficient quality and quantity to confidently meet the projected KPIs and support comprehensive studies. Based on this evaluation, we will identify any necessary changes or enhancements to the data collection strategies, ensuring our methods remain robust and responsive to project needs.

After construction Data Gathering for After Study | Post-construction, CCTA and project teams will undertake a detailed process to update and analyze the data collected up to six years after the project's completion. This extended period of data gathering is critical to observe long-term impacts and effectiveness. The collected data will be meticulously analyzed and compared against the pre-construction baseline to comprehensively articulate the project's outcomes. These insights will culminate in the After Study report, which will document the achievements and areas for improvement, providing valuable feedback for future infrastructure projects.

DATA COLLECTION PROCESS

All project documents, including technical reports and research results, photos, images, and videos taken during the project, will be considered project-generated data that will need to be tracked and maintained. Some of the work items may be considered deliverables or publications that will be submitted to the USDOT as described in Section 6 of this document. In order to maintain document control, CCTA will implement a Control Plan (CP) to manage all documents. The CP would be a roadmap to track, add, archive, and remove documents from the system. Document control will largely focus on project documentation; however, it is intended to be scalable to accommodate both project and non-project information. Some examples of non-project documents include product manuals, stand-alone drawings, and various procedural manuals that apply broadly to the program that can improve assets State of Good Repair.

This CP shall identify both technical and administrative direction for the control of changes and integrity of the program data and documentation. The CP will identify the configuration of the software and hardware at given points in time, systematically controlling changes to the configuration and maintaining the integrity and traceability of the configuration throughout the project's life cycle.

The Plan establishes a framework and workflow processes for managing documents, including:

1. Systematic capture filing, distribution, and archiving of project documents and data and relevant non-project documentation;
2. Maintenance of a single source of truth for documentation;
3. Indexing documents such that files are logically organized and easily retrievable;
4. Controlled distribution; and
5. Maintenance of historical files for defined time periods.

Document management is the process of organizing, storing, protecting, and sharing documents. The CP describes how to manage documentation and provide a consistent approach to the creation, update, and distribution of material. The CP provides an overview of the organization, activities, tasks, and objectives of the project and identifies the roles and responsibilities of Configuration Control Board (CCB) members and Configuration Management Office (CMO). CCTA will lead the CCB as the supervising agency, and the CMO tools (E-builder) will be used by all and maintained by staff. The processes defined in this CP will remain in place for the life of this program.

The document control measures will ensure that all documents are properly stored and versioned, accessible to authorized users, distributed based on approved distribution lists, and properly secured and archived. The documents will be stored on the E-builder platform provided by CCTA for the life of this program.

The following figure illustrates the Document Control and Configuration Management process that will be in place during the life of the program.

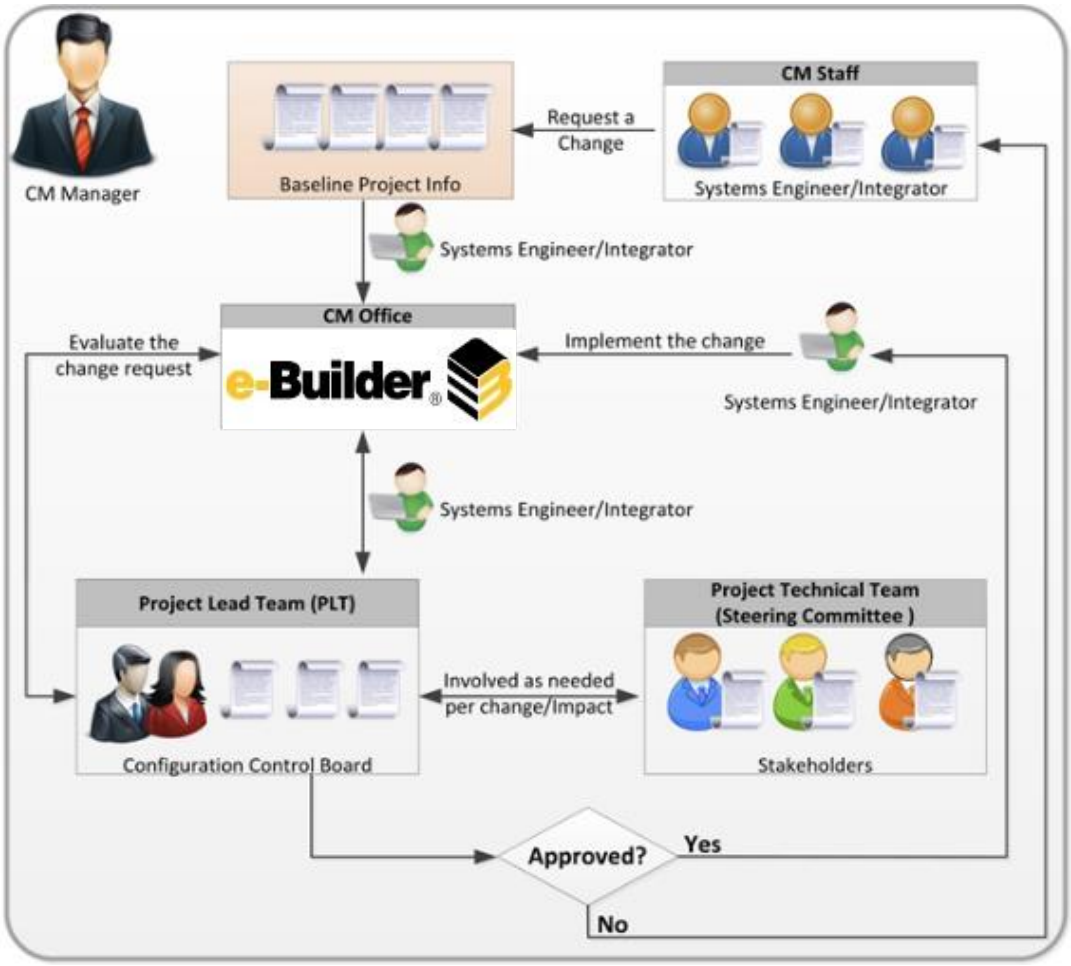


Figure 1 : Document Control Process

Data Collection by Project Component

Based on the methodology described in Section 2.5 of this document for data collection, CCTA will work with team members of each project component to provide their plan before and after the study in accordance with this DCP. Each project component that receives funding from this grant will be required to prepare and submit a DCP to CCTA. CCTA will then gather all the DCPs for each project component and generate the overall DCP for the *Innovate 680* program. As mentioned in Section 1.3 of this document, this section may be updated upon the release of Mega Grant's standard framework. Each DCP for each project component will need to include the following in their plan:

Project Component Overview | There will be an introduction section describing each project component. It will include a summary and purpose of the project component, its goals and objectives, the overall schedule from planning to deployment, and expected benefits and KPIs. It will also identify if the project component is purely construction or if it will improve traffic conditions, transit, environment, economics, or if it is a mix of a few areas and what are the characteristics of the project component in terms of cost saving and improving assets and state of good repair in the area.

Data Collection Plan Overview | There will be a section describing how the project component will handle digital data during and after the project component is deployed and delivered. This section will include a 2 to 3-page narrative describing the following:

1. The final project component data to be produced in the course of the project.
2. The expected outcome of the project component and how it can be monitored and/or verified.
3. How the data can be used to perform a before-after study and verify the accuracy of the KPIs and benefits expected from the project component's outcome.

Data Description | There will be a section to describe further the data that will be gathered in the course of the project component that can be used to perform before and after studies and verify the accuracy of the KPIs and expected benefits of the project component. In addition, it will include the following:

1. Describe the nature, scope, and scale of the data that will be collected.
2. Identify types of data (e.g., numerical data, image data, text sequences, video, audio, database, modeling data, source code, etc.)
3. Describe the characteristics of the data and their relationship to other data and provide sufficient detail so that reviewers can understand any disclosure risks that may apply. Discuss the value of the data over the long term.
4. Describe methods for creating the data (e.g., simulated, observed, experimental, software, physical collections, sensors, enforcement activities, researcher-generated database, tables, and/or spreadsheets, the instrument-generated digital data output such as images and video, etc.)
5. Describe the period of time data will be collected and the frequency of updates.
6. If existing data is used, what is the relationship between the data the project will collect and the existing data.
7. Indicate the sources of the data and the party responsible for managing the data.

Standards Used | There will be a section that identifies metadata standards used as well as anticipated formats that project component data and related files will use. If non-proprietary formats and tools are used, this section will identify them and the rationale for using such tools.

1. Identify the format(s) of the project component's data that will be collected for the before and after study and if they are open or proprietary.
   - If proprietary data formats are used, discuss the rationale for using those standards and formats.
   - Describe alternatives and how the data in proprietary and non-standard formats can be documented in a sharable and reusable format.
   - Indicate what tool or software is required to reach or view the data.
2. Describe how versions of data will be signified and/or controlled.
3. Identify documentation that will be created as part of the project component that can be used to make the data understandable to others for the purpose of before and after the study and measuring KPIs and expected outcomes of the project.
4. Describe the quality control measures and how the project component will verify the accuracy of the collected data.
5. Identify standards and metadata schema used to collect, manage, and store data.

Access Policies | This section describes if data needed from the project component's before-and-after study can be publicly shared or if it may include personally identifiable information, confidential business information, and/or classified information. This section will also describe the process for access restriction that may apply to the project component data, the process to make data anonymized and/or sharable to others, including a 3rd party, CCTA, and USDOT for the purpose of re-use and/or verification in terms of accuracy and performing before and after study.

Re-Use, Redistribution, and Derivative Products Policies | This section will describe who holds the intellectual property rights for the data created by the project component. Describes whether those rights will be transferred to a data archive and will identify any applicable copyrights to the data in a case where copyrighted instruments are used. It will also describe if data needed for the before-after study can be reused if there is a need for data citation through a license, and if there are terms of use or other policies that may apply.

Archiving and Preservation Plan | This section will identify where the project component's data will be archived and when the data will be available and provided to CCTA and USDOT designated storage and archive locations for the project delivery. In addition, this section will describe the following:

1. Identify where data will be stored prior to being sent to an archive.
    - Describe how back-up, disaster recovery, off-site data storage, and other redundant storage strategies will be used to ensure the data's security and integrity.
    - Describe how data will be protected from accidental or malicious modification or deletion prior to receipt by the archive.
    - Discuss data archive policies and practices for back-up, disaster recovery, off-site data storage, and other redundant storage strategies to ensure the data's security and integrity for the long term.
2. Indicate how long the chosen archive will retain the data.

Data Collection Schedule

A high-level schedule for the data collection plan is included in Section 1.2 of this document. As part of the program, CCTA will develop a detailed schedule for data collection to identify when each project component's DCP will be developed, what are the end dates for each project component and identify the time needed for before study and development of the baseline and what is the schedule for after study The schedule will be included in the appendix of this document and will be frequently updated as the project progress.

Data Management Strategies

CCTA will gather each project component's DCP and review them to identify strategies for managing each project component's data collection process, to collect, prepare, store, assemble, analyze, and distribute data needed for the before and after study, periodic monitoring, frequent updates, and development of the reports needed as part of this program. CCTA will also review project components DCPs to identify if there are duplicate efforts in gathering data. CCTA will work with team members of each project component to identify the schedule for the data collection and if efforts can be combined or coordinated to perform data gathering more cost-effectively while maintaining or improving the quality, scale, and accuracy. CCTA will also implement processes and procedures for data quality control and data archiving where applicable.

ASSEMBLING AND ANALYZING DATA

As outlined in Chapter 3, CCTA will instruct each project team to prepare a Data Collection Plan (DCP) for their respective segments: Shared Mobility Hubs (SMH), NB680 Express Lane Completion (Phase 1), and Coordinated Adaptive Ramp Metering (CARM). We will then collaborate with these teams to consolidate the data into a central repository. This centralization will facilitate data completeness assessment, help identify gaps, and avoid duplicative data collection efforts.

Specific Review and Simulation Model Development:

- CCTA will review each project's DCP to ensure all data types necessary for effective simulation models are collected. These models will simulate existing conditions and forecast the KPIs, which will be essential for assessing the impact of the projects.
- For the **SMH**, **NB680 Express Lane Completion**, and **CARM** projects, we will develop simulation models that represent the infrastructure's current physical geometry and the anticipated changes post-construction. These models will include details such as lane configurations, traffic volumes, and control systems to forecast how the changes will impact traffic flow, safety, and environmental factors.
- Each simulation will be tailored to the specific characteristics of the projects, ranging from microsimulation for detailed traffic behavior analysis to macroscopic models for broader traffic trends.

Data Calibration and Additional Analysis:

- The simulation models will be calibrated with actual data to verify the accuracy of the forecasted KPIs. This calibration will involve adjusting model parameters to align simulations closely with real-world observations.
- We will incorporate traffic control data, such as ramp metering rates from the CARM project, and detailed traffic movement data from the SMH and NB680 projects to enhance the models' precision.
- Additional travel demand data, such as traffic density, speed, and composition, will be integrated to refine the models further.

Advanced-Data Analysis Techniques:

- Advanced traffic engineering techniques, supported by guidelines from USDOT, Caltrans, and other authoritative sources, will be used to analyze the data. This includes employing ITS4US complete trip data to address mobility challenges for all travelers, particularly focusing on underserved communities.
- We will also utilize data from existing Transportation Performance Management (TPM) reports and Transportation Demand Management (TDM) studies to establish baselines and analyze the effectiveness of the implemented measures.
- Environmental, safety, and operational data will be analyzed to develop comprehensive reports on the project's impacts on air quality, quality of life, and state of good repair.

Innovative Data Management Practices:

- CCTA will use E-Builder and other project management tools to effectively maintain and archive all project-related data and documentation. This will ensure that all data, from initial plans to final analysis, is stored securely and is accessible for review and auditing purposes.
- Additionally, Building Information Modeling (BIM)[4] practices will be explored to enhance the visualization and management of construction projects. BIM will facilitate the integration of various data types and improve collaboration across project teams.

This comprehensive approach to assembling and analyzing data ensures that CCTA can effectively monitor project progress, evaluate impacts, and make informed decisions to optimize project outcomes based on detailed empirical evidence.

Maintaining Baseline data

CCTA recognizes the necessity of establishing and maintaining robust baseline data across multiple categories for this program. To effectively manage this, CCTA will collect, integrate, and synchronize the baseline data from each project—SMHs, NB 680 Express Lane Completion (Phase 1), and CARM - into a centralized data repository. This repository will serve as the foundational reference point for all subsequent analyses and reporting.

As the projects progress, CCTA will regularly update and refine this consolidated baseline to ensure it accurately reflects the current status and conditions before the implementation of project interventions. This ongoing maintenance is crucial for tracking changes and measuring the impact of the projects effectively.

A comprehensive report will be compiled once the baseline data is finalized and validated. This report will encapsulate all baseline metrics and initial conditions for the entire scope of the Mega Grant program, including detailed descriptions of project components. CCTA will then share this baseline report with the USDOT to provide a clear and comprehensive reference point that outlines the starting benchmarks against which the success and impact of the projects will be measured.

Monitoring Key Performance Indicators per Project

CCTA has outlined preliminary KPI measures in Section 1.4 of this document, which may be refined based on the Data Collection Plans (DCPs) provided by each project component, as detailed in Chapter 3. We will actively monitor these KPIs, continuously updating and verifying the baseline data to ensure the accuracy of the anticipated outcomes throughout the project's duration. If discrepancies arise or adjustments become necessary due to changes in project designs or alternative analyses, CCTA will promptly revise the existing conditions reports and KPI expectations. These updates will also be reflected in an amended before-study report submitted to USDOT, ensuring all stakeholders are informed of the current project status and anticipated impacts.

Forecasting and Simulating data

Forecasting and simulation methods critical for predicting the outcomes of KPI measures will be thoroughly documented in both the before and after study reports. These methods will adhere to the established guidelines, handbooks, and best practices as previously described in Chapter 4. In the event of unforeseen changes or the adoption of new forecasting techniques, CCTA will collaborate with USDOT

to adjust the forecasted data appropriately. This will ensure that our simulation models and studies remain accurate and relevant, facilitating the provision of updated reports to accurately reflect any new insights or developments.

Monitoring and Updating Data

CCTA will employ a combination of automated, semi-automated, and manual data collection methods to support the comprehensive before and after studies. Our approach is designed to allow for periodic updates to the data sets, which will enable ongoing recalibration of the simulation models and baseline data. This is crucial for maintaining the integrity of the KPI measurements and ensuring that the projects align with their defined goals. Regular data updates and model reruns will be conducted to validate the accuracy of the KPIs and confirm that the projects remain on track to meet their objectives, adjusting strategies as necessary to address any emerging challenges or opportunities.

Generating data reports

CCTA will perform and deliver the following reports:

1. Before Study (project baseline) report: before the date of construction of a project, CCTA will submit a Before Study report providing baseline data to analyze the project's long-term impact in accordance with the Mega Grant Standard framework established per Section 1.3 of this document.
2. After Study (Updated project baseline) report: Not later than six years after the date of substantial completion of a project, the eligible entity carrying out the project shall submit a project outcomes report that compares the baseline data to quarterly project data for the duration of the fifth year of the project after substantial completion.

ACCESS POLICIES AND SHARING

Policies for Access and Sharing

CCTA anticipates that the before and after study reports developed as part of this project can be shared publicly. CCTA will work with each project component team member to identify if any portion of the data cannot be shared publicly and may require access restrictions due to the presence of confidential or privacy-related data. If access restrictions are necessary, CCTA will work with a 3rd party to perform data anonymization prior to sharing the data with the public.

The project managers are expected to ensure that privacy and confidential information is protected before the archive or public release of data.

Policies for Re-use and Redistribution

It is expected that the researchers performing the before and after study reports to measure the ability of the project to meet the expected KPIs will cite their data sources and the license under which they used the data in the reports. Researchers are also expected to describe if they are transferring rights to the data archive and the process and methods used to perform the calculations to be available for reuse and redistribution.

CCTA will work with each project components team member to identify who has the right to manage the data, indicate who holds the intellectual property rights to the data, list any copyrights to the data and indicate who owns them, discuss any rights to be transferred to a data archive, and to describe how data will be licensed for reuse, redistribution, and derivative products.

Currently, CCTA anticipates that almost all the data needed to perform the before and after study will be collected from publicly available sources.

Plans for Archiving and Preservation

CCTA expects all the data gathered and collected under each project component to be archived in storage locations identified in Chapter 6 of this document. The data used to generate the before and after study reports required by Mega Grant will also be copied to a different archive repository that will conform to the USDOT data repository requirements. The data will be duplicated regularly to prevent catastrophic loss of information. Information will be backed up and mirrored at another site to provide a means of recovery in the case of a disaster.

DATA STORAGE

As part of this program, a series of data storages will be used to store gathered data, reports, documents, and work-in-progress data being assembled, analyzed, and archived. The contractors may include a copy of the data they need for their work on their file storage systems. They must provide a copy of the raw data, processed, assembled, analyzed, and forecast data to CCTA in near real-time and always keep an updated copy available to CCTA. Contractors will be required to upload a copy of the data to the following storage locations and CCTA will perform periodic verification of accuracy and completeness.

3. CCTA's E-Builder is a cloud-based system that will be used to store information related to construction projects, including all the information related to planning, design, procurement, construction, and operations, including the project's schedule and budget.
4. SharePoint Online on Microsoft Azure 365 Cloud;
    - Project reports and technical documents may be stored on CCTA and/or consultants' SharePoint for sharing data and work-in-progress technical reports between project team members for collaboration. Draft and Final documents and their backup data will be copied to CCTA's E-Builder.
5. Telegra Advanced Transportation Management System (ATMS) Local Storage servers at CCTA's IT Room;
    - Real-time and historical data gathered from infrastructure and devices along the freeway and arterials will be stored on the Telegra ATMS SQL Server database. This raw data will be gathered and used to generate reports where needed.
6. CCTA's Amazon Web Services (AWS) Cloud services and tools;
    - CCTA's AWS services will be used to store specific tools and data-gathering components needed for gathering information related to Transit, Ridership, Transfers between modes via Connection Protection, etc. This raw data will be gathered and used to generate reports where needed.
7. IBM's Engineering Lifecycle Management (ELM) system;
    - Information such as project goals, objectives, user stories, functional and performance requirements, system and subsystem requirements, design requirements, test requirements, test procedures, test results, and verification and validation processes will be stored on ELM in addition to the requirements traceability matrix for various components of the projects. This information will be used to assist with the preparation of Systems Engineering Documentation following FHWA's guidebook for Systems Engineering.
8. USDOT and National Library Websites for Data Sharing;
    - CCTA will work with USDOT to provide a copy of project-related Publications and technical data to applicable USDOT websites where needed.

Policies for Data Storage

All data storage utilized for this project will adhere to or surpass the data storage requirements set by the USDOT.

The program's data storage should facilitate access control and restrictions where necessary while also enabling public access to the data. These storage systems must meet or surpass cybersecurity standards, with robust security measures like SSL certificates and/or encrypted VPN connections in place to prevent unauthorized access or modification of archived reports and their data.