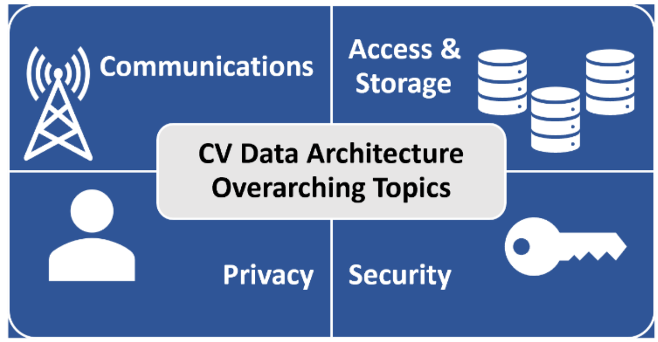


Model Connected Vehicle Data Architecture

FINAL

March 31, 2023

Prepared for:
The Connected Vehicle Pooled Fund Study
(University of Virginia Center for Transportation Studies)

Prepared by:
Athey Creek Consultants, LLC

Contents

Tables

Figures

List of Acronyms

AADT	Annual Average Daily Traffic
AI	Artificial Intelligence
API	Application Programming Interface
ATIS	Advanced Traveler Information System
ATMS	Advanced Traffic Management System
Austroads	Australian Road Authority
AVL	Automatic Vehicle Location
BSM	Basic Safety Message
C-V2X	Cellular Vehicle-to-Everything
CAMP	Crash Avoidance Metrics Partnership
CFR	Code of Federal Regulations
CI	Connected Intersections
CTI	Connected Transportation Interoperability
CV	Connected Vehicle
CV PFS	Connected Vehicle Pooled Fund Study
CVDF	Connected Vehicle Data Framework
DMS	Dynamic Message Sign
DOT	Department of Transportation
EDCM	Event Driven Configurable Messaging
FHWA	Federal Highway Administration
GLO	Green Light Optimization
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
I2V	Infrastructure-to-Vehicle
IMA	Intersection Movement Assist
IOO	Infrastructure Owner Operator
IRIS	Intelligent Roadway Information System
ITE	Institute of Transportation Engineers
ITIS	International Traveler Information Interchange Standard
ITS	Intelligent Transportation System
JPO	USDOT's Joint Program Office
JSON	JavaScript Object Notation
LC	Lane Closure
MDO	Managing Disruptions for Operations
MEC	Mobile Edge Computing
MPH	Miles Per Hour
MPO	Metropolitan Planning Organizations

NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NIST	National Institute of Standards and Technology
NTCIP	National Transportation Communications for Intelligent Transportation Systems Protocol
O-D	Origin-Destination
OBU	On-Board Unit
OEM	Original Equipment Manufacturer
PDMS	Probe Data Management System
PFS	Pooled Fund Study
PII	Personally Identifiable Information
PSID	Public Safety Identification
QA/QW	Queue Advisory / Queue Warning
RDI	Roadway Digital Infrastructure
RITIS	Regional Integrated Transportation Information System
RLVW	Red Light Violation Warning
RSU	Roadside Unit
RSZW	Reduced Speed Zone Warning
RTCM	Radio Technical Commission for Maritime Services
RWIS	Road Weather Information System
SCMS	Security Credential Management System
SPaT	Signal Phase and Timing
SRM	Signal Request Message
SSM	Signal Status Message
SSP	System Security and Privacy
SWIW	Spot Weather Information Warning
TIM	Traveler Information Messages
TIMDx	Transportation Incident Management Data Exchange
TLS	Transport Layer Security
TMC	Transportation Management Center
TMDD	Traffic Management Data Dictionary
TOSCo	Traffic Optimization for Signalized Corridors
UPER	Unaligned Packed Encoding Rules
USDOT	United States Department of Transportation
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VIN	Vehicle Identification Number
VRU	Vulnerable Road User
XML	Extensible Markup Language
WZDx	Work Zone Data Exchange

Executive Summary

Background

For Connected Vehicle (CV) data to be assets to state and local Infrastructure Owner Operators (IOOs), an organized approach is needed for communicating and managing the data. IOOs have limited resources to dedicate to ingesting, procuring, managing, storing, and providing CV data. The Connected Vehicle Pooled Fund Study (CV PFS) developed this *Model CV Data Architecture* to support IOO practitioners as they make decisions and implement processes for ingesting and/or providing CV data. Key sections of this report include:

- **Needs Addressed by CV Data.** Challenges and needs of drivers and IOOs that CV data can address.
- **The Model CV Data Architecture.** Use cases, requirements, and concepts that IOOs should consider for "how" to plan local CV data implementations.
- **Systems Integration Guidance.** Steps to navigate the model CV data architecture to reach key decisions about "what" CV data may be appropriate.

Key Concepts for CV Data

CV data communications may occur directly between roadside infrastructure and vehicle on-board units via cellular vehicle-to-everything (C-V2X) communications. However, CV data also includes indirect communications involving network cellular communications, typically through private vendors and aggregation portals that serve as third parties to communicate CV data between IOOs and vehicles, as illustrated below.

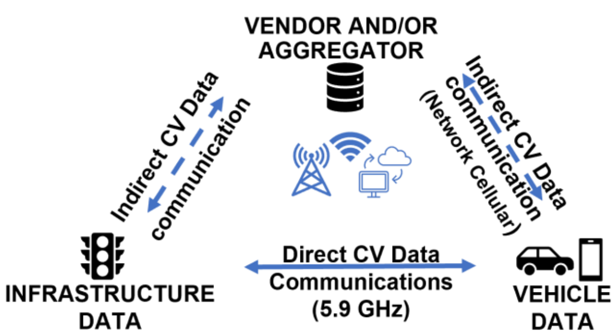


Figure ES-1: CV Data Communications

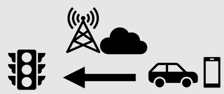
Three key concepts of CV data described in this document are:

- ***CV data is not just direct data***. This document describes the ingest of data from vendors as well as the aggregation and output of agency data via Internet feeds or Application Programming Interfaces (APIs).
- ***IOOs should follow a needs-based approach to ingest and provide CV data.*** Not all available CV data will be useful and worth expending resources; IOOs should ingest vehicle data based on IOO needs and provide infrastructure data based on driver needs and vehicle readiness to receive data.
- ***Consistent provision of CV data.*** IOOs should leverage national approaches and standards for connected intersections and work zone data, as well as emerging approaches such as the Transportation Incident Management Data Exchange (TIMDx).

Needs Addressed by CV Data

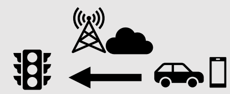
Critical data needs can be addressed by CV data. Specifically, this report presents challenges and needs of drivers and IOOs that CV data can address, defining a list of CV data. The document then describes use case examples for CV data uses that address various needs, including those presented below.

Ingest of Data: Direct



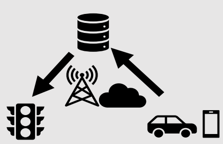
- Utah DOT uses signal preemption to improve transit vehicle on time status and snowplow efficiency.
- Utah DOT receives Basic Safety Messages (BSMs) from vehicles, which enable creation and broadcast of Traveler Information Messages (TIMs) to other vehicles.
- Tennessee DOT collects BSMs for various use cases (e.g., to understand energy efficiency and predict potential emissions and fuel savings).

Data Provision: Direct



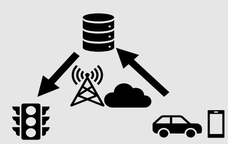
- Connected Intersections in Utah, Georgia, and Ohio were tested against current standards to ensure they will be trusted to support RLVW applications.
- Live Traffic Optimization for Signalized Corridors (TOSCo) is demonstrated for automated longitudinal vehicle control to optimize traffic movements.
- North Carolina DOT directly delivers work zone data to participating vehicles.

Ingest of Data: Vendor Data



- Minnesota DOT receives CV data aggregated into average travel speeds to display current travel times on dynamic message signs (DMS).
- Kentucky Transportation Cabinet uses crowdsourced data to identify road weather hazards and monitor work zone performance.
- Maricopa County DOT uses CV data to perform signal timings, eliminating the need for floating car surveys.
- Maryland DOT accesses CV data through an aggregation portal to support post-event analyses and planning for managing future events.
- Many agencies ingest vendor data to provide speed data on traveler information maps.
- Multiple agencies use vendor data to improve situational awareness about delays and incidents.
- Multiple agencies are exploring hard braking, acceleration, and telematics data for safety and planning purposes.

Data Provision: API or Data Feeds



- Current signal status reported on Internet feeds from centralized signal controllers in multiple cities are received by private vendors and communicated to vehicles to support green light optimization applications to reduce fuel consumption.
- Multiple IOOs share work zone data using the WZDx data specification via API data feeds.
- Multiple IOOs share event and condition data with OEMs and third parties to communicate to travelers and provide in-vehicle alerts.

Figure ES-2: CV data uses that address various needs

The Model CV Data Architecture

Research Approach

This *Model CV Data Architecture* is the result of a research project that included three activities for information gathering:

- **Review of recent research and development products.** A total of 44 resources were reviewed as background information about CV data.
- **Infrastructure Owner Operator (IOO) Interviews.** Fourteen (14) transportation agencies participated in interviews in 2022 to share their experiences receiving, using, and storing CV data, as well as their experiences sharing CV data to vehicles.
- **Private Industry Survey.** An internet-based survey was conducted to understand private sector vehicle data products and services, as well as interest from private companies in receiving infrastructure data.

Each of these research activities were focused around four overarching topic areas: communications, access and storage, privacy, and security.

Model CV Data Architecture Content

The core content of the *Model CV Data Architecture* are requirements for each of the four overarching topic areas shown in Figure ES-3, with supporting discussion about emerging CV trends that include the national strategy on roadway digital infrastructure, the SAE Probe Data Standard, and other emerging standards. These requirements were derived from use cases that were developed based on the interviews, survey, and research in this project. Additional concepts are also described for IOOs to consider as they implement the architecture. See Figure ES-4.

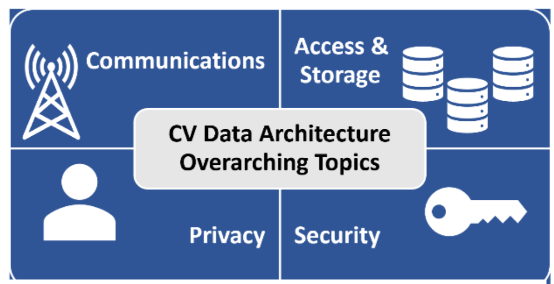


Figure ES-3: CV Data Architecture Overarching Topics

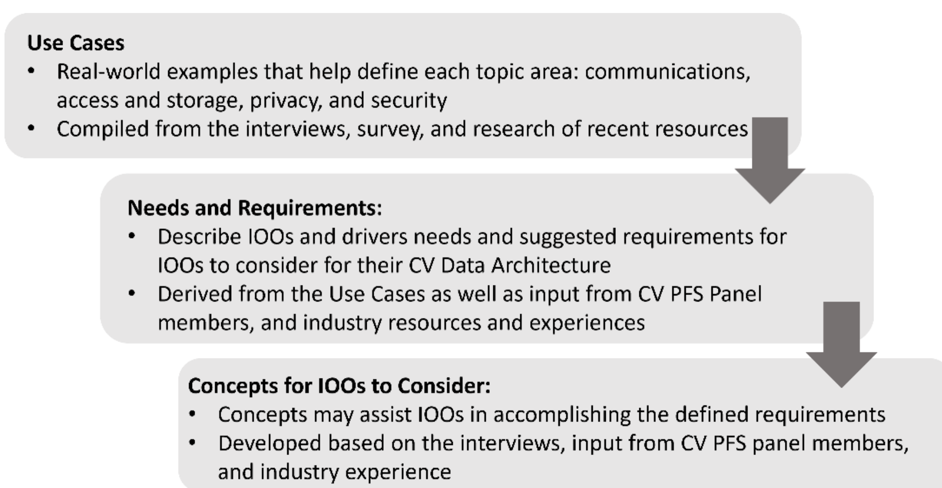


Figure ES-4: CV process and structure of topics addressed in this Model CV Data Architecture

Systems Integration Guidance for the Model CV Data Architecture

Chapter 3 of this document provides step-by-step Systems Integration Guidance, intended to support CV data decisions through a series of steps on how to navigate the *Model CV Data Architecture* to reach key decisions about what CV data may be appropriate, which data sources might best support the needs, and requirements and concepts to plan how to implement CV data.

The *Model CV Data Architecture* Systems Integration Guidance includes steps for ingesting CV data and providing CV data, as summarized below. These steps in the Guidance include references to the appropriate sections, needs, and information within the *Model CV Data Architecture* to support IOO practitioners as they make decisions and implement processes for ingesting and/or providing CV data.

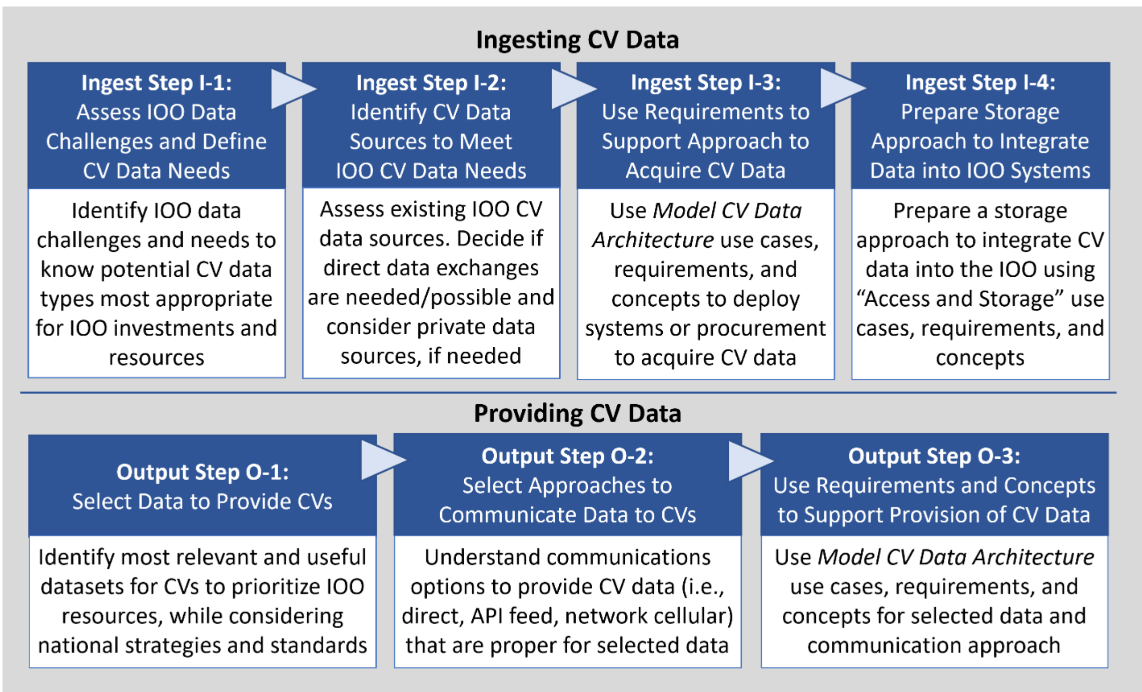


Figure ES-5: Steps for Ingesting CV Data and Providing CV Data

Preface: How to Use This Document

This *Model Connected Vehicle (CV) Data Architecture* document is intended to support a variety of users in transportation with the incorporation of data needs and sources within their agencies. The following table provides guidance to readers on how to approach using the content of this document for their specific needs.

Table P-1: **How to approach using the document content for specific needs**

Readers who want to…	Refer to…
Understand Background Information on the State of CV Data in 2023	Chapter 2 presents an overview of current CV data and the communications approaches that support them. Chapter 2 also discusses two "emerging" CV communications approaches (e.g., the Roadway Digital Infrastructure and the recently published SAE Probe Data Standard).
Identify Agency Priorities for Receiving or Provisioning CV Data	Section 4.1 (Data Needs and Uses) discusses vehicle drivers' and IOOs' challenges and data needs; and presents over 40 CV data types that may address one or more challenges of either the drivers or the IOOs. Section 4.1 concludes by grouping these 40+ data elements into eight Data Categories that are used to organize the remainder of the model data architecture.
Develop Local Plans for Deploying Infrastructure Systems to Communicate Directly with CVs (to send or receive CV data)	Sections 4.2-4.5 describe use cases, requirements, and concepts that IOOs should consider as they plan local CV data implementations; those that begin with **Direct** relate to direct communications with CVs.
Develop Local Plans for Procuring CV data from Private Vendors or Aggregation Services	Sections 4.2-4.5 describe uses cases, requirements, and concepts, and those that begin with **Vendor/Aggregation** are specific to these procurements.
Develop Local Plans for Provisioning Data to CVs Indirectly Through Internet Feeds	Sections 4.2-4.5 describe uses cases, requirements, and concepts, and those that begin with **API/Feed** (i.e., application programming interface or data feed) are specific to IOOs provisioning data through internet feeds to reach CVs through third-party providers.
Understand Current CV Data Sources	Section 4.6 describes CV data sources and relates these sources to the requirements and concepts described in the data architecture.
Understand How to Implement the Contents of this Document Based on their Agency Needs	Chapter 3 (Systems Integration Guidance) includes a series of steps describing how to navigate this document to reach key decisions about "what" CV data may be appropriate, "which" data sources might best support the needs, and how to identify requirements and concepts to plan "how" to implement CV data.

Overall Navigation of this Document

The header of each page of this document for Chapters 1-4 always displays the overarching sections for the whole document and highlights what section you are currently reading. The header changes, page by page, to identify specific subsections.

1.0 Introduction

For Connected Vehicle (CV) data to be assets to state and local Infrastructure Owner Operators (IOOs), there must be an organized approach for communicating and managing the data. The sooner that a critical mass number of IOOs organize around consistent CV data architectures, the sooner the benefits of CV will be adopted and scaled. Note that the definition of CV in this document is clearly outlined in Chapter 2.0 and is **not intended to reflect only data exchanged within the 5.9 GHz spectrum.**

While there is currently no national plan for CVs, IOOs are considering and deploying systems and infrastructure to communicate with connected vehicles and increasingly procuring CV data sources from private vendor sources.

Document Purpose

A resource for IOOs to make decisions about the CV data they produce and share, as well as the CV data they receive and use.

The purpose of this effort, *Model CV Data Architecture*, conducted by the CV Pooled Fund Study (PFS) was to create a resource for IOOs to use as they make decisions about the CV data they produce and share; and the CV data they receive and use. The information presented is to assist IOOs as they craft their local CV data architecture or append their agency data architecture to include CV data.

This document was developed based on reviews of technical resources, interviews with IOOs, and input provided by original equipment manufacturers (OEMs) and private vendors through an online survey.

1.1 Document Structure

Following this document overview, Chapter 2 presents an overview of CV data, including a definition, a summary of the vision for using CV data, progress toward that vision, and a description of various communications approaches for CV data. Chapter 3 contains *Systems Integration Guidance* for the use of this model CV Data Architecture. Finally, Chapter 4 the *Model CV Data Architecture* includes six sections:

1. *Data Needs and Uses*. This section identifies what gaps or challenges may be addressed by CV data and encourages a needs-based approach to CV data.
2. *Communications Requirements*. This section describes communications needs, requirements, and concepts for CV data.
3. *Access and Storage Requirements*. This section describes needs, requirements, and concepts for data access and storage for both the data received and sent by IOOs.
4. *Privacy Requirements*. This section provides needs, requirements, and concepts for IOOs to consider as they preserve user privacy.
5. *Security Requirements*. This section includes a description of needs, requirements, and concepts for CV data security.
6. *Data Sources*. This section summarizes currently available data sources and map data sources to CV data to support IOO decision making.

Appendix A contains links to all identified resources that were reviewed as part of this effort and References contains the specific resources cited in this document.

1.2 Target Audience

This document was created for decision makers and technical staff of state and local IOOs to support them in creating CV data architectures in their agencies or amending existing data architectures to add CV related content. Some examples of the technical staff that are expected to benefit from this document include:

- An agency's Intelligent Transportation Systems (ITS) staff responsible for making strategic and/or technical decisions about CV related deployments by the agency.
- An agency's planning group that may be purchasing and using CV data as part of their planning processes.
- An agency's operations groups that may be purchasing and using CV data for real-time information dissemination to travelers while also sharing agency gathered data with CVs.
- An agency's maintenance staff who may be collecting CV data from their maintenance vehicles and supplementing data with purchased data from private vendors to assist in prioritizing roadway maintenance.

There are many other instances of agency staff who are able to use CV data to support and enhance their business processes, and through this document they should be able to incorporate CV data within almost all facets of an IOO agency, as desired.

2.0 Connected Vehicle Data and Communications Overview

This chapter provides an overview of CV data and potential communications pathways for CV data. For this project, a focus only on CV data directly broadcasted between vehicles and infrastructure was too narrow a perspective, so an effort was made to specifically define a broader meaning of "CV data."

2.1 Connected Vehicle Data Definition and Background

This section defines CV data, provides context about the longstanding vision and potential role of CV data, and describes the current state of progress toward achieving that vision.

2.1.1 Definition of Connected Vehicle Data

For this report, CV data includes data that is communicated either directly or indirectly (i.e., via private vendors and/or applications) between IOOs and vehicles and has been triaged into the following latency categories:

- *Low latency,* defined here as 300 ms or less from when it is generated*,*
- *Real-time*, defined here as 5 seconds or less from when it is generated,
- *Near real-time*, defined here as 5 minutes or less from when it is generated, and
- *Historical data*, defined here as data with a latency higher than 5 minutes.

This report describes three types of CV data sources, as defined in Table 1, which are vehicle data, infrastructure data, and private vendor data (i.e., vehicle or infrastructure data that is collected, processed, and distributed by third-party vendors and aggregators).

Table 1: Description of the three primary types of CV data sources

Connected Vehicle Data	*Connected Infrastructure Data*	*Private Vendor Data*
• Includes data directly communicated to infrastructure, to vehicles around them, or to external third parties. • Communications can occur as direct (i.e., vehicle based 5.9 GHz cellular vehicle-to-everything (C-V2X) broadcasts) or indirect through network cellular based data exchanges between the vehicles and a central server or system. • Vehicle communications may communicate data continuously or on-demand (e.g., transit vehicles may only broadcast signal request messages when signal priority is needed).	• Includes data describing signal controller status, driving conditions, work zones, and other similar data. Includes both: ○ **Direct** 5.9 GHz C-V2X broadcasts of standardized messages from roadside units, and ○ **Indirect** as back-office data is shared using APIs or internet feeds that are typically ingested by third party applications or vendors that communicate the data to vehicles using network cellular data exchanges.	• Data that is collected by third-party vendors (typically through network cellular communication with vehicles) that is processed and provided to buyers (e.g., to IOOs via internet data exchanges), including: ○ Individual vehicle data (e.g., speed, heading, measured conditions); and ○ Aggregated data (e.g., average speed or travel times calculated from individual vehicle data). • Third-party vendors or aggregators typically process and deliver CV data to IOO users.

2.1.2 The Connected Vehicle Vision and Role of Data

The big picture vision for CVs has long been to enable a transportation future with improved safety, mobility, and system efficiencies through direct two-way communications between vehicles (i.e., vehicle-to-vehicle (V2V)), as well as direct or indirect two-way communications between infrastructure systems and vehicles or vulnerable users. This vision relies on CV data to improve situational awareness and enable an increasingly automated array of vehicle and infrastructure applications that support improved decision-making capabilities.

2.1.3 Progress Towards Recognizing the Vision for Connected Vehicle Data

The research and interviews conducted throughout this project confirm that the vision described above is not yet fully realized. Nonetheless, a lot of CV applications and commercial products are available that have proven to be beneficial with the current percentage of vehicles that are connected. Many of these CV applications that are operational today involve private vendors that receive data from vehicles and provide this data to IOOs. IOO's use of this data includes real-time or near real-time operational uses (e.g., using speed or travel time data to populate dynamic message signs (DMS) with travel times) as well as retrospective uses (e.g., using CV data to assess performance during a past winter storm, supporting signal retiming, or validating planning models). A literature review for this project identified examples that illustrate how extensive IOO use of CV data is. One example of this is a 2021 survey conducted through a National Cooperative Highway Research Program (NCHRP) effort, *NCHRP Synthesis 561: Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation*[1] that gathered information from 45 state departments of transportation (DOTs) plus the District of Columbia about the current state of practice for GPS-based probe data sources. These findings indicated that:

- *Speed data* was used or purchased by all 46 (100%) agency respondents, including real-time or historical average speed data on one or more road segments.
- *Origin-destination data* was used or purchased by 26 (57%) of agency respondents (i.e., data that describes the origin and destination of a trip).
- *Location-based services data* was used or purchased by 20 (43%) of agency respondents (i.e., data obtained from private vendor applications on smartphones that periodically report the location of users).
- *Route/trajectory data* was used or purchased by 10 (22%) of agency respondents (i.e., data describing the path and routes of a trip from an origin to a destination).

Figure 1 presents high-level categories identified in this NCHRP effort about how the 46 IOO respondents use Global Positioning System (GPS) probe-based data.

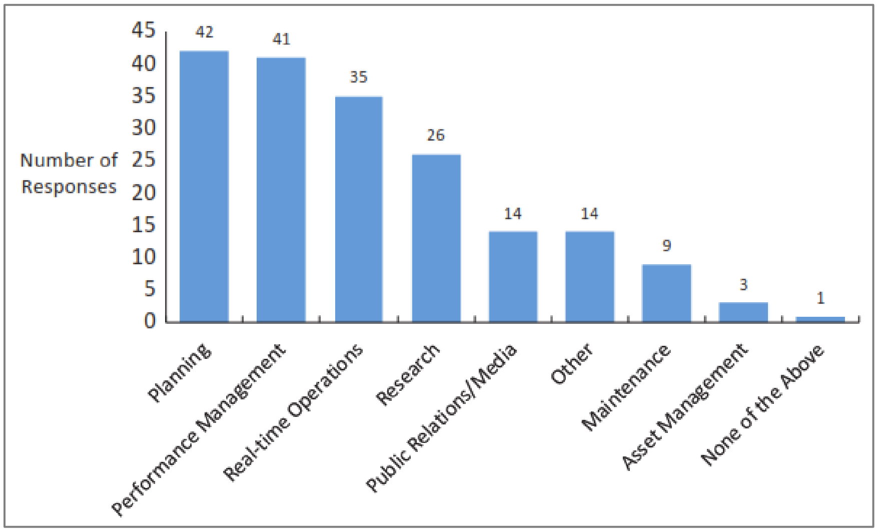


Figure 1: Uses of GPS-based probe data reported by 46 surveyed IOOs (Source: NCHRP Synthesis 561: Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation[1])

Research in this project found fewer examples of IOO data communicated to vehicles. Four examples are:

- The relatively limited but expanding number of connected intersections being deployed nationally that broadcast signal phase and timing (SPaT) and MAP information for CVs using low latency 5.9 GHz cellular vehicle-to-everything (C-V2X) communications.
- The provision of IOO work zone data to API or internet feed interfaces to communicate work zone information to private vendors or OEMs using the Work Zone Data Exchange (WZDx) specification.
- The provision of data from central traffic signal controllers describing current signal status (i.e., SPaT/MAP messages) over internet APIs that enable private sector companies to download the data and communicate over network cellular to their applications in vehicles, supporting systems such as Green Light Optimization (GLO), best route selection, and countdown to green.
- The provision of road conditions, closures, and other event information that are either broadcast from roadside units (RSUs) or communicated by the internet to private vendors that disseminate the information directly to in-vehicle applications.

Although the full vision of CV impacts is still not a reality and may be a long way off, both IOOs and industries continue to expand the availability and use of CV data as technologies, capabilities, and processes evolve and improve. That said, vehicle capabilities to directly broadcast, receive, and use IOO data, as well as IOO options to access fully representative datasets of CVs on the roadway, still need to mature and evolve to attain the original vision.

In this regard, there have been indications that vehicle capabilities to directly receive and use CV data are planned for some models of production vehicles starting in 2024, and regulatory certainty regarding C-V2X communications in tandem with the pending establishment of processes for connected infrastructure increases the probability that the CV vision will come to fruition in the relatively near future.

2.2 Summary of Current Connected Vehicle Data Communications Approaches

This section describes current CV data communications approaches for Vehicle to IOO and IOO to Vehicle data exchanges.

2.2.1. Vehicle to IOO - Current CV Data Communications Approaches

The current approaches used to communicate data from vehicles to IOOs that were documented in the interviews conducted in this study can be summarized as one of three data paths.

2.2.1.1 Vehicle to IOO Data Path #1: Direct

As illustrated in Figure 2, Vehicle to IOO Data Path #1 includes vehicles equipped with on-board units (OBUs) broadcasting data using the 5.9 GHz spectrum that is received by IOO deployed RSUs. There is no 'handshaking' or confirmation of data communications, these are broadcasts of data that may be received by any device within range of the broadcast.

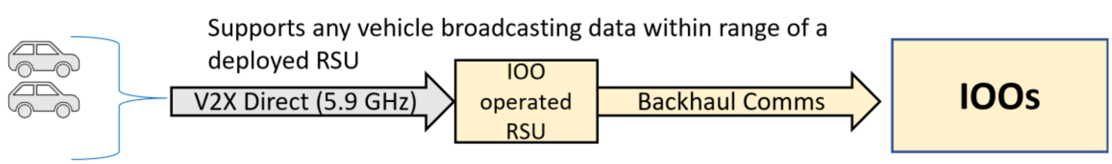


Figure 2: Illustration of vehicle to IOO data Path #1: V2X direct CV data communicated from vehicle to IOO

From the CV perspective, the in-vehicle radios that broadcast data are connected to the in-vehicle application(s) that generate data into messages. From the IOO perspective, the RSUs that receive data may either communicate the data (as received) through backhaul communications to a central location or implement some form of mobile edge computing (MEC) to either process and integrate the data into systems at the roadside or communicate the processed data to a central location through backhaul communications. Therefore, MEC is an enabling technology approach that can support the backhaul communications needed by direct CV data paths but is not part of the direct communications.

Advantages of this approach include the fact that the broadcast nature of data exchange allows any CV with an OBU to send data to be received by any radio within range, and that the same OBUs that perform V2V communications will support vehicle-to-everything (V2X) Direct communications.

> **Ohio DOT Example**
> The DriveOhio Communications Master Plan Appendix B[2] includes communications requirements for over 100 CV applications, and notes when a wired backhaul connection would enhance the application.

A disadvantage of this approach is the expense to deploy and operate each RSU at locations where data is to be collected, together with the need for backhaul communications. However, when RSUs are deployed at locations with existing traffic signals or ITS equipment, backhaul communications may already exist or be planned, and the cost for communications will likely be negligible. Additionally, required backhaul capabilities for direct CV data may differ, depending on the applications being supported at a specific location. As such, other factors should be considered beyond the availability of backhaul communications at a specific site. The DriveOhio Communications Master Plan describes when backhaul communications are required or would enhance a CV application.[2]

2.2.1.2 Vehicle to IOO Data Path #2: Private Vendors and Aggregation Portals

Vehicle to IOO Data Path #2 involves private vendors and aggregation portals communicating over network cellular and is referred to in this document as 'indirect' data communication. While the interviews in this study identified many examples of CV data being exchanged using network cellular communications, each example identified in the member interviews included some form of private vendor, aggregation service, or intermediary system. The most common example is network cellular communication of vehicle data to a central server operated by an OEM or a private vendor. Vehicle to IOO Data Path #2 is illustrated in Figure 3 below.

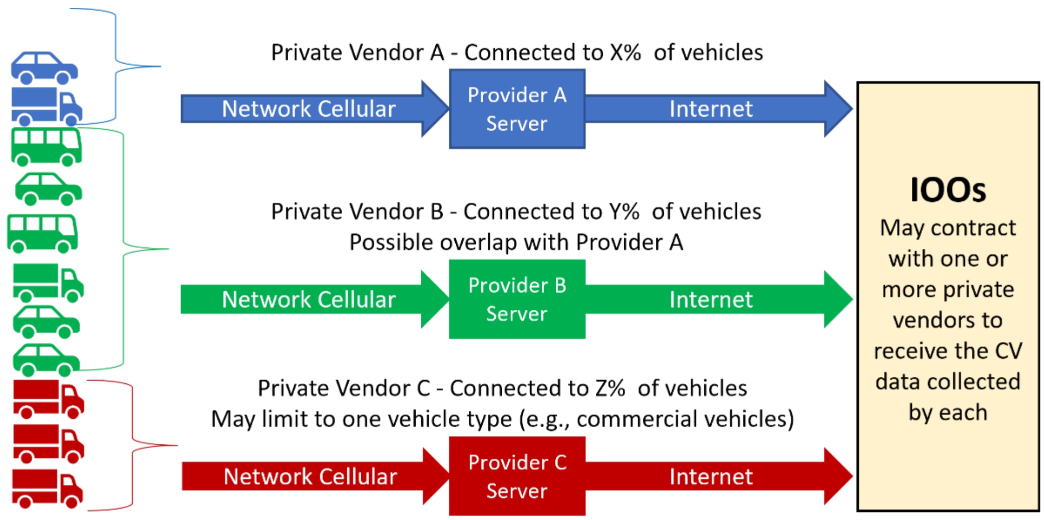


Figure 3: Illustration of Vehicle to IOO Data Path #2: Private Vendor and Aggregation Portals

From the IOO perspective, IOOs receive CV data by the private vendor applications or aggregation portals through internet communications.

Advantages to Data Path #2 include the fact that the private vendors are able to collect data wherever data communication over network cellular is available and therefore not limited to locations where RSUs are deployed to receive data. Another advantage is the large number of connected cars that send data using network cellular, estimated by Statista to be 84 million cars in 2021.[a] For these reasons, private vendors have been able to establish a market for CV data and have introduced IOOs to the option of purchasing CV data.

Disadvantages include:

- Because data is not broadcast direct from vehicle to IOO, no examples of Data Path #2 accomplishing low latency (i.e., less than 300 ms latency) or real-time data exchanges (i.e., less than 5 second latency) were cited in the interviews. However, examples of near real-time data exchanges (i.e., less than 5 minutes) were noted.

[a] https://www.statista.com/statistics/1155517/global-connected-car-fleet-by-market/

- As illustrated in Figure 3 above, IOOs would only have access to the CV data from the private vendor(s) they subscribe to and may require separate data connections to various CV data providers. IOO input in this project discussed the fact that cellular MEC does not currently provide cross provider communication, but it may in the future. If introduced, this cellular MEC option may incur additional latency between vehicles or users that use different cellular providers.
- Due to the need for reliable cellular communications, this data path is not possible (or is limited) for remote, rural locations where reliable cellular communications are not available.

2.2.1.3 Vehicle to IOO Data Path #3: IOO Internal Application

Vehicle to IOO Data Path #3 includes situations when IOO vehicles communicate directly to IOO operated applications. Examples of this are data communications with freeway assist vehicles, snowplows, or other maintenance vehicles. The communications may use network cellular, or an agency operated radio communications. See Figure 4.

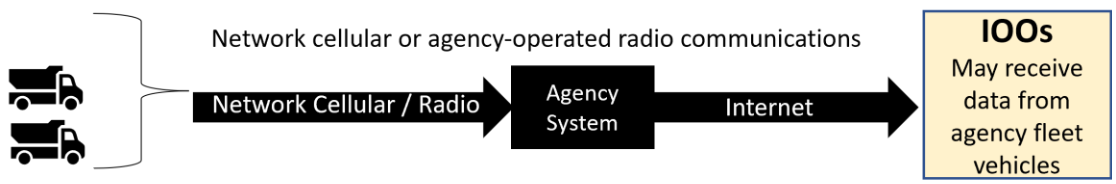


Figure 4: Illustration of vehicle to IOO Data Path #3: IOO internal CV data communication from fleet vehicles

This data path has generally evolved from earlier approaches where IOO fleet vehicle operators communicated using voice communications over radio. In current systems, a combination of voice and data are communicated, depending upon on-board applications. Examples of the data communicated from vehicles includes snowplow position, application rates of chemicals, vehicle position, and various data entered manually by vehicle operators to describe their actions or observations. The data that are received are commonly integrated into dispatching software or decision support systems.

The data communicated through this data path typically have very specific purposes or roles to support the agency (e.g., monitoring snowplow progress and chemical treatment rates) and therefore is different from receiving data from general vehicles traveling on the roads. This has both advantages and disadvantages.

One advantage of this approach is that IOOs have control over both the on-board applications sending the data and the system receiving the data. For this reason, IOOs can install specific sensors or monitoring equipment and configure the communications to send this data at the frequency desired. During winter weather events, the tracking of the snowplows allows for display using traveler information systems and for DOT performance monitoring.

A disadvantage of this data path is that the number of vehicles is limited to IOO equipped vehicles. This is not a disadvantage for the intended use, but it does limit additional uses of the data for other purposes.

2.2.2 IOO to Vehicle - Current CV Data Communications Approaches

While Section 2.2.1 discussed data paths from vehicles to the IOO, this section discusses data paths from the IOOs to the vehicles. The current approaches used to communicate data from IOOs to vehicles can also be decomposed into three different paths, described below as: direct, API/feeds, and IOO internal applications.

2.2.2.1 IOO to Vehicle Data Path #1: Direct

IOO to Vehicle Data Path #1 is direct communications. As illustrated in Figure 5, IOOs communicate data to RSUs and the RSUs broadcast the data as standardized messages using the 5.9 GHz spectrum for any OBU within range to receive.

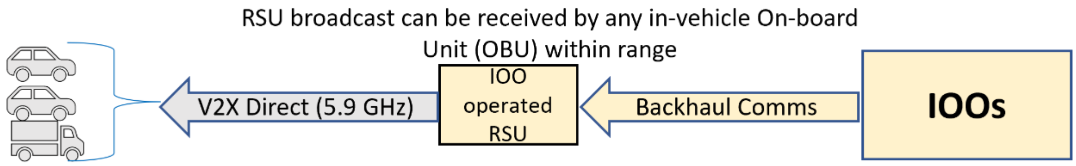


Figure 5: Illustration of IOO to vehicle data Path #1: V2X direct communications of CV data to vehicles

There is no 'handshaking' or confirmation of data communications, these are broadcasts of data that may be received by any device within range of the broadcast. IOOs may use direct communication to send data collected or generated at the site of the RSU (e.g., signal phase and timing (SPaT) data from a local signal controller) or may generate data centrally and communicate it to one or more RSUs for broadcast (e.g., work zone data). Direct data communications typically involve standardized messages. For example, SPaT data broadcasts by RSU is communicated using the SAE J2735 standard message. This enables standardization in the OBUs receiving the messages and processing the messages into usable data.

MEC may perform a role in data processing at or near the RSU to process data and/or formulate needed messages for broadcast by the RSU. Therefore, MEC is an enabling technology approach that can support the data assembly, backhaul communications, or message creation needed by direct CV data paths, but is not part of the direct communications.

Advantages of this approach include the fact that the broadcast nature of data exchange allows any CV with an OBU to receive data from the RSU when in range, and the latency and reliability of the broadcasts allow it to support safety applications that require less than 300 ms latency. Additionally, there is a defined industry approach for security credentialling to allow vehicles that receive direct communications to trust that the data is from the IOO and to prevent unauthorized broadcasts of infrastructure data.

A disadvantage of this approach is that RSU deployments are expensive, and coverage is limited to where they are deployed (with approximately 300 meters of range) and connected by backhaul communications. A second disadvantage is that vehicles must be equipped with OBUs to receive the data broadcasts.

2.2.2.2 IOO to Vehicle Data Path #2: API/Feeds

IOO to Vehicle Data Path #2 is referred to as API/Feeds. In this data path, IOOs make their data available over internet-based APIs or standard internet feeds (API/Feeds) either as standard messages or data. Private vendors or OEMs pull the data from the APIs or Internet feeds, and then communicate them to vehicles wirelessly, typically using network cellular communications. See Figure 6 below.

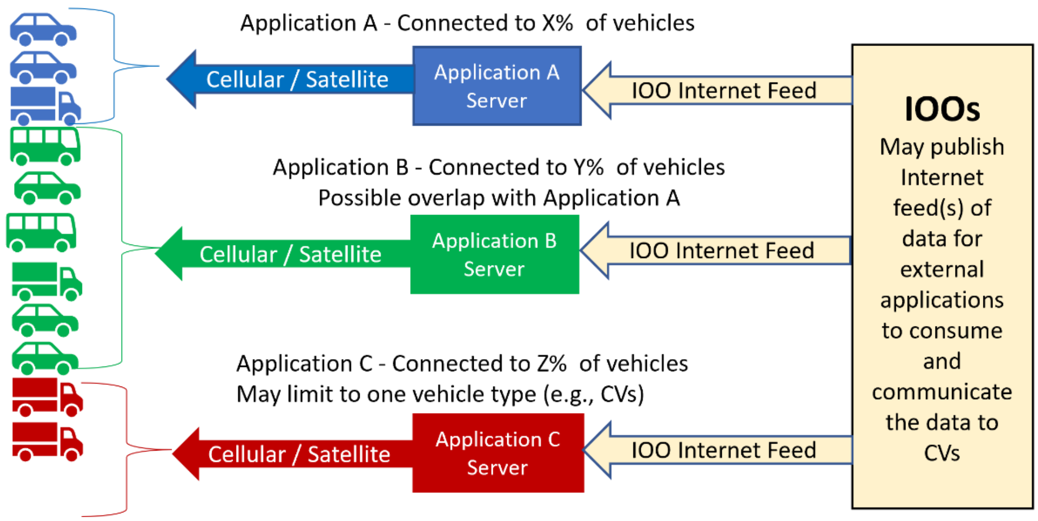


Figure 6: Illustration of IOO to Vehicle Data Path #2: API/Feed CV Data Communications from IOO to Vehicles

Examples of this data path include publishing SPaT data from central signal systems to the internet to support OEMs and private vendors communicating data to on-board applications. Note that this SPaT data is not likely to be in the SAE J2735 format as when it is when communicated in IOO to Vehicle Data Path #1 as bandwidth is not as crucial in this data path.

Other examples include IOOs publishing data describing work zones (e.g., using WZDx) or crashes/events, and private vendor or OEMs pulling these from internet access points and disseminating them to vehicles they connect with.

A variation on Data Path #2 was described by Arizona DOT in their pilot deployment of "virtual RSUs." In this system, Arizona DOT communicated data to the cellular provider (Verizon) and the data was communicated to vehicles based on their location (simulating a vehicle passing an RSU and receiving data pertinent to that location).

The CV PFS conducted a project specifically to address IOO communication of CV data through third parties. One project outcome was the recommended approach of a Connected Vehicle Data Framework (CVDF) API. The project outcomes, including detailed requirements for the CVDF are available at [Using Third Parties to Deliver Infrastructure-to-Vehicle (I2V)].

Advantages to this approach are that centralized internet-based communications from IOOs are less expensive than deploying and operating RSUs to cover the road network. Also, this enables IOOs to publish data to one location and enable all private vendor applications and OEMs to consume the data and communicate the raw or processed data to their recipients.

Disadvantages to this approach include:

- The additional latency introduced by the private vendor or OEM pulling/receiving data from the IOO internet feed and then communicating it to vehicles.
- Currently there is not consistency or universal geographic coverage of the same data being provided by all IOOs, thus creating constraints on either OEMs or the private vendors that wish to provide IOO generated data to the CVs they are connected to.
- Travelers using one cellular service plan may receive some alerts that customers of other cellular service plans from another provider may not receive.
- In rural locations where reliable cellular communications are not available or are limited, this data path may have technical challenges that make it not a viable option until cellular communications improve.

2.2.2.3 IOO to Vehicle Data Path #3: IOO Internal Application

In IOO to Vehicle Data Path #3, the IOOs rely on local communications with their internal systems that communicate to their own vehicles. The communications may use network cellular or agency operated radio communications. See Figure 7.

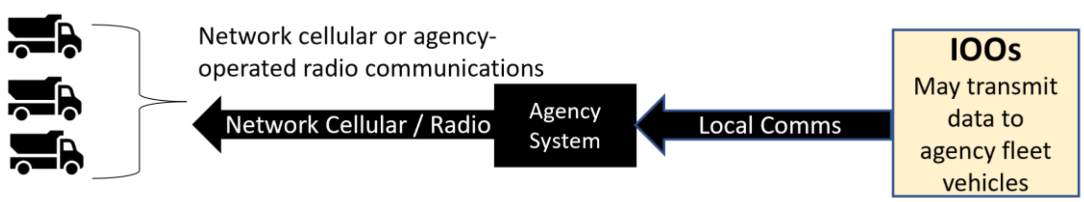


Figure 7: Illustration of IOO to vehicle data path #3: an IOO internal CV data communication to fleet vehicles

The advantages and disadvantages to this data path are similar to the Vehicle to IOO data path described in Section 2.2.1.3. The IOOs have control over both the system sending data and the on-board applications receiving the data.

A disadvantage of this data path is that the number of vehicles is limited to IOO equipped vehicles. This is not a disadvantage for the intended use, but it does limit the additional uses of the data for other purposes.

2.3 Emerging CV Data Communications Concepts

There are two emerging concepts related to CV data communications identified during the research and interviews conducted for this project (i.e., digital infrastructure and probe data management subsystems) that may impact and/or provide alternatives to the approaches identified above. These are described in this document as background and included in the concepts described below.

2.3.1 Emerging Concept #1: National Strategy for Roadway Digital Infrastructure

While a universally recognized definition of digital infrastructure as it relates to CV data has not been formally standardized, a working definition within an FHWA led initiative to evolve a national strategy for roadway digital infrastructure was shared in March 2023 as *"Roadway Digital Infrastructure (RDI) are the collective public and private technology assets that create, exchange or use data or information to improve the transportation system by the provision of existing and new services for travelers, businesses and agencies."* Building on this tentative definition, IOOs are already managing digital infrastructures today, purchasing data, governing their own sources of data, and providing them to others. Private sector CV vendors have their own digital infrastructures that collect, communicate, manage, and deliver CV data to their customers. Data service providers and utilities that communicate, store, process, and aggregate data are parts of multiple digital infrastructures. Related to content in this document, digital infrastructures include a combination of edge computing, cloud storage and services, V2X direct communications, network cellular communications, and data management systems to assemble, aggregate, and communicate data between IOOs and CVs. It is likely that RDI includes the different communications technology approaches identified previously in this document.

In the future, the scope and complexity of digital infrastructures required to manage the data deluge is expected grow and are likely to be created, both by public and private sectors. The rising challenge has become not a lack of data but a question of how to best utilize all these data, data services, and enabling utilities in the most efficient manner possible to ensure data governance, quality, availability, and open sharing. An emerging concept that is likely to influence CV data is FHWA's efforts initiated in 2023 to develop a national strategy for roadway digital infrastructure. IOOs are encouraged to participate in and contribute to this national strategy.

> Roadway Digital Infrastructure (RDI) are the collective public and private technology assets that create, exchange, or use data or information to improve the transportation system by the provision of existing and new services for travelers, businesses, and agencies.

2.3.2 Emerging Concept #2: Probe Data Systems

In June 2022, SAE released the standard titled SAE J2945/C_202206 "Requirements for Probe Data Collection Applications."[3] This standard defines the communications interface for vehicles participating as probe vehicles to communicate data to a probe data management system (PDMS) either using the cellular network or 5.9 GHz C-V2X communications. In the use cases presented in this standard, IOOs would periodically communicate probe data configuration requests to the vehicles requesting specific probe data, and vehicles would use the J2945/C standard to reply and communicate the probe data to the PDMS. The use of IOO generated requests for probe vehicle data is a different approach than the concept of CVs broadcasting Basic Safety Messages (BSMs) at a rate of 10 Hz. Instead, this affords the IOO the opportunity to request specific data from vehicles in specific areas (or for all areas).

For example, IOOs using a PDMS could:

- Identify a geo-fence area around work zones and request all vehicles within the work zone to send their speed data.
- Identify a request for vehicles on bridges or selected stretches of road to send data describing roadway friction or anti-lock brake activations when temperatures are 31-33 degrees (and there is uncertainty about whether roads are icy).

Vehicles that receive the data request could respond if appropriate. Figure 8 illustrates a conceptual example of how a PDMS could be used for IOOs to request and receive data from vehicles through both network cellular and C-V2X communications.

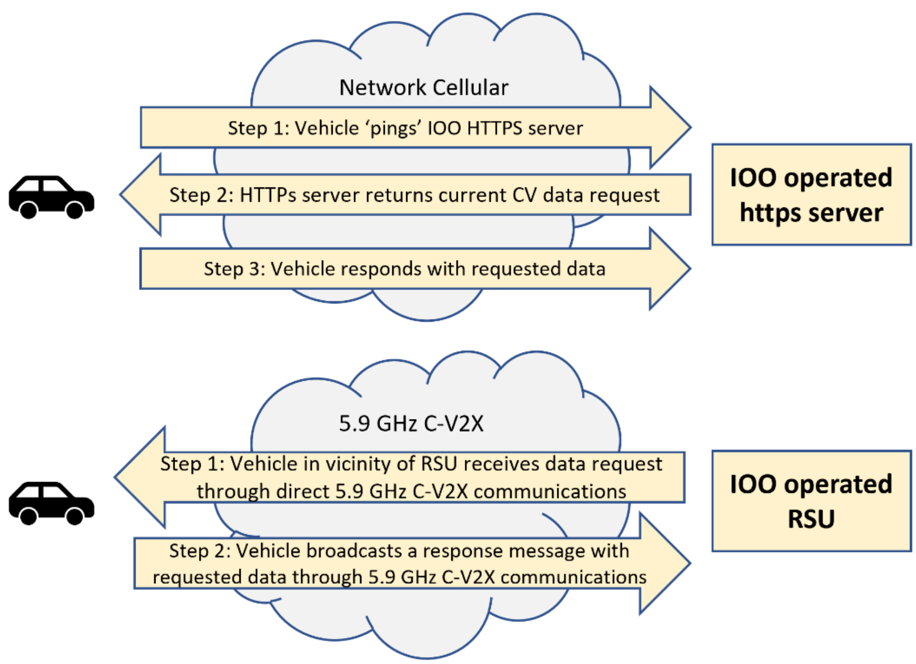


Figure 8: Conceptual illustration of the use of SAE probe data standard

This standard would support two-way communications between IOOs and vehicles using either direct (i.e., 5.9 GHz C-V2X) or indirect (i.e., network cellular) communications and would not require intermediary systems. This concept was initially defined in the Event Driven Configurable Messaging (EDCM) and Queue Advisory / Queue Warning (QA/QW) systems described in earlier efforts of the CV PFS[4] and Federal Highway Administration (FHWA).[5]

2.3.3 Emerging Concept #3: Emerging Standards

The United States Department of Transportation (USDOT) development of the WZDx data feed has been very successful at defining a format for exchanging work zone data. A consortium led by ITE and SAE are progressing to develop a standard based on the WZDx. Building upon this model, USDOT's Joint Program

Office (JPO) and FHWA are also collaborating to develop a Transportation Incident Management Data Exchange (TIMDx) with the focus on incident management and response. These emerging standards, as well as any others to support national standardization of IOO data such as road weather or incident data, will support the communication of IOO generated information to vehicles as CV data. Similarly, the CV data described in this report will likely serve as sources of data to populate TIMDx messages.

2.3.4 Potential Impacts of Emerging Concepts

The potential impacts of these emerging concepts are described here as they will relate to the requirements and concepts presented below.

Potential Impact #1: Open Connectivity to All CVs. The SAE J2945/C standard might enable the flow of data from all CVs to IOOs using a probe data system, either delivered by private vendors or directly to/from the IOOs. This could allow IOOs to benefit from data collected by all CVs in their area.

Potential Impact #2: Increased IOO Data for CVs. The WZDx and TIMDx standards, as well as the national strategy for roadway digital infrastructure, might be catalysts that help to enable IOO provided data in North America that is as consistent and comprehensive as possible, enabling OEMs and private vendors to build applications that rely on ubiquitous coverage and consistency.

Potential Impact #3: More Data Driven Decision Making for IOOs. The national strategy for digital infrastructure might be a catalyst that improves data governance and consistency of all CV data, potentially enabling IOOs to better use CV data to make more efficient and data-driven decisions that benefit all facets of their agency, including planning, operations, and maintenance. As part of this increased decision-making, future planning by IOOs should include IOOs exploring the costs of deploying, supporting, and maintaining the digital infrastructure necessary to provide stable CV data communications. Regardless of the overall approach towards the digital infrastructure (e.g., purchase data, partner to receive data, deploy systems to receive data), digital infrastructure investments are expected to require a significant change in investment strategy and cultural change.

3.0 Model CV Data Architecture Systems Integration Guidance

The goal of this Systems Integration Guidance is to help readers of this CV Data Architecture make use of the content defined in the Model CV Data Architecture in Chapter 4 by providing steps to follow when considering local deployments, both for receiving and provisioning CV data.

3.1 IOO Ingest of CV Data

The content in this section supports decision-making for IOOs as they develop their approach towards ingesting CV data from vehicles. This section is structured according to four Ingest "I" steps described in Figure 9.

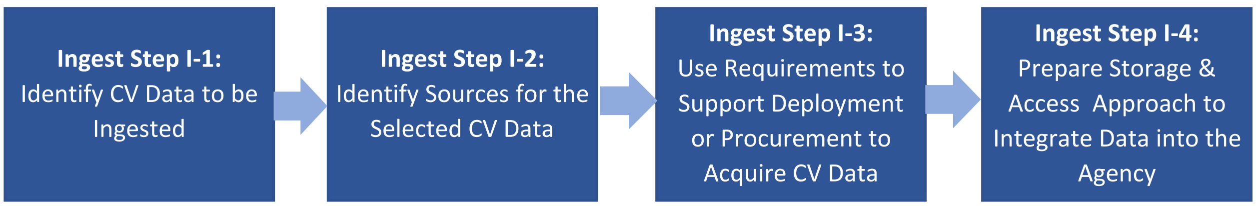


Figure 9: "I" Steps of this systems integration guidance for IOOs to prepare to "Ingest" CV data

3.1.1 Ingest Step I-1: Identify CV Data to be Ingested

Agency practitioners are recommended to follow a need-based approach to identify CV data to be ingested. By first identifying agency data challenges and needs, potential CV data types can be pursued that would be the best use of agency investments and resources.

Ingest Step I-1 Guidance:

- *Identify applicable IOO challenges, data needs, and CV data to potentially address those needs*. Refer to Section 4.1.2 of this document and Table 3 (on page 27 of this document) for common IOO data challenges, data needs, and candidate CV data. Select one or more data from Column 3 "potential CV data Needed by IOOs".
- *Determine the CV Data Category(ies)*. Use Table 4 (on page 29 of this document) for each CV data need identified above to identify the Data Category(ies), which will be the basis for identifying relevant requirements In future steps.

3.1.2 Ingest Step I-2: Identify CV Data Sources for the Selected CV Data

Once an agency has identified one or more needed CV data they wish to ingest and has determined the CV Data Categories that are related to the data, the intent of Step I-2 is to assist in identifying which of the CV data sources may meet the data needs.

Ingest Step I-2 Guidance:

- *Inquire about existing CV data sources that may already by procured by your agency*. It is possible that other groups at your agency (e.g., planning, maintenance, or operations) may already purchase CV data that is also available to other agency groups. Similarly, your agency may have a subscription to a data portal that enables access to CV data or relationships with local universities with access to CV data. Existing data sources may eliminate your need for new CV data.
- *Determine whether direct or indirect data ingests are possible.* Using the Data Category(ies) identified in Step I-1, refer to the data category descriptions on Table 4 on page 29, to understand if direct or indirect/vendor data sources are options for this ingest.
- *Evaluate direct data sources before selecting a direct approach.* If Table 4 identified a direct data source as an option, consider the percentage of vehicles sharing CV data directly (i.e., broadcasting CV data or responding to requests for probe data messages) before confirming direct as the selected approach. Alternately, assess whether local fleet vehicles are sending CV data directly to meet data needs, as it may be possible that the number of public fleet vehicles could be sufficient to meet data needs.
- *Evaluate vendor data sources before selecting to pursue an indirect data source.* If Table 4 identified indirect/vendor data sources, review the use cases for communications of CV data in Section 4.2.1 for examples of IOOs successfully procuring and using private vendor data or aggregation services, with special attention to bullets labelled **"Vendor/Aggregation."** Note that limited examples of costs are included. Also, review Sections 4.6.1.1 and 4.6.1.2 to better understand private vendor data sources.

3.1.3 Ingest Step I-3: Use Requirements to Support Deployment of Systems or Procurement of Services to Acquire CV Data

After completing Step I-2, your agency should have determined that the CV data needs can either be met by deploying infrastructure to receive CV data directly or by procuring private data services. **One of the key benefits of this document is that the use cases, requirements, and concepts included in Chapter 4 can be used to support deployment or procurement activities**.

Ingest Step I-3 Guidance:

- *Review use cases.* Review the uses cases for access, privacy, security, and communications in Sections 4.2.1, 4.3.1, 4.4.1, and 4.5.1 for the Data Category(ies) being considered. The use cases include insights into how other IOOs are accomplishing and benefitting from the ingest of CV data.
- *Consider costs of CV data ingest.* It is important to understand cost estimates for initiation and ongoing operations of CV data ingest, whether it is vendor data or infrastructure to receive direct data. The use cases described in Sections 4.2.1, 4.3.1, 4.4.1, and 4.5.1 also include examples of costs shared by agencies interviewed in this project for both direct and indirect data ingest. Communicate with other IOOs to gather the most current cost estimates for direct and indirect data ingests.
- *Understand requirements to consider.* The tables in Sections 4.2-4.5 organize requirements by Data Category and communication approach (Direct or Vendor). Requirements are found in Section 4.2.2 for data communications, Section 4.4.2 for privacy, and Section 4.5.2 for security. Each requirements table is structured such that:
    - Requirements are arranged in rows indicating what Data Category(ies) it relates to.
    - Each requirement begins with either **"Direct"** or **"Vendor/Aggregation"** to identify the communication approach.
    - Some tables have an initial row with general requirements applicable to all Data Categories.
- *Use the Concepts*. Concepts for each topic area (presented in Sections 4.2.3, 4.3.3, 4.4.3, 4.5.3) relate to the selected Data Category and communication approach.

3.1.4 Ingest Step I-4: Prepare Storage Approach to Integrate Data into the Agency

Finally, in Step I-4 of this systems integration guidance, practitioners should consider specific requirements and concepts when developing a CV data storage approach.

Ingest Step I-4 Guidance:

- *Understand access and storage.* Section 4.3 has insights for storing the data and providing access to agency staff, and includes use cases, requirements, and concepts that begin with either **"Direct"** or **"Private/Aggregation."** The requirements table is divided into rows for each Data Category.
- *Identify requirements and supporting concepts.* Identify these in Section 4.3.2 and Section 4.3.3, respectively, based on the Data Category you are considering and the communication approach your agency is pursuing.

3.2 IOO Provision of CV Data

The content in this section supports decision-making regarding what CV data your agency might output or provision for access by CVs. This section is structured into three “O” steps shown in Figure 10 for agencies to “output” data.

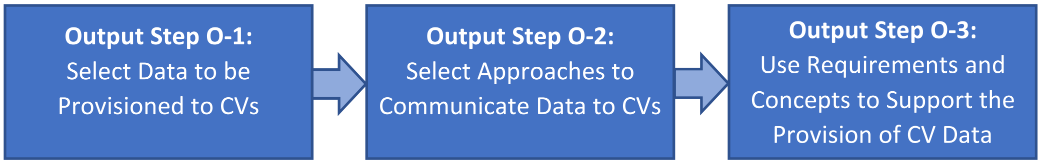


Figure 10: “O” Steps of this systems integration guidance for an IOO to provide CV data “Out” to vehicles

3.2.1 Output Step O-1: Select Data to be Provisioned to CVs

Agencies maintain an extensive array of data, and it is important to identify what datasets are most relevant and useful for CVs. This will help to prioritize agency resources necessary to prepare those datasets and any systems, interfaces, or infrastructure to support communication of that data from the IOO to CVs, either directly or via a private vendor. Depending on the data identified, part of this step may include the identification of locations to deploy infrastructure for direct V2X broadcasts or additional resources for processing data into accepted industry data standards and specifications or to ensure data quality, for instance.

Output Step O-1 Guidance:

- *Consider national deployment strategies and CV applications.* There are three national actions that readers of this document are recommended to pursue when provisioning CV data. These include:
    - Work zone data provisioned over API/feed communications using the WZDx; and
    - Connected Intersections (CI) deployments following the Connected Transportation Interoperability (CTI) family of standards.
    - The Traffic Incident Management Data Exchange (TIMDx) described in Section 2.3.

- *Familiarize with vehicle/driver challenges, needs and CV data to address needs.* In addition to the two national approaches identified above, Table 2 of this report identifies challenges that drivers of vehicles face, and CV data needs to address these challenges, as well as potential CV data needed. IOOs should review this table when making decisions about provisioning data.

- *Identify CV data output priorities for your agency.* Using the outcomes of the previous bullets, determine priorities for CV data outputs. Agencies are encouraged to balance agency and industry priorities to provide the most benefit. In other words, there is little value in providing data that will not be consumed by CVs. It may also be beneficial to prioritize the provision of CV datasets that align with agency priorities (e.g., if an agency strategic plan emphasizes work zone safety due to high work zone crash rates, work zone data may be an early priority) in order to dedicate agency funding and resources more easily.

- *Determine applicable CV Data Category(ies).* Use Table 4 for each CV data prioritized above from Table 1 to determine what CV Data Category(ies) to use in future steps to identify relevant requirements for providing the CV data.

3.2.2 Output Step O-2: Select Communications Approach(es)

Agencies should provision CV data to CVs using communications that meet industry requirements and that are consistent with how the data will be used by CVs. This step will support IOOs in determining the options for communications and informing them of resources to select the best approach(es).

Output Step O-2 Guidance:

- *Understand communications options.* Refer to Table 4 to understand which communications options are possible for each data category identified in Step O-1.
- Review Section 2.2.2 to understand more context around each communications approach for provisioning IOO CV data (i.e., direct V2X broadcast, API/feeds).
- *Evaluate vehicle readiness to receive direct data communications before selecting a direct approach.* If Table 4 identified direct data communications as an option for your data category(ies), consider the percentage of vehicles equipped to receive CV data directly (i.e., with OBUs or on-board applications able to receive direct data communications) before confirming direct as the selected approach. Alternately, assess whether local fleet vehicles (e.g., transit, snowplows, maintenance vehicles) are equipped to receive CV data directly to meet data needs, as it may be possible that the receipt by fleet vehicles could be sufficient to justify direct broadcasts.
- *Evaluate vendor or OEM use of IOO provisioned data before selecting to pursue an Indirect data approach.* If Table 4 identified indirect communications using an API/feed, review the use cases for communications of CV data in Section 4.2.1 for examples of IOOs successfully outputting data to support CVs, with special attention to bullets labelled **"API/feeds."** Also, review Section 4.6.2 to better understand IOO to vehicle data sources.

3.2.3 Output Step O-3: Use Requirements to Support the Provision of CV Data

Step O-3 describes how the requirements, use cases, and concepts included in this document can be used to support the provision of selected CV data.

Output Step O-3 Guidance:

- *Review use cases.* Review the uses cases for access, privacy, security, and communications in Sections 4.2.1, 4.3.1, 4.4.1, and 4.5.1 for the Data Category(ies) being considered. The use cases include insights into how other IOOs are accomplishing the output of CV data.
- *Consider costs of CV data output.* It is important to understand cost estimates for all aspects of CV data output. The use cases described in Sections 4.2.1, 4.3.1, 4.4.1, and 4.5.1 also include examples of costs shared by agencies interviewed in this project, primarily for direct data output.
- *Identify appropriate requirements.* Select requirements for access, privacy, security, and communications in Sections 4.2.2, 4.3.2, 4.4.2, and 4.5.2 for the Data Category being considered and selected communications approach for provisioning IOO CV data.
- *Use the Concepts*. Concepts for each topic area (presented in Sections 4.2.3, 4.3.3, 4.4.3, 4.5.3) relate to the selected Data Category and communication approach.

4.0 Model Connected Vehicle Data Architecture

This chapter presents the contents of the CV PFS Model Data Architecture based on the research and interviews conducted within this project. The model architecture can be thought of as three parts:

Part 1: Definition of data needs and uses for CV Data.

Section 4.1 presents challenges and needs of drivers and IOOs that CV data can address. This section defines a list of CV data that is currently or planned to be communicated and groups these data into categories that are used in the remainder of the model Data Architecture.

Part 2: Topic-Specific Requirements IOOs Should Consider for Their Local CV Data Architecture

Sections 4.2 – 4.5 address the four core topic areas of the model data architecture (Section 4.2 addresses communications, Section 4.3 addresses access and storage, Section 4.4 addresses privacy, and Section 4.5 addresses security).

The process followed within each topic area was to document use cases, derive needs and requirements from these use cases (and other inputs), and finally to develop concepts to consider when addressing the requirements. This process and structure used for these sections is illustrated in Figure 11.

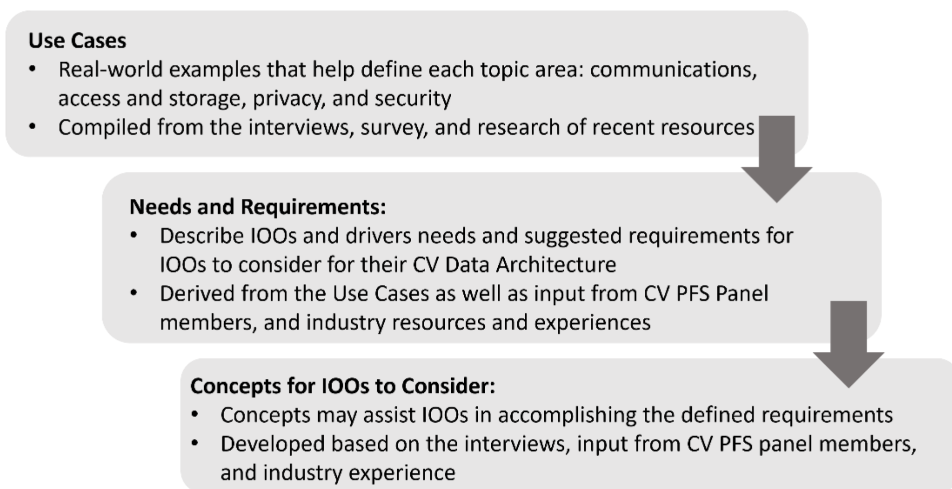


Figure 11: CV process and structure of topics addressed in this Model CV Data Architecture

Part 3: CV Data Sources

The final component of the CV data architecture (Section 4.6) discusses the extent to which the requirements and concepts are addressed by current and potentially future CV data sources. Specifically:

- The use cases, needs and requirements, and concepts for CV data that IOOs can ingest are presented separately for "**Direct**" and "**Vendors/Aggregation**" given the differences between how these data may be communicated, ingested, processed, used, and managed.
- The use cases, needs and requirements, and concepts for CV data that IOOs can share with vehicles are presented separately for "**Direct**" and "**API/Feeds**" given the differences of each approach.

4.1 Data Needs and Uses

The need for data is underwritten by the transportation industry's systemic need to improve safety, mobility, and efficiency. In 2022, traffic crashes are a safety risk; travel delays cause inefficient uses of time for commuters, deliveries, and leisure travel; and the impacts of fuel consumption on the environment are not sustainable. In order for IOOs using this model CV data architecture to consider a "needs based" approach to managing CV data, this section describes the challenges and needs of drivers and IOOs that are most likely to be addressed by CV data. The data needed to address these challenges will form the basis for later sections of this architecture (See Sections 4.2 – 4.5).

4.1.1 Drivers Challenges, Data Needs, and Types of Data Needed

Currently, drivers observe and process a massive amount of visual and audio input as they drive, in the form of roadway signs, traffic signals, position and movement of other vehicles and pedestrians, horns, sirens, in-vehicle sensor alerts, and others. These human observations allow drivers to travel at today's safety and efficiency levels. However, traffic crashes and congestion are still occurring. Therefore, there are still challenges and needs from drivers for improvements that can be addressed by CVs. Table 2 identifies the highest-level challenges, data needs, and the type of data needed to address some of the most documented driver challenges.

Table 2: Driver Challenges, CV Data Needs of In-Vehicle Applications, and Data Needed to Address the Challenge

Driver Challenges	Data Needs of In-vehicle Applications to Address Driver Challenges	Potential CV Data Needed by Vehicles
Safety at signalized intersections	Low latency, secured, reliable access to infrastructure data to support red light violation warning (RLVW) applications[7]	• Low latency SPaT, MAP, Radio Technical Commission for Maritime Services (RTCM)
Safety during inclement weather or conditions	Inclement Driving Condition Data to support Spot Weather Information Warning (SWIW) applications[6]	• Location specific inclement driving conditions (including spot specific hazards such as ice) • Road weather conditions • Advisory speeds • Road closure notifications
Safety when approaching work zones, lane closures, and reduced speeds	Information describing locations and impacts of work zones to support Reduced Speed Zone Warning (RSZW) and Lane Closure (LC) Applications[7]	• Work zone data (e.g., lane closures information, workers present, posted speed limit, Advisory speed, traffic queues) • Current speed limits in locations of variable speed limits
Safety when approaching curves	Information required to determine driver warnings[7]	• Posted speed limit • Advisory speed • Road geometry • Roadway material information

Driver Challenges	Data Needs of In-vehicle Applications to Address Driver Challenges	Potential CV Data Needed by Vehicles
Fuel efficiency and travel delays	Access to network-wide signal status data	• Real-time network-wide signal phase and timing, MAP messages • Road closures and detours
Transit or Snowplow stops at intersections	Low-cost signal priority / preemption approach	• Real-time signal preemption status
Conditions and restrictions impacting commercial vehicles	Information about temporary vehicle restrictions Information about parking and available services	• Commercial vehicle restrictions (width, height, weight) • Parking/services information for commercial vehicles

4.1.2 IOOs Challenges, Data Needs, and Types of Data Needed

IOOs rely on data as they plan, build, maintain, and operate the transportation network. Data collection devices, driver surveys, and driving studies were early mechanisms for data collection and are still used today. As technologies have increased, IOOs have been able to collect more data at a lower cost, but funding is a constraint on the amount of data collected. CV data offers two distinct advantages to address IOOs' data challenges:

- CVs may offer data not previously possible (e.g., vehicle specific sensor data); and
- CVs may offer extended coverage and/or reduced costs of data collection as compared to non-CV data collection approaches (e.g., speed data available network-wide vs. only at locations of sensors).

Table 3 identifies examples of challenges and needs for data that IOOs experience and the potential CV data to address the challenges. The BSM is an SAE standard CV data message that includes a series of data elements intended to identify the required data to be broadcast by vehicles participating in V2V data communications. The data elements identified below that are part of the BSM Core Data (aka Part 1) are denoted with an asterisk.

Table 3: IOOs Challenges, Data Needs, and Potential CV Data to Address the Challenge

IOO Challenge	IOO Data Needs	Potential CV Data Needed by IOOs
Limited situational awareness of current travel speeds or locations of queues	Low-cost data to support operations	• Near real-time vehicle speeds* • Near real-time average speeds • Near real-time volumes • Near real-time segment travel times • Near real-time start and end of traffic queues

IOO Challenge	IOO Data Needs	Potential CV Data Needed by IOOs
Limited information about events or incidents prevents analyses and planning for future responses	Low-cost data to support event analyses and planning activities	• Vehicle speeds* • Vehicle heading* • Vehicle acceleration* • Average speeds • Traffic volumes • Segment travel times • Start and end of traffic queues • Trip trajectories
Limited information about current road and weather conditions	Low-cost data options to support real-time maintenance and operations	• Near-real-time spot specific atmospheric condition • Near-real-time spot specific pavement friction
Limited information about inclement weather start/end times and recovery times to clear conditions	Low-cost data options to support maintenance and operations performance management	• Spot specific atmospheric conditions • Spot specific pavement friction • Vehicle speeds • Average speeds
Situational awareness of areas prone to crashes	New data sources to understand areas prone to incidents/events	• Harsh braking occurrences • Vehicle brake system status*
Situational awareness of pavement conditions	Supplemental data to what is available now to support pavement management	• Abnormal pavement (e.g., potholes) • Pavement marking conditions
Situational awareness of missing or damaged roadway assets	Supplemental data to what is available now to support asset management	• Road sign conditions (or missing)
Transit or snowplow vehicle stops at intersections reduce efficiencies	Data to indicate when authorized vehicles will benefit from priority or preemption	• Real-time signal request messages from authorized vehicles when appropriate
Occurrences of vehicle collisions	Data to support vehicle collision and/or vulnerable road user (VRU) applications	• Vehicle position, speed, heading • Notifications of air-bag deployments

*** Data elements included in the SAE BSM. Full contents of the BSM include time (seconds), vehicle position, position accuracy, state of vehicle transmission, speed, heading, steering wheel angle, acceleration, brake system status, and vehicle size.**

4.1.3 CV Data Categories

Table 2 and Table 3 included over 40 examples of CV data to address the needs of drivers and IOOs. In order for this document to organize the needs, requirements, and concepts for these CV data, the CV data were grouped into categories based on commonalities of the data, and these categories will be used for the remainder of this document. Eight CV data categories were derived and are presented in Table 4 below. By grouping these data into categories, the intent is not to exclude or ignore any data, but rather to organize needs and requirements in a useful way. Some CV data may fit into multiple data categories.

Table 4: IOO and Driver Data Needs Grouped to Derive CV Data Categories

Data Category	Data Category Description
Data Category #1: Low Latency Individual Vehicle Data (Vehicle to Infrastructure) *Data include:* • *Vehicle position, speed, and heading* • *Notifications of air-bag deployments*	Data describing individual vehicle movements, activities, or data collected that may be used by IOOs to support safety applications or activities. Total latency from vehicle to IOO is defined here as less than 300 ms. Examples of the use of these data may include an IOO role in Intersection Movement Assist (IMA), and receiving and resending BSM data to other vehicles. These data would be communicated: • **Direct** from the *vehicle* to IOO infrastructure using the BSM.
Data Category #2: Immediate Individual Vehicle Data (Vehicle to Infrastructure) *Data include:* • *Real-time signal request messages from authorized vehicles when appropriate* • *Near real-time vehicle speeds* • *Near real-time spot specific atmospheric condition* • *Near real-time spot specific pavement friction*	Data describing individual vehicle movements, activities, data collected, or requests that are needed immediately to support IOO operations activities. Includes both: • Real-time data: Data communicated in 5 seconds or less from when generated (e.g., Signal Request Messages (SRMs) that need to be timely to support signal priority responses). • Near real-time data: Data communicated in 5 minutes or less from when generated (e.g., vehicle location, vehicle speed, pavement friction that might help formulate TIMs). These data may be communicated: • **Direct** from vehicle to the IOOs using a vehicle broadcast of the BSM or SRM; • **Indirect** via **Private** vendors. • **Indirect** from vehicle to IOOs using Hypertext Transfer Protocol (http) communications of the probe data message standard.
Data Category #3: Individual Vehicle Data (Vehicle to Infrastructure) *Data include:* • *Vehicle speeds* • *Vehicle acceleration* • *Spot specific atmospheric conditions* • *Spot specific pavement friction* • *Vehicle brake system status* • *Pavement marking conditions* • *Road sign conditions (or missing)*	Data describing individual vehicle movements and activities that are **not** needed or used by the IOO immediately. These data are typically used for after event analyses or to determine roadway improvements. These data would have a communication latency of more than 5 minutes, although this category includes data that is communicated within 5 minutes and stored by the IOO for later use. These data may be communicated: • **Direct** from CV to the IOO using the BSM or probe data messages; • **Indirect** via **Private** vendors; or • **Indirect** from vehicle to IOOs using http communications of the probe data message standard.

<table>
<tr><th>Data Category</th><th>Data Category Description</th></tr>
<tr><td>Data Category #4: Near real-time Aggregated Vehicle Data
(Vehicle to Infrastructure)

Data include:
• Near real-time average speeds
• Near real-time volumes
• Near real-time segment travel times
• Near real-time start and end of traffic queues
• Harsh braking occurrences</td><td>Data that are aggregated, typically by private vendors, to add value before delivery (e.g., to ensure privacy and make the data more usable for specific applications), and not communicated directly from CV to the IOO. These data are needed in near real-time, defined as within 5 minutes from when the data are generated, to support IOO operations. For example, speed data from vehicles over a segment of roadway may be aggregated to describe the segment travel time and communicated to IOOs for display on DMS.

These data may be communicated:
• Indirect via Private vendors.</td></tr>
<tr><td>Data Category #5: Aggregated Vehicle data
(Vehicle to Infrastructure)

Data include:
• Average speeds
• Traffic volumes
• Segment travel times
• Start and end of traffic queues
• Trip trajectories
• Abnormal pavement (e.g., potholes)
• Pavement marking conditions
• Road sign conditions (or missing)</td><td>Data that are aggregated, typically to add value, and are not needed or used by the IOO in real-time or near real-time. These data are typically used for planning or after event analyses. These data would have a communication latency of more than 5 minutes from when the data is generated. The aggregation of vehicle position data to define origin-destination (O-D) tables is an example of aggregated vehicle data available from some vendors.

These data may be communicated:
• Indirect via Private vendors.</td></tr>
<tr><td>Data Category #6: Low Latency Infrastructure Safety Data
(Infrastructure to Vehicle)

Data include:
• Low latency SPaT, MAP, RTCM</td><td>Data describing the current status of the infrastructure that needs to be delivered with low latency to support vehicle safety applications such as RLVW. Total latency from infrastructure to vehicles is defined here as less than 300 ms from when the data is generated.

These data would be communicated:
• Direct from IOO RSUs to the in-vehicle OBUs.</td></tr>
</table>

Data Category	Data Category Description
Data Category #7: Real-time Infrastructure Data (Infrastructure to Vehicle) *Data include:* • *Real-time network-wide signal phase and timing, MAP messages* • *Real-time signal preemption status* • *Location-specific inclement driving conditions (including spot specific hazards such as ice)*	Data describing the current status or conditions of the infrastructure that are intended for real-time use by in-vehicle applications, but do not need to be communicated with low latency. Total latency of communication would be less than 5 seconds from when the data is generated. These data may be communicated: • **Direct** from the IOO to the vehicle using RSUs broadcasting messages such as Signal Status Messages (SSM). • **Indirect** from the IOO to the vehicle if vehicles use network cellular to connect to the IOO APIs or data exchange to request and receive data. • **Indirectly** by IOOs making data available to an **API/Feed** for private vendors to download data and communicate the data to their in-vehicle applications.
Data Category #8: Infrastructure Data (Infrastructure to Vehicle) *Data include:* • *Road weather conditions* • *Advisory speed* • *Position and content of curve warning road sign* • *Road Geometry* • *Roadway material information* • *Road closure notifications* • *Work zone data* • *Current speed limits in locations of variable speed zones* • *Commercial vehicle restrictions* • *Parking/services information for commercial vehicles* • *Posted speed limit* • *Location specific inclement driving conditions (including spot specific hazards such as ice)*	Data describing the current status or conditions of the infrastructure that do not need to be communicated within 5 seconds from when the data are created. Final delivery of data to the vehicle may be on a low latency cycle of broadcast (e.g., data to support curve warning systems may be broadcast frequent enough to be received timely by approaching vehicles). Therefore, total latency of communication would vary based on the nature of the application and individual data elements (see more descriptions in the Communications Requirements section). These data may include manually entered data by IOO staff, automatically recorded data (e.g., road weather information system (RWIS) sensors), or data that IOOs receive from external sources, including CVs. There may be a time latency from when conditions or events start until the time the report is recorded by the IOO, therefore latency is described as the time from when it is available to the IOOs (not specifically from when it occurs). Data may be described using the WZDx specification, TIMs that include International Traveler Interchange Standard (ITIS) codes, or the emerging “managing disruptions for operations (MDO)” standard. These data may be communicated: • **Direct** from the IOO to the vehicle using RSUs broadcasting messages such as TIM. • **Indirect** from the IOO to the vehicle by vehicles using network cellular to connect to IOO APIs to request and receive data. • **Indirectly** by IOOs making data available to an **API/Feed** for private vendors to download the data and communicate the data to their in-vehicle applications.

Figure 12 illustrates the flow of CV data for each Data Category defined above.

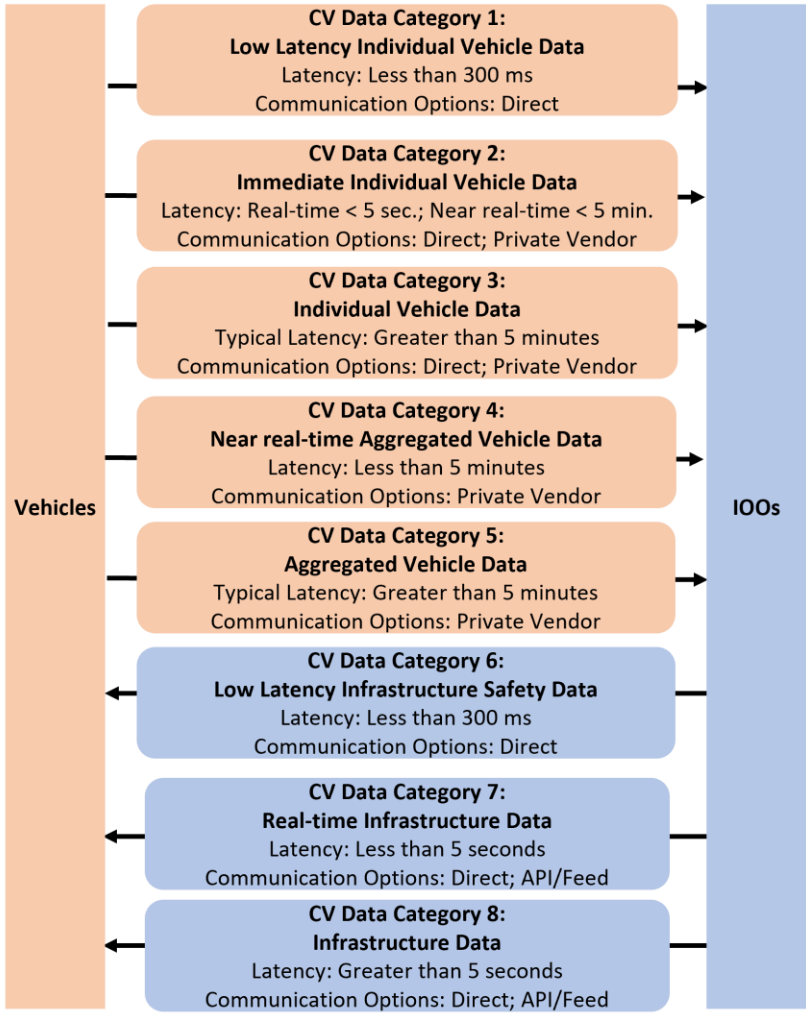


Figure 12: CV Data Categories Flow of Data for IOOs and Vehicles

These categories of data form the basis for the remainder of this Model CV Data Architecture.

4.2 Communications Requirements

Research in this project sought to document both the current and emerging approaches for communicating CV data between the vehicles and the IOOs in order to derive requirements and concepts for IOOs to consider for the communications aspect of their CV Data Architecture.

> To support readers as they select requirements most applicable to their local situation, CV data that IOOs can ingest are identified as:
> - ***Direct***
> - ***Vendor***
> - ***Aggregation***
>
> Note that a combination of two or more of these approaches may be a preferred solution.

4.2.1 CV Data Use Cases Related to Communications

The following are example use cases gathered from interviews and research conducted in this project. They describe elements of CV data communication to be reflected in the requirements included in Section 4.2.2.

Category 1: Low Latency Individual Vehicle Data

- **Direct.** There are examples of discussions about safety applications such as IMA where the IOO will perform a role in receiving data from vehicles and resending it to other vehicles that may be approaching the vehicle or common intersections, but no examples of deployments or demonstrations were identified in this outreach.

Category 2: Immediate Individual Vehicle Data

- **Direct.** Utah DOT receives requests for traffic signal priority (in the form of SRMs) from approaching transit vehicles and requests for traffic signal preemption from approaching snowplows. The signal requests are formatted as standardized SRMs and broadcast from OBUs using 5.9 GHz communications and received by RSUs at signalized intersections. These requests are processed and used to provide priority/preemption when appropriate.
- **Direct.** Utah DOT described their use of direct communications as they receive CV data broadcasted by vehicles passing by RSUs communicated as BSMs with extensions for weather and pavement friction data. In these situations, the use of direct communications allows Utah DOT to receive the vehicle data (direct from the vehicles) limiting latency and enabling Utah DOT to act on the data received. The use of direct communications enables vehicles in range of RSUs to communicate even if the area is lacking cellular coverage.
- **Direct.** Caltrans described their test bed in Palo Alto, California that equipped three buses to broadcast V2X communications to request transit signal priority.
- **Direct.** IOO estimated costs to install a RSU to receive data broadcasts are estimated to be $10,000-$16,000 and would vary depending upon whether supporting infrastructure is required to also generate messages to be broadcast. Also, backhaul communication would be required to move data to a central location and to send commands to the RSU. In situations where the CV data is solely used at the site of the RSU to generate new messages to be broadcast to other vehicles, backhaul communication needs may be reduced.
- **Vendor.** Pennsylvania DOT described a pilot project in Pittsburgh to use 4G LTE to receive locations of buses in real-time and then through a central system send the priority requests to the signals.

- **Vendor.** In Palo Alto California, Caltrans and University of California-Berkeley operate a testbed that receives individual vehicle data formatted as BSMs from a private vendor. The private vendor operates detection technologies at the intersection to detect vehicles (including location, speed, heading) and assemble BSM Part 1 messages for delivery to the testbed in near real-time, reporting approximately 10 BSMs per vehicle per second.

Category 3: Individual Vehicle Data

- **Vendor.** The Maricopa Association of Governments described that they receive individual vehicle data from a vendor and store the data to conduct queries that generate data to replicate floating car surveys, supporting corridor-wide signal retiming.
- **Vendor.** Maryland DOT works with a private CV data provider that equips Maryland DOT fleet vehicles with applications that communicate their position. This allows Maryland DOT to access the positions and movements of their fleet vehicles in near real-time, and also allows Maryland DOT to provide access of their data to the HAAS Alert system (using a direct feed from the CV data vendor to HAAS Alerts).

Category 4: Near real-time Aggregated Vehicle Data

- **Vendor.** Minnesota DOT procures segment travel time data from a private vendor. Minnesota DOT's Intelligent Roadway Information System (IRIS) accesses the data immediately and generates messages for display on DMS located throughout the metro area. The IRIS software communicates the travel times to DMS for display. No storage or access to the data is needed beyond this immediate use.
- ***Vendor.*** The city of Montreal described their use of private vendor data. They access the data through the vendor's API and use incident, speed, and travel time data in near real-time.
- ***Vendor.*** Costs for CV data procured from vendors and aggregation services through indirect CV communications vary based on many factors, including roadway miles, average daily traffic volumes, types of data procured, delivery mechanisms, and duration of agreement. Examples of vendor procured CV data cited include:
    - Aggregated near real-time CV data used for traveler information costs $400,000-$1,000,000 per year.
    - Aggregated CV data describing origin-destination patterns costs in excess of $1,000,000 per year.
    - Automatic Vehicle Location (AVL) data typically has lower costs, such as $100,000/year.

Category 5: Aggregated Vehicle Data

- **Aggregation.** Maryland DOT and Georgia DOT users access historical data through the Regional Integrated Transportation Information System (RITIS) data portal. This eliminates their need to store data and provides a readily-available query for any authorized user to access the data.

CV data that IOOs provision to CVs are identified as:

- ***Direct; and/or***
- ***API/Feed;***

to support readers as they identify requirements most applicable to their local situation.

Category 6: Low Latency Infrastructure Safety Data

- **Direct.** Feedback from the OEM community during the ITE Connected Intersection project suggested the strong need for in-vehicle applications to access low latency infrastructure safety data describing the current status of signalized intersections to support RLVW applications.
- **Direct.** Another example of CVs using low latency infrastructure safety data was the Traffic Optimization for Signalized Corridors (TOSCo) project demonstration. In TOSCo, strings of vehicles that follow each other closely form where the vehicle performs the longitudinal control (i.e., adjusting speeds and spacing from the vehicle immediately ahead). Signalized intersection status is used by vehicles to adjust speed as they either maintain speed through the intersection or stop/accelerate as appropriate.
- **Direct.** IOO estimated costs to install a RSU at an intersection to be $10,000-$16,000, assuming the signal controller is able to output the data required, there is space in the cabinet for new equipment, and there are backhaul communications to the signal. This excludes the engineering to prepare the site.

Category 7: Real-time Infrastructure Data

- **API/Feed.** The City of Ottawa described a pilot project[7] where the city's traffic operations branch made real-time signal system data available through APIs to signal timing prediction algorithms operated by third-party providers. In-vehicle units designed in-house by the City of Ottawa displayed information to drivers recommending speeds to achieve improved fuel consumption. Overall fuel savings of 5% were recognized when considering all daily travel of the test fleet.

Category 8: Infrastructure Data

- **API/Feed.** Iowa DOT noted they share work zone data formatted in the WZDx, as well as their advanced traffic management system (ATMS) and 511 data through an open portal, providing access to private vendors (and others). They are not aware of who is consuming the data.
- **API/Feed.** Arizona DOT discussed the high costs of deploying RSUs as a constraint on the coverage that could be offered. They are continuing to share condition data (e.g., work zone information) available over internet APIs for private vendors.

4.2.2 Needs and Requirements for CV Data Communications

Based on the use cases described in Section 4.2.1 and other material researched in this project, a series of ***needs and requirements for CV data communications*** have been derived as documented in Table 5.

Table 5: Communications candidate needs and requirements by CV data category

Category	Candidate Communications Needs	Candidate Communications Requirements
Category 1: Low Latency Individual Vehicle Data	• In order to support infrastructure involved safety applications (e.g., IMA) IOOs need to receive low latency communications of vehicle data.	• **Direct.** In locations where infrastructure involved safety applications are deployed, IOOs shall deploy RSUs and supporting systems to receive BSMs, extract the data from messages, and format the data into messages for rebroadcasts. • **Direct.** IOOs shall establish systems to meet latency requirements specific to the applications.
Category 2: Immediate Individual Vehicle Data	***Real-time Communications:*** • Fleet vehicles sending SRMs need to communicate the message frequently enough that it is received in time to be processed to activate priority or preemption.	• **Direct.** IOO fleet vehicle signal request messages shall be communicated from vehicle to infrastructure within 5 seconds from when the request is generated.
	Near real-time Communications: • IOOs that use individual vehicle data to support operations need to receive the data within 5 minutes to ensure the data describes current conditions. • IOOs may need these data in locations where cell service is not available. • IOOs need private vehicles to communicate BSMs frequent enough to receive broadcasts while the vehicle is in range of an RSU.	• **Direct.** IOOs deploying RSUs to receive vehicle broadcasted BSMs to collect immediate individual vehicle data shall deploy supporting backhaul communications to process, aggregate, and deliver the data to the systems that rely on the data within 5 minutes. • **Direct.** IOOs shall consider deploying RSUs at locations where individual vehicle data is required and may not be served by network cellular communications. • **Vendor/Aggregation.** Based on the requirements of operational systems using these data, IOOs shall specify communications latency requirements, when procuring data. Recommended to be 5 minutes maximum for operational uses.

Category	Candidate Communications Needs	Candidate Communications Requirements
Category 3: Individual Vehicle Data	• For situations when IOOs are using individual vehicle data for planning or post-event analyses, the need for communication latency is driven by individual IOOs schedules for using data. • When collecting CV data describing traffic/roadway conditions, IOOs need CV data to represent as many vehicles as possible, avoiding over/under representation of vehicle classifications and driving patterns.	• **Vendor/Aggregation.** IOOs shall describe their specific latency, delivery, data standards and formatting when procuring internet delivered CV data from private vendors or aggregation sources. • **Vendor/Aggregation/Direct.** IOOs shall deploy infrastructure, procure services, or seek digital infrastructure approaches to receive CV data from as many vehicles as possible. • **Vendor/Aggregation/Direct.** IOOs shall take steps to avoid over/under representation of vehicle classification or travel patterns.
Category 4: Near Real-time Aggregated Vehicle Data	• Aggregated data is used by IOOs for operations. IOOs need to receive these data in time for the data to support operational systems (e.g., travel times to DMS, broadcast TIM/BSMs).	• **Vendor/Aggregation.** IOOs shall procure services or seek digital infrastructure approaches to receive aggregated data that represent as large a percentage of vehicles traveling in the area of coverage as possible. • **Vendor/Aggregation.** Based on the requirements of operational systems using these data, IOOs shall specify aggregation periods (e.g., averaged over 5 minutes) and communications latency requirements when procuring data.
Category 5: Aggregated Vehicle Data	• Some IOO needs for aggregated vehicle data do not require near real-time communications. For planning or event analyses, aggregated vehicle data communicated with a latency of more than 5 minutes meets the needs of the IOO.	• **Vendor/Aggregation.** IOOs shall procure services or seek digital infrastructure approaches to receive these data from as many vehicles as possible while also avoiding over/underrepresenting vehicle classification or travel patterns. • **Vendor.** IOOs shall describe their specific delivery, data standards, and formatting when procuring internet delivered CV data from private vendors.

Category	Candidate Communications Needs	Candidate Communications Requirements
Category 6: Low Latency Infrastructure Safety Data	• Selected in-vehicle safety applications require low latency, secure communications between IOO and the vehicles.	• **Direct.** IOOs shall consider deploying CIs to communicate low latency infrastructure data at locations with higher-than-average red light running incidents. • **Direct.** IOOs shall consider current and emerging OEM deployments of V2X based in-vehicle RLVW or optimization applications in their decisions for connected intersection deployments. • **Direct**. Low latency infrastructure safety data to support RLVW applications shall be broadcast with total latency of less than 300 ms from when the data is generated and in compliance with detailed requirements defined by the CTI family of standards. • **Direct.** IOOs shall consider deploying CIs along corridors that are likely to benefit from traffic optimization applications (e.g., TOSCo applications) and at a time when production vehicles are equipped with traffic optimization applications. • **Direct**. Additional detailed requirements for communications at connected intersections to support TOSCo applications are available at the Crash Avoidance Metrics Partnership (CAMP) LLC website. • **Direct.** IOOs shall participate in other OEM collaborative efforts (as needed) to determine future requirements for low latency infrastructure safety data communications.

Category	Candidate Communications Needs	Candidate Communications Requirements

Category 7: Real-time Infrastructure Data	• Some connected vehicle applications need to receive data timely enough to issue driver alerts in real-time (less than 5 seconds). These include:
	  • Network-wide SPaT and MAP message communication (e.g., to support green light optimization or countdown to green applications)
	  • SSMs sent from the infrastructure to vehicle to support signal priority or preemption systems.
		• **API/Feed.** IOOs selecting to support network-wide green light optimization, green light countdown, or other similar applications shall publish network-wide SPaT/MAP messages no less frequent than every 5 seconds.
		• **API/Feed.** IOOs shall provide internet access to infrastructure status data (SAE J2735, National Transportation Communications for ITS Protocol (NTCIP) 1202, Traffic Management Data Dictionary (TMDD) v3.03d) as described in the CV PFS Interface Control Document for Using Third Parties to Deliver I2V).
		• **API/Feed.** IOOs shall consider deployment of a CVDF as described in the CV PFS Interface Control Document for Using Third Parties to Deliver I2V) to encourage consistency in IOO data delivery to third-parties. Published resources including a Concept of Operations, System Requirements, and Interface Control Document for CVDFs can be found at the CV PFS website.
		• **Direct**. IOOs shall broadcast SSMs at locations where it is needed to support signal preemption/priority using established standards no less frequent than 5 seconds after the message is generated.

Category	Candidate Communications Needs	Candidate Communications Requirements
Category 8: Infrastructure Data	• OEMs and private vendor application providers supporting driver alert applications need internet access to infrastructure data in order acquire the data and communicate it to in-vehicle applications. • OEMs and private vendors need consistency in the communication standards used for both V2X and internet exchanges of messages.	• **API/Feed.** For work zones that close one or more lanes, IOOs shall make available over internet access descriptions of the work zone activities formatted in the WZDx format within 20 minutes or less from the time of closure, as per CFR Title 23 Part 511 updated 11/10/22. • **API/Feed.** For roadway or lane blocking incidents, IOOs shall make descriptions of the incident available in either the TIM message, RSM message, or the planned "managing disruptions for operations" and/or TIMDx data exchange format (USDOT developing) within 10 minutes or less from time of verification for incidents within metropolitan areas and 20 minutes or less from time of verification for incidents outside metropolitan areas, as per Code of Federal Regulations (CFR) Title 23 Part 511 updated 11/10/22. • **API/Feed.** For hazardous driving conditions and road or lane closures or blockages caused by adverse weather, IOOs shall make available over internet access information about the hazard formatted in the TIM message, RSM message, and/or TIMDx data exchange format (USDOT developing) within 20 minutes or less from the time of closure, as per CFR Title 23 Part 511 updated 11/10/22.

Category	Candidate Communications Needs	Candidate Communications Requirements
Category 8: Infrastructure Data (continued)	• CVs will need to receive data timely enough to generate advisories and warnings to drivers, but low latency or real-time communications from the onset of conditions are not needed. • CVs equipped to receive direct communications need to receive data communicated as standard messages through V2X direct communications.	• **Direct.** Each IOO shall periodically review the number of vehicles equipped with V2X OBUs equipped with driver alert or warning applications to determine where/when there is enough penetration to achieve benefits of deploying RSUs to broadcast infrastructure data directly to vehicles. • **Direct.** IOOs shall consider deploying RSU broadcasts of travel disruption information in areas prone to hazards and when there are vehicles equipped to receive these messages. • **Direct.** IOOs shall consider deploying RSU broadcasts of work zone information for work zones on freeways and interstates when there are vehicles equipped to receive these messages. • **Direct.** IOOs shall consider deploying RSU broadcasts of data to support curve warning systems at IOO determined curves where data is needed and when there are vehicles equipped to receive and use these messages. • **Direct***:* Where RSU broadcasts are determined appropriate, IOOs shall broadcast infrastructure data using WZDx messages, TIM/RSM messages with ITIS codes to represent data, or other future standards at a frequency of at least 1 Hz to enable vehicles to receive the message within the broadcast range of the RSU.

4.2.3 Concepts for CV Data Communications

The following bullets present a series of concepts for IOOs to consider in order to address the needs and meet requirements for communication of each CV data category.

Vehicle to Infrastructure Data Categories (Categories 1-5)

- **Direct.** For IOO deployed RSUs receiving data directly from vehicles, the backhaul communication from the RSU to the IOO central systems is a significant topic and challenge. IOOs should explore any existing connectivity plans (e.g., regional traffic management connectivity plans) for communicating data between field equipment and central servers and leverage existing plans and/or communications systems to support backhaul communications. Both wired and wireless communications should be considered. IOOs might find that many elements of their local traffic management system may be connected or about to be connected to central systems using high-speed connection, and RSUs may take advantage of the same communication line.
- **Private/Aggregation.** While private vendor products delivering individual and aggregated vehicle data both immediately and historically are benefitting IOOs and represent the current industry

norm, IOOs should understand if the services they procure represent all vehicle classifications and driving patterns. If procuring and conflating multiple data sources, IOOs should be cautious of duplicative data from multiple sources.

- **Emerging.** IOOs are encouraged to pursue emerging communications concepts, such as the approaches described in the planned National Strategy on Roadway Digital Infrastructure and probe data standard that may help to expand IOOs access to additional vehicles and not require IOOs to either select one data source or procure multiple data sources.

Category 6: Low Latency Infrastructure Safety Data

- **Direct.** FHWA has sponsored Institute of Transportation Engineers (ITE), SAE, and partner agencies to conduct a well-organized and productive CI Program that has produced the CTI family of standards describing the communications of data to support connected intersections. In early 2023, SAE's Connected Transportation Interoperability Committee initiated activities to continue to advance the CTI family of standards. IOOs should monitor, participate in, and adhere to these standards and requirements as they deploy connected intersection infrastructure.

Category 7: Real-time Infrastructure Data

- **API/Feed.** Private vendor applications are more likely to be successful if consistent infrastructure status data is available from as many IOOs as possible. IOOs should support national trends, such as the publication of work zone data in the WZDx format. IOOs are encouraged to track FHWA efforts related to TIMDx.
- **Emerging.** The national strategy for roadway digital infrastructure may help IOOs better understand how to effectively benefit from existing and future digital infrastructures for data communications.

Category 8: Infrastructure Data

- **API/Feed.** There are examples of IOOs making infrastructure data available over internet connections (e.g., Extensible Markup Language (XML) / JavaScript Object Notation (JSON) feeds, data portals). While these meet communication requirements, IOOs should consider options for provisioning data directly to vehicles, through data exchanges or through approaches outlined in the national strategy for roadway digital infrastructure.
- **API/Feed.** CFR Title 23 Part 511 was established in 2010, specifying requirements for IOOs to make consistent data available describing driving conditions and road hazards. CFR Title 23 Part 511 still exists and was updated in November 2022. While originally developed to support traveler information, these guidelines could form a basis for consistency in data provided for CV data.
- **API/Feed.** The CV PFS project titled "Using Third Parties to Deliver I2V" has published multiple deliverables detailing approaches and standards use for delivering infrastructure data to third parties for dissemination to CVs. Specifically, the use of a CV Data Framework is described. Florida DOT is an agency that is proceeding with developing a CV Data Framework following this model.
- **Emerging.** The national strategy on digital infrastructure is expected to define enablers to help expand IOOs communications of infrastructure status data, potentially reducing costs and enabling communications directly to vehicles.

4.3 Access and Storage Requirements

This section includes uses cases, needs and requirements, and concepts that IOOs may consider for storing and providing access to CV data throughout their organization.

4.3.1 CV Data Use Cases Related to Access and Storage

The following are example use cases gathered from interviews and research conducted in this project and describe elements of access and storage to be reflected in the requirements included in Section 4.3.2.

Categories 1-5 All Vehicle Data

- **Direct.** Georgia DOT noted that they are still defining their approach for storing BSM data received at RSUs and have indicated they would likely not store all BSMs, but rather those of relevance (something significant happening or BSMs in the vicinity of a work zone or other event). Their current practice is to store the individual BSMs (vs. aggregating), but they are considering some processed format for storage. They currently are storing SRMs sent by transit vehicles for performance evaluation.
- **Direct.** Ohio DOT indicated that if they do store BSMs the duration of storage may not exceed 8-10 hours to allow examination of specific conditions. Ohio is considering an aggregated BSM methodology that would allow them to count occurrences of vehicles. For example, they might count the number of vehicles (by BSM reports) traveling between 30-40 miles per hour (MPH) in a designated time period, and the number of vehicles within vehicle classifications (defined by length of vehicle). This would further support traffic management operations in that algorithms could be built with artificial intelligence (AI) models to examine BSMs and the presence of low-speed vehicles on a higher speed roadway, thus denoting a possible issue on that segment of road. Additionally, large variances in vehicle speeds could be a key contributor to higher accident probability.
- **Direct.** Utah DOT RSUs receive BSMs from passing vehicles and create warning messages formatted as TIMs to send to other vehicles. Utah DOT needs their cloud-based platform (Cirrus) to access this data immediately when received in order to generate timely messages.
- **Direct/Vendor.** Caltrans and UC Berkeley have noted experiences with data storage based on their testbed as follows:
  - 800 bytes is a safe upper estimation size of one BSM (Unaligned Packed Encoding Rules (UPER) encoded).
  - For an intersection with annual average daily traffic (AADT) of 40,000 vehicles, the daily BSM (part 1 only) storage size is approximately 32 GB. When the testbed stores the BSM payload into data files, the average data size per intersection per day is approximately 1.5 GB unzipped and 300 MB zipped.
- ***Vendor.*** Iowa DOT receives historical CV data and compresses and converts the data to reduce the size and allow faster queries. Users then access the data from their data analysis systems.
- **Private Vendor.** The City of Ottawa described their planning group's use of historical aggregated origin destination data. They currently only have access to aggregated data but indicated interest in possibly accessing the raw data to better understand the routes taken.

- **Vendor.** Several of the agencies interviewed in this project described how two or more groups in their agency either procured CV data from the same vendor or were considering procuring CV data from the same vendor, illustrating the need for open sharing within agencies about what data is being procured or received and sharing access to others.
- **Aggregation.** Maryland DOT and Georgia DOT users access historical data through the RITIS data portal, eliminating the need for local storage.

Category 6: Low Latency Infrastructure Safety Data

- ***Direct.*** Members participating in validation of CIs relied on the storage of signal controller data for comparison against SPaT message content.

4.3.2 Needs and Requirements for Access and Storage of CV Data

Based on the use cases described in Section 4.3.1 and other material researched in this project, a series of ***needs and requirements for CV data access*** have been derived as documented in Table 6.

Table 6: Candidate needs and requirements for access and storage by CV data category

Category	Candidate Access and Storage Needs	Candidate Access and Storage Requirements
General (Multiple CV Data Categories)	• IOOs need data to be accessible to staff and/or software systems when and where needed. • IOOs may need partner agencies (e.g., universities, metropolitan planning organizations (MPOs)) to have access to the data.	• **Vendor/Aggregation.** If procured, procurement agreements shall specify delivery approaches to enable storage and access by individuals or local query systems. • **Vendor/Aggregation.** Procurement agreements shall clearly allow sharing and access for other groups in the agency and partner agencies. • **Vendor/Aggregation/Direct.** If CV data is stored locally, data storage shall take steps to enable access and use by others in the agency. • **Vendor/Aggregation**. If relying on storage by vendor or aggregation service, IOO's procurement agreements shall define minimum storage and availability periods, including survivability of contract end dates. • **Direct/Vendor.** IOOs shall consider the need to store BSMs during debugging and system testing, as well as limited storage to support research. • **Direct/Vendor.** IOOs shall consider the need for long-term storage of BSMs against the storage demands and costs of storage. • **Direct/Vendor.** IOOs should plan for an estimated storage of 800 bytes per BSM received. For an intersection with 40,000 AADT, preliminary daily BSM storage size should be allocated as 32 GB. • **Vendor/Aggregation.** If IOOs are receiving individual vehicle data or aggregated vehicle data from vendors or aggregation services, IOOs shall consider multi-year storage of the messages received and shall plan for storage using specifications provided by the vendor.
Category 1: Low Latency Individual Vehicle Data	• Due to the low latency use of these data, long-term storage has minimal benefits for operations but may be needed for crash investigations, troubleshooting, or to verify data received/broadcast.	• **Direct.** When receiving low latency individual vehicle data, IOOs shall develop local policies for the duration of storage of low latency individual vehicle data, based on the applications supported.

Category	Candidate Access and Storage Needs	Candidate Access and Storage Requirements
Categories 2-5 (Individual Vehicle and Aggregated Vehicle Data)	• No specific needs beyond the General Candidate Needs presented above.	• No additional requirements beyond the General Requirements presented above.
Category 6: Low Latency Infrastructure Safety Data	• **Access:** CVs need access to signal status data (i.e., SPaT, MAP, RTCM) that meets the strict requirements defined in the CTI family of standards. Storage of multiple aspects of the data (e.g., signal controller output time, message creation time, broadcast time, etc.) are needed during the validation process. • **Access:** Corridor-wide access to data is needed to achieve full mobility and efficiency benefits.	• **Direct.** IOOs shall store the SPaT/MAP/RTCM messages broadcast by CIs, as a minimum as needed for CI validation and described in the CV PFS developed Connected Intersections Test Plan and Connected Intersections Detailed Testing Log. • **Direct.** IOOs shall consider the need to store the SPaT/MAP/RTCM messages broadcast for future debugging or to support queries of data availability and make storage decisions based on agency policies and procedures. • **Direct.** IOOs shall consider current and emerging OEM deployments of V2X based in-vehicle RLVW or longitudinal control applications in their decisions for CI deployments.
Categories 7-8: Infrastructure Data	• Green light optimization applications need signal control status in as many locations as possible to deliver optimal benefits.	• **API/Feed.** IOOs shall consider publishing real-time internet feeds of signal controller status data for third-party providers and OEMs to consume.

4.3.3 Concepts for CV Data Access and Storage

The following bullets present a series of concepts for IOOs to consider in order to address the needs and meet requirements for CV data access and storage.

Categories 1-5 (All Vehicle Data):

- **Vendor/Aggregation/Direct.** There are examples of best practices to help prevent stove-piping of data use and encourage more groups in an IOO to be aware or, understand, and ultimately use the data sources. Examples include:
  - Using a data source catalog/assessment tool to describe an approach to tracking and describing each data source. Fields that might be populated for each data source include data source, description, ownership, format, size, cost, security level, restrictions.

- Suggestion that data stored be organized using a "regular file system" structure and that the structure, data sets, and access policies are managed to accommodate end users' needs and encourage use.
- Data can be more discoverable if a searchable metadata repository is used.

These and other best practices are well documented for big data management, although not typically called out as CV data. The NCHRP report "Guidebook for Managing Data from Emerging Technologies for Transportation"[8]

- **Vendor/Aggregation/Direct.** The volume of CV data, and data in general, that is produced and consumed at an agency has increased dramatically in the past decade, and this trend is expected to continue with the deployment of new connected infrastructure and vehicles. It is important for IOOs to plan for a future state when determining access and storage needs, lest the procured system requires replacement or significant expansion in a short timeframe, long before the system lifecycle. For instance, IOOs that receive BSMs at RSUs note that a limited number of OBUs broadcasting these messages but anticipate more data within five years and have therefore designed CV data access and storage to scale up in the near-term. To illustrate this point, a USDOT study, *Integrating Emerging Data Sources into Operational Practice: State of the Practice Review*,[9] projected the growth in agency data storage needs from 2016 through 2026.
- **Direct.** IOOs receiving BSM or SRM data are encouraged to follow the approaches of Utah, Ohio, Georgia, and California and explore different approaches for BSM storage before settling on the selected approach.
- **Direct.** IOOs are encouraged to consider the privacy aspects of storing BSM data as the storage of multiple BSMs can introduce privacy concerns that exceed those of individual BSMs.
- **Vendor/Aggregation/Direct.** IOOs are encouraged to consider cloud storage approaches when storing CV data.
- **Vendor/Aggregation/Direct.** IOOs are encouraged to consider their experiences with public records requests for other data, camera images, and video when defining their CV data storage approach.

Categories 6-8 (All Infrastructure Data):

- **Direct/API/Feed.** The Australian Road Authority (Austroads) developed multiple guidance documents for provisioning CV to connected and automated vehicles.[10] A recurring recommendation is that road authorities collaborate to establish consistency in the back-end systems that provision data to provide private information providers with access to real-time data that is consistent with similar levels of service across all agencies.

4.4 Privacy Requirements

A major concern shared in both the public and private sectors is the protection of personally identifiable information (PII), which could include individual identities and associated information like social security numbers, date and place of birth, current address, as well as license plate or vehicle information like the vehicle identification number (VIN). PII is a target for hackers, and data breaches are "too common for comfort."

Beyond the inclusion of PII in a data field, there are also concerns about privacy that may result from storage of large volumes of CV data. For example, even if anonymized, access to data describing repeated trips from an origin to a destination could be used to derive knowledge of an individual's repeated travel (especially if the travel occurs in rural areas with limited residences and locations of services).

Further, in order to gain the public's trust, both public and private sector entities sometimes go to great lengths to assure users that their privacy will be protected, often by not collecting PII in the first place. The National Institute of Standards and Technology (NIST) Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)[11] provides recommendations for IOOs to consider for CV data. Additional resources for data privacy considerations are available in Appendix A.

4.4.1 CV Data Use Cases Related to Privacy

The following are example use cases gathered from interviews and research conducted in this project. They describe elements of CV data privacy to be reflected in the requirements included in Section 4.4.2.

Categories 1-5 (All Vehicle Data):

- **Vendor.** Private vendor CV data providers have a variety of privacy procedures in place to ensure that PII is not included with CV data that is procured by IOOs. These policies can often be found on the private vendor website.
- **Direct.** Colorado DOT noted that for fleet vehicles and OEM vehicles interfacing with their system the Basic Safety Message (BSM) vehicle IDs change every 5 minutes or 2 km to protect privacy of the vehicle and occupants. They also noted that BSM data is processed as it is received to remove any potentially identifiable information (e.g., make, model, year).
- **Direct/Vendor.** Some states like California and Texas have robust laws about data privacy, and so the use of CV data with PII at the IOOs are either not allowed, strongly discouraged, or governed by specific policies and procedures.
- **Vendor.** Even the perception of privacy concerns can be a challenge. Multiple CV PFS sites noted the challenges convincing executive staff to allow purchase of private vendor data as the fear that the agencies might be tracking individuals (or perceived by others as tracking them).
- **Vendor.** Iowa DOT noted that the historical aggregated vehicle data they procure identifies "block areas" not specific addresses, adding anonymity and reducing risks of privacy concerns.
- **Vendor.** California described how all data received is anonymized which results in the data being heavily aggregated and sometimes manipulated beyond what is considered to be reasonable.
- **Direct/Vendor.** During an interview for this project Kentucky Transportation Cabinet noted the importance of anonymized data since the CV data is subject to public records requests.

- **Vendor.** The City of Ottawa noted that data describing scooter trips includes some data about the rider and introduces potential privacy risks.
- **Direct.** The New York City, Tampa, and Wyoming CV Pilot sites all developed privacy plans[12,13,14] that primarily emphasize the protection of PII by focusing on a triad of:
  - Confidentiality: preserving authorized restrictions on information access and disclosure.
  - Integrity: guarding against improper information modification or destruction
  - Availability: ensuring timely and reliable access to and use of information
- ***Direct.*** Ohio DOT has developed comprehensive and detailed Security and Privacy Requirements, which address privacy concerns related to data received through V2X deployments including interfaces at the transportation management center (TMC), RSU, and OBU.

4.4.2 Needs and Requirements for Privacy of CV Data

Based on these use cases and other material researched in this project, a series of ***needs and requirements for CV data privacy*** have been derived as documented in Table 7.

Table 7: Privacy candidate needs and requirements by CV data category

Category	Candidate Privacy Needs	Candidate Privacy Requirements
General (All CV Data Categories)	• IOOs need to limit PII in CV data to the greatest extent possible in order to reduce the risk of unauthorized sharing or access.	• **Vendor/Aggregation/Direct.** IOOs shall make every effort to collect, acquire, and store CV data in a way that contains no PII. • **Vendor/Aggregation/Direct.** IOOs shall develop and implement retention and use policies for CV data (even if believed to contain no PII) to prevent any retention or access to PII (i.e., to the extent possible, PII should not be acquired or retained). • **Vendor/Aggregation/Direct.** IOOs shall prevent conflation of non-sanitized and de-anonymized CV data with other data sources. • **Vendor/Aggregation/Direct** IOOs shall ensure their handling of CV data meets or exceeds any established policies and procedures for all data handling and data exchanges defined by their agency or at the state level.
	• Even in the absence of PII, IOOs need to take steps to prevent any pseudo privacy risks that could come with volumes, geographic coverage, or date/time ranges of data.	• **Vendor/Aggregation/Direct.** IOOs shall perform a risk analysis to ensure that safeguards are in place to minimize the risks of derivation of PII from conflated or large volumes of CV data.
	• IOOs need to understand what PII, if any, is contained in the CV data in order to understand how to manage and use the data.	• **Vendor/Aggregation/Direct.** IOOs shall understand what PII may be contained in all CV data in order to understand how to effectively manage and use the data in regard to privacy.

Category	Candidate Privacy Needs	Candidate Privacy Requirements
		• **Vendor/Aggregation**. IOOs shall understand privacy policies that apply to the CV data source (i.e., agency policies, contracts with private vendors). • **Vendor/Aggregation/Direct.** IOOs shall understand whether removal of PII from CV data would limit usefulness of the data.
Categories 1-3 (All Individual Vehicle Data)	• Even in the absence of PII, IOOs need to take steps to prevent any pseudo privacy risks that could come with volumes, geographic coverage, or date/time ranges of data.	• **Direct.** IOOs shall minimize BSM storage volumes of collected BSMs by limiting 1) size, 2) coverage area of BSM transmissions, and 3) time span per configured limits to the extent possible to prevent correlation and subsequent privacy losses. • **Direct.** BSMs received from non-agency vehicles shall be managed in a way that removes any PII or vehicle-identifying information before storing or sharing data. • **Direct.** BSMs received from agency vehicles shall retain vehicle-identifying information as needed for applications (e.g., to validate signal requests) and shall be managed per local agency approach for retaining agency generated data. • **Direct.** IOOs may implement policies to access, manage, use, and store CV data that differ for agency vehicle BSMs, non-agency BSMs, and data from private vendors. • **Vendor/Aggregation.** Data received from private vendors shall require (and be tested to verify) removal of all PII before IOOs receive data.
Categories 4-5 (All Aggregated Vehicle Data)	• **Privacy.** IOOs need to ensure that temporal aggregation of CV data does not introduce any pseudo privacy risks.	**Vendor/Aggregation.** When acquiring aggregated CV data, IOOs shall review the data periodically to ensure that PII sensitive topics such as travel patterns to or from specific addresses cannot be derived from the data. **Vendor/Aggregation.** IOOs shall consider specifying that aggregated data received from private vendors identifies “block areas” not specific addresses (e.g., in origin destination tables), adding anonymity and reducing risks of privacy concerns.
Categories 6-8 (All infrastructure Data)	• No specific needs for privacy as infrastructure status data is available to the public.	• No requirements for privacy beyond standard privacy requirements related to each agency’s data.

4.4.3 Concepts for CV Data Privacy

Below are a series of concepts for IOOs to consider in order to address the needs and meet requirements for privacy of each CV data category.

Overarching Concepts for CV Data Privacy in Vehicle to Infrastructure Data

- **Direct.** Ohio DOT has developed detailed Security and Privacy Requirements related to V2X exchanges that can be accessed and viewed at the link provided.
- **Direct.** The National Institute of Standards and Technology (NIST) Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)[11] provides recommendations for IOOs to consider for CV data. Additional resources for data privacy considerations are available in Appendix A.
- **Direct.** Because most IOO applications of CV data do not require any PII (exceptions include tolling applications that use CV data and signal request messages that require validation that the request came from a valid vehicle), both IOOs and private vendors often make a significant effort to anonymize collected CV data through either aggregation or generalization of origins and destinations before it is shared or used, or avoid the collection of PII in the first place.
- **Vendor/Aggregation.** IOOs have and should continue to work with private vendors to understand their privacy policy (e.g., which often protects users by removing or anonymizing data elements containing PII).
- **Direct / Vendor.** If IOOs do receive CV data with PII, they may consider processing the data, as needed upon receipt, in order to remove the PII before it is stored, shared, or used by IOO staff.
- **Direct/Vendor/Aggregation.** IOO business practices on data privacy are often governed by established policies and procedures for all data and data exchanges that occur, either within their agency or at a state level, and are not specific to CV data (e.g., California and Texas have strict data privacy laws that limit IOO access to CV data with PII, and Texas follows a standard process to evaluate the potential risks of exchanging data with private vendors regarding both security and privacy).
- **Direct/Vendor/Aggregation.** Given universal concerns over privacy, national security standards that consider data privacy would be beneficial for IOOs and CV data providers to follow in order to streamline the process of ensuring security for CV data, rather than each agency working independently in silos.

Categories 1-3: All Individual Vehicle Data

- **Direct.** BSMs that IOOs receive at the roadside from CVs are inherently anonymized per the SAE J2735 standard. However, IOOs need to be aware that the assembly and storage of large volumes of BSM messages (or the data from the messages) can introduce pseudo PII as travel patterns could be derived from the large data sets.
- **Direct/Vendor/Aggregation.** If individual CV data needs to be stored, IOOs should consider implementing data retention policies that minimize the duration of time that CV data with potential PII are archived.

Categories 4-5: All Aggregated Vehicle Data

- **Vendor/Aggregation.** The processing and aggregation of data performed by private vendors should anonymize data. For example, an IOO receiving average segment travel times or segment volume data will not have any tie to PII or the potential to derive PII. An example of an exception would be O-D trip tables if the locations were specific to addresses (or could allow derivation of addresses). IOOs should consider this when specifying blocks of areas for the trip origins and destinations.

4.5 Security Requirements

Security considerations for CV data are primarily driven by three factors:

1. The presence of PII in the CV data, which requires robust security to prevent hackers from unauthorized access of private individuals' data.
2. Securing direct communications to allow the infrastructure system or vehicle receiving data to trust that the data is from the designated source.
3. Protection of IOO and CV systems to prevent unauthorized access and actions that could put the traveling public at risk.

4.5.1 CV Data Use Cases Related to Security

The following are example use cases gathered from interviews and research conducted in this project. They describe elements of CV data security to be reflected in the requirements included in Section 4.5.2.

General to All Data Categories

- Security for CV data that IOOs receive, access, and store has all the data security concerns of other data collected by the IOOs. However, security is a greater concern in some instances, including the following:
  - **Direct.** Because V2X message broadcasts are received by OBUs and RSUs without any 'handshaking,' it is possible that OBUs or RSUs may broadcast erroneous and possibly damaging messages, either intentionally or accidentally. A Security Credential Management System (SCMS) has been established and continues to evolve that includes clear security mechanisms to prevent these.
  - **Direct/Vendor.** When any PII is attached to the CV data additional security accommodations should be made, as appropriate.
- **Direct.** Caltrans noted that for their existing test beds, all BSM data is stored in one location (on University of California-Berkeley servers) where the data is secured.
- **Direct.** Ohio DOT has developed Security and Privacy Requirements which address security at the V2X platform, roadside equipment, OBU, and TMC as well as the CV application interfaces at the TMC, RSU, and OBU. Additionally, Ohio DOT's Security Checklist covers general security considerations, as well as those focused on servers, OBUs, and RSUs during planning, design, procurement, deployment, and operations.

Vehicle to Infrastructure Data Categories

- **Vendor.** Feedback from members interviewed in this project identified an example where a private vendor CV data provider included security requirements for how IOOs communicate and store their data (to prevent inappropriate or unauthorized use and access to the private vendor data).

4.5.2 Needs and Requirements for CV Data Security

Based on the use cases described above and other material researched in this project, a series of ***needs and requirements for CV data security*** have been derived as documented in Table 8.

Table 8: Security candidate needs and requirements by CV data category

Category	Candidate Security Needs	Candidate Security Requirements
General (All CV Data Categories)	• **Security.** IOOs need to address security concerns in all aspects of agency systems to ensure that CV data and supporting infrastructure are secure from unauthorized access.	• **Direct.** IOOs shall perform security enrollment for new V2X devices and equipment in a secure environment using an approved process and equipment to validate that it is authentic. • **Direct.** IOOs shall secure communications of CV data between the roadside and TMC. • **Vendor/Aggregation/Direct.** IOOs shall secure infrastructure, TMCs, and other facilities against unauthorized access to ensure data integrity through standard security protocols (e.g., strong password protection and expiration, database encryption, etc.). • **Direct.** IOOs shall ensure infrastructure and devices are routinely upgraded to the latest industry standard security protocols. • **Direct.** IOOs shall ensure roadside devices allow for security access and update capabilities by authorized providers. • **Vendor/Aggregation/Direct.** IOOs shall, at a minimum, secure CV data according to agency security policies and principles for other datasets, including archives, access, and communications considerations. • **Vendor/Aggregation/Direct.** IOOs shall consider requiring users to register for authorization to access and use data according to established rules and policies, including CVs or private vendors accessing IOO data or internet or cloud interfaces. • **Vendor/Aggregation.** IOOs shall secure CV data received by private vendors to prevent unauthorized access, sharing, and use that would violate the terms by which the data was procured or collected.
Categories 1-3 (All Individual Vehicle Data)	• **Security.** IOOs need to understand that CV data broadcasts are secure and authentic.	• **Direct.** IOOs shall ensure that the deployed RSUs use the SCMS to verify the authenticity of the security signature of CV data received at the RSU (e.g., determine that the BSM was received by an OBU participating in SCMS). • **Direct.** IOOs receiving and processing signal request messages shall ensure the RSUs and SCMS

Category	Candidate Security Needs	Candidate Security Requirements
		provider verify that received SRMs possess the SRM public safety identification (PSID) and system security and privacy (SSP) permissions, before SRMs are considered for signal priority. • **Direct.** IOOs shall be able to verify the security credentials of data received from vehicles prior to its use (e.g., for signal priority) via the functionality provided within the RSU and between the RSU and SCMS provider. • **Vendor/Aggregation.** No specific security requirements beyond the general requirements.
Categories 4-5 (All Aggregated Vehicle Data)	• **Security.** No specific needs beyond the General Needs presented above.	• **Vendor/Direct.** IOOs receiving CV data that has been aggregated using edge computing or if the IOO performs aggregation on the edge, the transmission of aggregated data to back end systems shall be secured using transport layer security (TLS) or some other method to ensure that malicious data is not being transferred to IOOs for use in decision making.
Category 6: Low Latency Infrastructure Safety Data	• **Security.** CVs operating safety applications need IOO data broadcasts to be secured and accurate (i.e., verified they are sent by a trusted IOO).	• **Direct.** IOOs shall secure communications between infrastructure systems generating data (e.g., signal controllers) and the RSU. • **Direct.** IOOs shall deploy an SCMS certification mechanism signaling to vehicles that the information should be trusted. • **Direct.** IOOs shall understand and implement any verification of data required to receive additional SCMS flags indicating the data is verified to be used by safety applications such as RLVW (as outlined in the CV PFS CI Guidance Document). • **Direct.** IOOs shall verify the integrity of data received by external sources (e.g., RTCM) before rebroadcasting it. • **Direct.** IOO messages that are signed by an IOO (or local area of IOO jurisdiction) shall be valid only for the geographic area covered by the jurisdiction.
Categories 7-8 (Real-time and Non-Real-time Infrastructure Data)	• **Security.** CVs operating applications need a mechanism to understand that received infrastructure status data is from an IOO source.	• **Direct, API/Feed, Network Cellular.** IOOs shall provide a mechanism for CVs to recognize the IOO as source of the infrastructure status data (e.g., SCMS for RSU broadcasted data, trusted agency ownership of IOO operated data feed for API/Feed). • **Direct.** IOOs shall secure communications between infrastructure systems generating data (e.g., signal controllers) and the RSU.

Category	Candidate Security Needs	Candidate Security Requirements
		• **Direct.** IOOs shall deploy an SCMS certification mechanism signaling to vehicles that the information should be trusted. • **Direct.** IOOs shall verify the integrity of data received by external sources (e.g., RTCM) before rebroadcasting it.
	• **Security.** CVs need to know that the IOO source is the authorized entity for the jurisdiction to which the data applies.	• **Direct, API/Feed, Network Cellular.** IOO messages that are signed by an IOO (or local area of IOO jurisdiction) shall be valid only for the geographic area covered by the jurisdiction.

4.5.3 Concepts for CV Data Security

The following are a series of concepts for IOOs to consider in order to address the needs and meet requirements for security of each CV data category.

Overarching Concepts for CV Data Security

- **Direct.** IOOs shall consider developing specific privacy requirements for local CV deployments and sharing examples of these with other IOOs. Ohio DOT has developed detailed Security and Privacy Requirements related to V2X exchanges that can be accessed and viewed at the link provided.
- **Direct/Vendor/Aggregation.** Given the universal concerns over security, national security standards that consider data privacy would be beneficial for IOOs and CV data providers to follow in order to streamline the process of ensuring security for CV data. IOOs should consider advocating for these national security standards if/when a national CV Deployment Plan is developed.
- **Direct/Vendor/Aggregation.** IOO business practices on data security are often governed by established policies and procedures for all data and data exchanges that occur, either within their agency or at a state level, and are not specific to CV data (e.g., Texas follows a standard process to evaluate the potential risks of exchanging data with private vendors regarding both security and privacy).
- **Direct/Vendor/Aggregation.** At a minimum,
  - CV data should be secured in a manner that is consistent with traditional IOO datasets.
  - IOOs should implement SCMS for V2X communications of CV data.
  - Data exchanged with private vendors should follow standard internet and/or cloud security procedures (as would be for other IOO data exchanges).
- **Direct/Vendor/Aggregation.** As the national strategy for RDI advances, IOOs should insist security concerns are addressed by new exchange processes.

4.6 Data Sources

This section presents the types of CV data available, the sources that provide them, and how these data sources meet the requirements presented in the previous sections. Specifically, as shown in Figure 13, a section on vehicle to IOO data sources (Section 4.6.1) presents four types of data sources and describes how they meet requirements for access, privacy, security, and communications. A second section (Section 4.6.2) likewise presents information about three types of available IOO to vehicle data sources and describes how they meet the requirements.

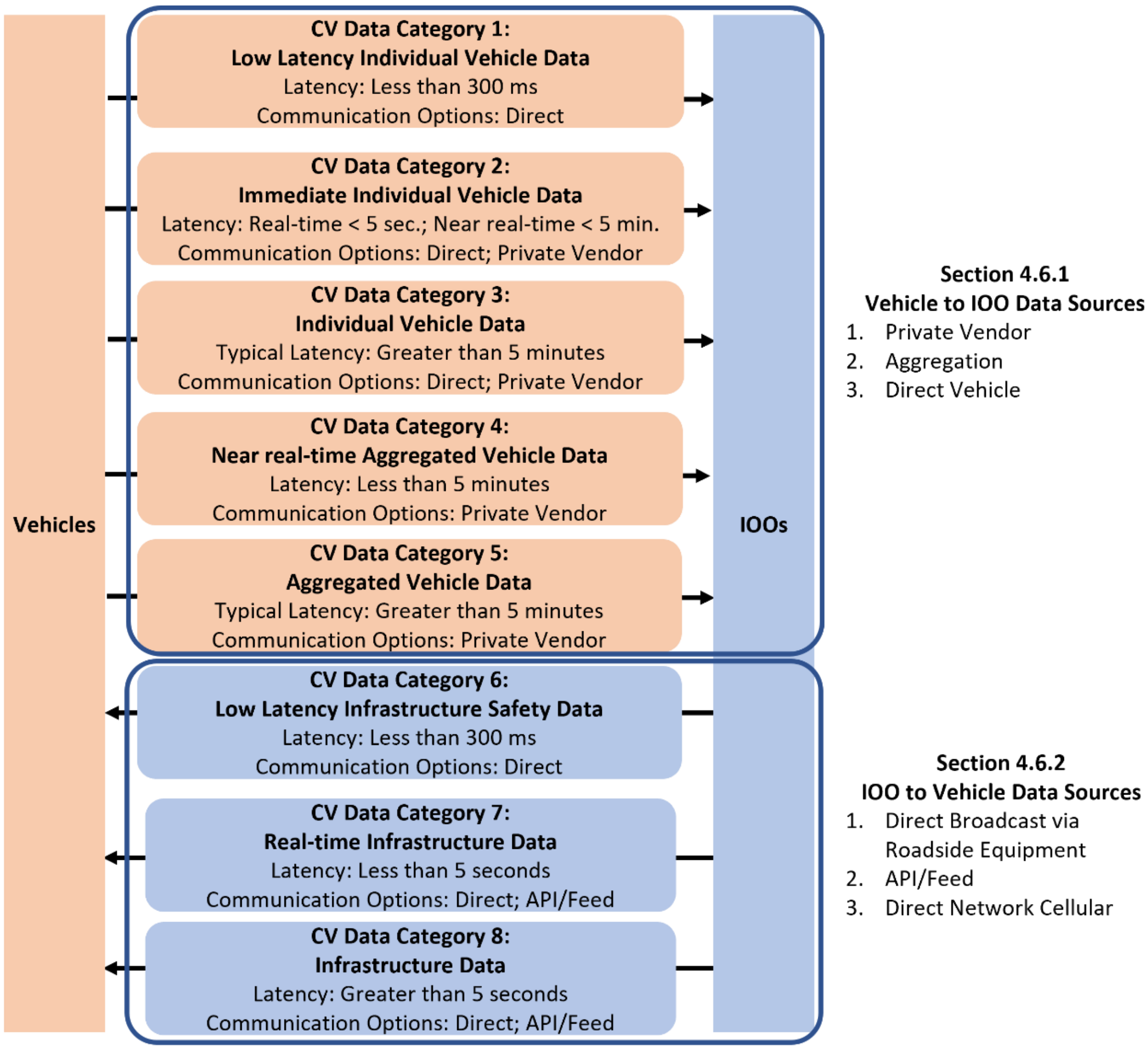


Figure 13: CV Data Sources related to CV Data Categories

4.6.1 Vehicle to IOO Data Sources

In general, sources of data that are transmitted from vehicles to IOOs can be organized within one of three types:

1. **Private vendor.** Private vendors of CV data offer a menu of data offerings that IOOs can procure, including immediate or historical vehicle or aggregated data. These data sources serve as an intermediary between the vehicles and IOOs by collecting data from CVs, potentially adding value by applying quality control or aggregation, and providing specific datasets that IOOs may be able to directly use with minimal additional effort (e.g., segment travel times for real-time traveler information maps).
2. **Aggregation.** Aggregation portals are another form of private vendors that ingest CV data from a variety of public and private sources in order to repackage it or provide a single access point for IOO purposes. These providers generally use data fusion tools that account for the data quality, prioritize the CV data sources for different segments in order to provide one or more streamlined data feeds that IOOs can apply for various use cases with increased assurance of their reliability and accuracy.
3. **Direct Vehicle**. Data in this category is communicated directly by the vehicle to the IOO, either as broadcast data from vehicles (e.g., in the form of SAE J2735 Basic Safety Messages) or by vehicles responding to probe data message requests and sending SAE J2945/C Probe Data Report Messages. IOOs receive this data in real-time, typically as a standardized message that requires processing to be used for various IOO applications.

Current and anticipated data offerings for these three types of vehicle to IOO data sources are mapped to the CV data categories in Table 9. Data offerings identified in Table 9 are available that are capable of meeting all the requirements described in previous sections, however IOOs must also ensure that the procurement and IOO systems include any additional considerations to meet the requirements. The number of private vendor sources identified as part of this effort that are available for each CV data category are indicated in parenthesis in the table below.

Table 9: Types of vehicle to IOO data sources by CV data category

Types of Vehicle to IOO Data Sources	CV Data Categories				
	#1 Low Latency Individual Vehicle	#2 Immediate Individual Vehicle	#3 Individual Vehicle Data	#4 Near Real-time Aggregated	#5 Aggregated Vehicle Data
Private Vendor Sources		✓ (3+)	✓ (3+)	✓ (5+)	✓ (10+)
Aggregation Sources		✓	✓	✓	✓
Direct Vehicle Sources	✓	✓	✓		

✓ = Currently Available

4.6.1.1 Summary of Private Vendor Sources

As part of this effort, a survey was sent to known CV data providers, that resulted in 10 responses. Additionally, information was gleaned from interviews with IOOs who procure and use CV data from various providers. As a result, 12 private vendors of CV data were reviewed and considered for content in this section. The names of private vendors who responded to the survey and those identified in the interviews are not included in this report because the intent is not to call-out services or features of any individual vendor, but rather to reflect the overall availability of data offered collectively by private vendor sources.

Specific examples of CV data provided by one or more private vendors include:

- Segment-based travel time and speed data, either immediate or historical, generated from aggregation of individual vehicle data.
- Individual vehicle location, speed, and heading assembled into BSM messages, provided by a vendor monitoring vehicles approaching intersections and converting vehicle detections into BSMs.
- Volume data, including vehicle miles travelled and turning movement count data generated from aggregating CV data.
- Origin-destination data created by aggregating individual CV data and delivering it as O-D tables.
- Pavement condition data and asset condition monitoring, generated from data collected by vehicle sensors or mobile phone/dashcam images.
- Work zone monitoring, using data collected through mobile phone images or dashcams.
- Safety analytics data, including hard braking occurrences aggregated from individual vehicle CV data.
- Signal analytics, based on aggregated historical data from CV and device data.
- Roadway analytics, including identification of bottlenecks and congestion determined by aggregating CV data.
- Weather roadway conditions, determined by aggregating CV data.

The private vendor sources for CV data are compared to the requirements in the following bullets:

- **Access & Storage.**
  - ***CV Data Access.*** Several requirements in Table 6 describe requirements related to how and when IOO staff or operational systems need to access CV data. These requirements are met by private vendor sources either by IOO staff using the private vendor web interface, or via IOO web systems that automatically pull the data from vendor web addresses or get the data pushed from vendor websites.
  - ***Access and Use.*** Table 6 also includes requirements related to who can access and use data. In general, IOOs may have needs for outside consultants or university researchers to access the data and similarly will likely wish to encourage additional IOOs to use the data procured. Use case examples described examples when sharing of data is allowed as well as examples of limitations on sharing of private vendor source data, but in general, access and use of vendor-provided data varies depending on the terms of the contract with the IOO, such that data use and sharing may be restricted to within the IOO or only be used for specific purposes.
  - ***Storage.*** Some private vendors store data and provide an API that allows authorized IOO users to access the CV data, but local storage may be preferred by the agency in order to ensure access after a contract period ends. Agency storage preferences should be addressed as part of procurement.

- **Privacy.** Privacy is addressed by vendors either aggregating (e.g., average speeds or travel times for a segment) or anonymizing (i.e., no individual identities, license plate, or vehicle information) the data such that any individual vehicle data no longer contains any PII. Many vendors include privacy policies on their websites to detail how PII is protected, used, and shared, if at all; some note that they do not process, use, or distribute PII, while others detail a very stringent privacy policy regarding PII that notes how PII is protected and secured.
- **Security.** Security is typically addressed in a way that is consistent and not distinct from the standard security processes or procedures in place to protect all IOO data and data exchanges. Some vendors detail data security procedures on their websites; for example, multiple private vendors' websites note that data is protected at all stages of the data lifecycle using industry standards for security.
- **Communications.** Communications approaches can vary based on whether data is provided as a historic data set or in real-time/near real-time. Specifically, CV data may be provided by a "push" of the data by the vendor to an IOO web address, a "pull" of the data from the vendor web address by the IOO, or an IOO-initiated query on the vendor web interface.

4.6.1.2 Summary of Aggregation Sources

Aggregation sources offer similar services to those described above for private vendors, particularly since data from one or more sources is included in the portal. One potential advantage of aggregation sources is that data provided to IOOs may be customized and ready for immediate use for specific IOO applications, given the data processing and data fusion tools that the portals use to clean and prioritize data from multiple sources.

- **Access and Storage.** Aggregation portals can provide remote access and storage capabilities in a "one stop shop," or directly push the processed data to IOO systems so that the IOO does not need to invest in in-house processing and storage capabilities.
- **Privacy.** Aggregation portals may provide an additional layer of protection for PII by conducting data processing to de-identify or anonymize CV data which would otherwise need to be done by the IOO.
- **Security.** As part of the provided service, aggregation portals may provide some additional data security services that would otherwise be needed at the IOO.
- **Communications.** Communications approaches for aggregation portals are similar to those described for private vendor sources above in that CV data may be provided by a "push" of the data by the aggregation source to an IOO web address, a "pull" of the data of the data from the aggregation source web address by the IOO, or an IOO-initiated query on the aggregation web interface.

4.6.1.3 Summary of Direct Vehicle Sources

Direct vehicle sources provide individual vehicle data that is communicated to IOOs using pre-defined standardized messages through vehicle broadcasts that are received by roadside infrastructure. In 2016, the National Highway Traffic Safety Administration (NHTSA) issued a Notice of Proposed Rulemaking on V2V communications technologies that would have mandated new light vehicles broadcast the BSM at a frequency of 10 times per second (10 Hz) to support V2V safety applications. This rulemaking never went into effect and there is no mandate; however, there have been ongoing actions by OEMs to advance the deployment of OBUs to broadcast BSMs at 10 Hz. There are examples of vehicle models broadcasting BSMs but a timeline of expansion of these is not defined at this time. There are also examples (noted in the use cases) of IOOs deploying OBUs to broadcast BSMs, either on their fleet vehicles or through partnerships with private companies to deploy CV functionality to broadcast BSMs in private vehicle using aftermarket devices. IOOs that receive these OBU broadcasts can process the messages and use the data contained in each message or aggregate messages to derive data about vehicle speeds, travel patterns, and roadway conditions, for instance.

Another example of direct vehicle data is authorized fleet vehicles broadcasting signal request messages (SRMs) to request signal priority (e.g., transit vehicles behind schedule) or signal preemption (e.g., emergency vehicles or snowplows).

Likewise, the SAE J2945/C data standard could allow for probe messages from individual vehicles to provide either isolated data (e.g., current vehicle speed) or aggregated information (average speed over the past timeframe). The SAE J2945/C standard allows probe messages to be communicated both from OBU to RSU as well as using network cellular communications.

CV direct vehicle data are compared to the requirements in the following bullets:

- **Access and Storage.** IOO systems at the roadside have direct access to receive and process the BSMs and SRMs and use the data in the messages for applications locally. However, access to the data received at RSUs by staff or central systems would require backhaul communications and may be supplemented by MEC at the roadside to pre-process the data to reduce the volume of data that needs to be sent to the back office. Given the potential volume of data from every CV broadcasting BSMs 10 times per second, it will be important for IOOs to develop plans for data retention and processing that aggregate data or purge unnecessary data elements to reduce storage requirements. Table 10 illustrates this point with projections from a USDOT study, Integrating Emerging Data Sources into Operational Practice: State of the Practice Review, on the growth in agency data storage needs from 2016 through 2026.

Table 10: Summary of data storage for a typical agency (Source: Integrating Emerging Data Sources into Operational Practice: State of the Practice Review)

Source	Data Volume per day per device	2016	2026 Estimate
Connected Travelers via 3rd party/opt-in app	~500KB	1.8TB	370TB
CVs (commercial)	~1KB per segment	5TB	294TB
CVs (commercial – future)	~500KB (non-aggregated)	~Zero	365TB
CVs (public/dedicated short-range communications)	~600KB per interaction + 500KB (probe)	~Zero	6546TB
Connected infrastructure	~173MB (signal), ~5.8MB (other device), ~4GB (closed-circuit television)	38TB	4244TB
Other sources	~5.4GB (3D Video)	1TB	4.6TB
Total		45.8TB	11823.6TB

- **Privacy.** Privacy is built into the SAE J2735 standard; for example, the BSM does not directly identify individuals or specific vehicles and is broadcast in a very limited geographical range.[15] Privacy concerns may increase as IOOs compile large amounts of BSMs by enabling travel patterns to emerge from repetition vehicle location capture. However, there are examples of approaches that IOOs use to minimize risks of privacy outlined in Table 6 (and additional details are documented in the Ohio DOT comprehensive and detailed Security and Privacy Requirements).

- **Security.** Security is addressed by the SCMS certificates to authenticate any messages received from CVs. Additionally, standard IOO practices used to encrypt and secure storage and communications for any IOO data sources should be applied.

- **Communications.** Vehicle direct data is received at roadside equipment with V2X communications, and then communicated to the IOO back office via backhaul communications or cellular, depending on what communications mechanisms are available at the roadside.

4.6.2 IOO to Vehicle Sources and Data Delivery Mechanisms

The requirements for IOO to vehicle data are accomplished by a combination of data sources IOOs possess and CV delivery approaches to delivering them. That is, IOOs are a single source of their data that is collected, managed, and stored either in the field or the back-office, and then transmitted to vehicles via three primary mechanisms, as shown in Figure 14, and discussed in the bullets below.

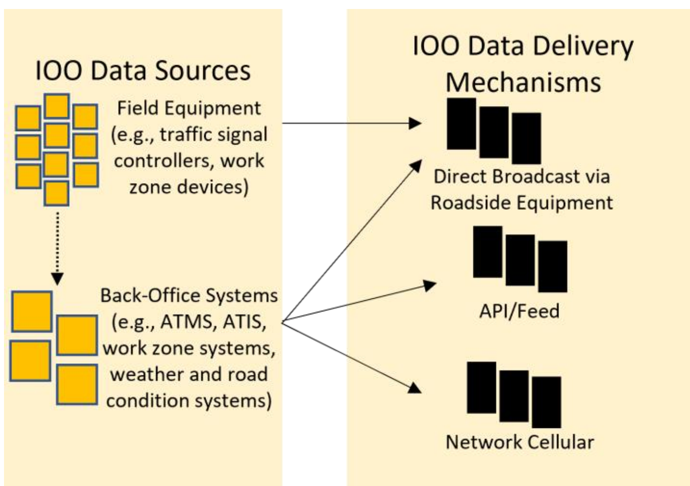


Figure 14: IOO data sources and mechanisms for transmitting CV data

1. **Direct.** Data in this category is broadcast directly to CVs from the IOO via roadside equipment using V2X communications. CVs in the vicinity of the broadcast and operating OBUs receive this data in real-time as standardized messages. The OBUs process the data and the data is used for various applications that generate appropriate responses (e.g., direct vehicle actions, in-vehicle messages to drivers, or generation and broadcast of messages back to the roadside equipment).

2. **API/Feed.** Data from IOO central systems (e.g., central signal controller systems, advanced traveler information systems (ATIS), ATMS, work zone systems) are provided to internet accessible locations where private vendors, third party application providers, fleet operators, or

OEMs may ingest the data and communicate it wirelessly to CVs. Note that some of the data contained within IOO back-office systems includes data collected at the roadside by field equipment. Examples of IOO data broadcast from central systems like ATMS/ATIS include road conditions, congestion and incident information, and work zone data (e.g., WZDx data feeds).

3. **Direct Network Cellular.** There are examples where IOOs work directly with network cellular providers for communication with vehicles. Arizona DOT piloted 'virtual RSUs' where data was shared with the cellular carrier who then delivered the data to vehicles based on their location and the defined geo-fence for the data. Another example of direct network cellular would be an IOO operating an https server that allows CVs to ping the http server and receive Data Probe Message Requests (i.e., requests for data based on current conditions and vehicle location). The vehicle could then send the data to the http server using the SAE Data Probe standard. Additionally, the https server could reply to vehicle requests by sharing standardized messages.

Current and anticipated data offerings for these types of IOO to vehicle data delivery sources are mapped to the CV data categories in Table 11.

Table 11: Types of IOO to vehicle data sources by CV data category

Types of IOO to Vehicle Data Sources	CV Data Categories		
	#6 Low Latency Infrastructure Safety	#7 Real-time Infrastructure Data	#8 Infrastructure Data
Direct Broadcasts via Roadside Equipment	✓	✓	✓
API/Feed		✓	✓
Direct Network Cellular		✓	✓

4.6.2.1 Summary of Direct Broadcasts via Roadside Equipment Data Sources

IOO field equipment generate data comprised as messages (e.g., CIs generate SPaT messages) and back-office systems (e.g., TMS, ATIS) generate messages describing conditions (e.g., work zones using WZDx, driving conditions using TIM messages). This data is then broadcast to CVs via roadside equipment in generally the same manner as described for direct vehicle sources above. This IOO data has the potential to improve mobility and safety, depending on number of CVs available to receive and use this data in on-board applications.

- **Access.** Data broadcast from roadside equipment is available to any CVs that are equipped with on-board units to receive and process the data.
- **Privacy.** No privacy concerns are present, since broadcast data is public, IOO-generated data.
- **Security.** Security is addressed with SCMS certificates attached to messages that broadcast both from vehicles and from RSUs. The SCMS enables recipients of direct CV data to verify certificate signatures to ensure that data is from a trusted source. Agency policies and procedures for securing backhaul communications, back-office equipment, and agency data are also applicable when data originates at the IOO back office.

- **Communications.** Roadside equipment data is broadcast to CVs with V2X communications. Note that some of this data may originate at the IOO back office and be sent to the roadside equipment via backhaul communications or cellular, depending on what communications mechanisms are available at the roadside.

4.6.2.2 Summary of API/Feed Data Sources

IOO central systems contain a variety of data that may be valuable for CVs. A central traffic signal control system may contain SPaT data that can be used across the entire IOO network by CVs for safety or fuel efficiency applications that do not have strict latency requirements, and freight carrier fleets may benefit from size and weight restrictions data that is updated for current construction activities. A common approach used to share data for traveler information has been the use of API/Feed posts of infrastructure data. This approach works for communicating CV data to private vendors and OEMs that have wireless communications capabilities to send data to the vehicles.

- **Access.** Data from API/Feed sources are generally provided at an IOO data exchange website, portal, or data feed that private vendors can access. IOOs may require private vendors and other private vendor users to register by creating an account and accepting IOO data use policies, for example, before being able to access and use the data.
- **Privacy.** No privacy concerns are present since provided data is public, IOO-generated data.
- **Security.** Security is addressed via standard IOO policies and procedures to encrypt communications and storage, as well as login or account requirements to authenticate data recipients.
- **Communications.** IOO central systems data is provided via the internet.

4.6.2.3 Summary of Direct Network Cellular

Direct network cellular examples are still pilot tests or concepts as no existing operational system was identified in the research. However, as the virtual RSU concept advances, as the SAE probe data standard is tested and eventually implemented, and finally, as the RDI is defined and implemented, this source of infrastructure data is expected to expand.

- **Access.** Access should be available wherever network cellular communications are operated and the IOO is providing data requests or data sharing.
- **Privacy.** No privacy concerns are present, since provided data is public, IOO-generated data.
- **Security.** Security is addressed via standard IOO policies and procedures to encrypt communications and storage, as well as login or account requirements to authenticate data recipients.
- **Communications.** Direct network cellular will rely on network cellular communication to the vehicles.

4.7 Summary of Model CV Data Architecture

Sections 4.1-4.6 contain the content identified by the CV PFS as most relevant to support IOO practitioners as they make decisions and implement processes for ingesting and/or providing CV data. The use cases, requirements, and concepts presented in these sections address both ***direct and indirect CV data communications*** with the focus on four key topics that are critical when considering CV data:

1. **Communications.** To be valuable to either the IOO or the vehicle, CV data must be communicated within required latency and using national standards governing both quality and format.
2. **Privacy.** The creation of CV data by vehicles (or individuals) carries with it the potential for personally identifiable information (PII) to be associated. Any PII must be protected.
3. **Security.** For vehicles and infrastructure systems to trust the data they receive, as well as to protect the IOO and CV systems, CV data requires security as the messages are created, communicated, and stored.
4. **Storage and Access.** The value that IOOs receive from CV data relies on it being accessible to the humans and systems that need it, as well as cost effective storage for future retrieval.

The intent is that the contents of these sections will be valuable to IOOs as they make internal decisions about managing CV data.

Appendix A: Recent Research and Development Projects Review

A review of recent (since 2015) documents related to CV data was conducted to inform the development of this Model CV Data Architecture. Collectively, these resources describe best practices and identify clarifications covering the breadth of topics and scenarios that IOOs need to understand to support their use of CV data. The list of identified resources reviewed for this effort are listed below. The resources reviewed included a variety of content and contained examples such as:

- Formally published documents of national research efforts (e.g., NCHRP reports, USDOT funded research, PFS research),
- Project deliverables (e.g., CV Pilot Site Resources),
- FHWA website briefings, and
- Private sector company web pages.

Table 12: Resources

Resource (Year)	Link
NCHRP Research Report 952: Guidebook for Managing Data from Emerging Transportation Technologies (2020)	https://www.trb.org/Main/Blurbs/180826.aspx#:~:text=The%20TRB%20National%20Cooperative%20Highway,can%20begin%20to%20shift%20%E2%80%93%20technically%2C
Framework for Managing Data from Emerging Transportation Technologies to Support Decision-Making (2020)	https://nap.nationalacademies.org/catalog/25965/framework-for-managing-data-from-emerging-transportation-technologies-to-support-decision-making
Integrating Emerging Data Sources into Operational Practice—Capabilities and Limitations of Devices to Collect, Compile, Save, and Share Messages from Connected and Automated Vehicles and Connected Travelers (2018)	https://rosap.ntl.bts.gov/view/dot/34985
Integrating Emerging Data Sources into Operational Practice: Opportunities for Integration of Emerging Data for Traffic Management and TMCs (2017)	https://rosap.ntl.bts.gov/view/dot/34175
Integrating Emerging Data Sources into Operational Practice: State of the Practice Review (2016)	https://rosap.ntl.bts.gov/view/dot/35143
NCHRP Synthesis 561: Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation (2021)	https://www.trb.org/Main/Blurbs/181749.aspx
ENTERPRISE PFS: Assess Speed Data for Traffic Management (2019)	https://enterprise.prog.org/projects/assess-speed-data-for-traffic-management/
Event Driven Configurable Messaging (EDCM) Concept of Operations – Version 1.6 (2020)	https://pronto-core-cdn.prontomarketing.com/2/wp-content/uploads/sites/2896/2020/09/EDCM_ConOps_Version_1_6_FINAL_20200828.pdf
NOCoE: Connected & Autonomous Vehicles (CAV) Data	https://data.transportationops.org/connected-autonomous-vehicles-cav-data

Resource (Year)	Link
CV Pilot Deployment Program: Driving Towards Deployment: Lessons Learned From the Design/Build/Test Phase (2018)	https://rosap.ntl.bts.gov/view/dot/37681
CV Pilot Deployment Program Phase 2: Data Privacy Plan: Wyoming (2017)	https://rosap.ntl.bts.gov/view/dot/32295
CV Pilot Deployment Program Phase 2: Data Privacy Plan: New York City (2016)	https://rosap.ntl.bts.gov/view/dot/32311
CV Pilot Deployment Program Phase 2: Data Privacy Plan: Tampa (2017)	https://rosap.ntl.bts.gov/view/dot/32034
NIST Big Data Interoperability Framework: Volume 4, Security and Privacy	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-4r2.pdfv
NIST Guide to Protecting the Confidentiality of Personally Identifiable Information (PII) (2010)	https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf
Data Needs Assessment for Making Transportation Decisions in Virginia (2015)	https://www.virginiadot.org/vtrc/main/online_reports/pdf/15-r23.pdf
Guide for State Department of Transportation Safety Data Business Planning (2017)	https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17047.pdf
NHTSA: Vehicle Data Privacy	https://www.nhtsa.gov/technology-innovation/vehicle-data-privacy#:~:text=NHTSA%20takes%20consumer%20privacy%20seriously%2C%20but%20does%20not%20have%20a,to%20enforce%20consumer%20privacy%20laws
Personal Data in your Car – National Automobile Dealers Association and the Future of Privacy Forum (2017)	https://fpf.org/wp-content/uploads/2017/01/consumerguide.pdf
Future of Privacy Form (FPF): Data and the Connected Car - Infographic (2017)	https://fpf.org/wp-content/uploads/2017/06/2017_0627-FPF-Connected-Car-Infographic-Version-1.0.pdf
Flow Lab IO privacy policy (2012)	https://flowlab.io/privacy
Connected Vehicles and Your Privacy (2019)	https://www.its.dot.gov/factsheets/pdf/Privacy_factsheet.pdf
How the USDOT is Protecting the Connected Transportation from Cyber Threats	https://www.its.dot.gov/factsheets/pdf/cybersecurity_factsheet.pdf
NHTSA Vehicle Cybersecurity	https://www.nhtsa.gov/technology-innovation/vehicle-cybersecurity
USDOT ITS Cybersecurity Research Program	https://www.its.dot.gov/research_areas/cybersecurity/index.htm
Utilization of CV Data to Support Traffic Management Decisions (2017)	https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/research/reports/fdot-bdv29-977-21-rpt.pdf
NCHRP Cybersecurity Implications of CV/AV Technologies on State and Local Transportations (2017)	https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4254
NCHRP Planning Data Needs and Collection Techniques for CV/AV Applications	https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4257
NCHRP Data Management Strategies for CV/AV Applications for Operations	https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4258

Resource (Year)	Link
Austroads AP-R672-22 Supporting Cloud Connected Road Users	https://austroads.com.au/publications/connected-and-automated-vehicles/ap-r672-22
Austroads Research Report AP-R662B-21 Road Authority Data for Connected and Automated Vehicles Module 2, Guidance for Incidents Data Provision to Connected and Automated Vehicles (2021)	https://austroads.com.au/publications/connected-and-automated-vehicles/ap-r662-21/media/AP-R662B-21_RADCAV_Module_2_Guidance_for_Incidents.pdf
Iteris ClearGuide	https://www.iteris.com/oursolutions/traffic-analytics-software/clearguide
Airbiquity	https://www.airbiquity.com/
Otonomo	https://otonomo.io/
Brand Motion	https://brandmotion.com/
Ceve	https://www.ceve.io/
Wejo	https://www.wejo.com/traffic-and-transportation-agencies
mobileye	https://www.mobileye.com/
Streetlight Data	https://www.streetlightdata.com/
Payver	https://blyncsy.com/payver/
INRIX	https://inrix.com/industries/public-sector/transportation-agencies/
i-Probe	https://i-probe-inc.com/services.html
ITS International: Connected Car Data Standard (2016)	https://www.itsinternational.com/its10/news/here-automotive-companies-move-forward-connected-car-data-standard
Markup - "Who Is Collecting Data from Your Car?" (2022)	https://themarkup.org/the-breakdown/2022/07/27/who-is-collecting-data-from-your-car
NHTSA Cybersecurity Best Practices for the Safety of Modern Vehicles (2022 release)	https://intelligenttransportationsocietyofamerica.cmail19.com/t/d-l-fujjic-tdirktjtth-b/
CV PFS Using Third Parties to Deliver I2V: Concept of Operations (2019)	https://engineering.virginia.edu/sites/default/files/Connected-Vehicle-PFS/Projects/3rd%20Party%20I2V/Task3-1%20SwRI-10-24479_ConOps_Final.pdf
CV PFS Using Third Parties to Deliver I2V: System Requirements (2019)	https://engineering.virginia.edu/sites/default/files/Connected-Vehicle-PFS/Projects/3rd%20Party%20I2V/Task3-2_System%20Requirements%20Walkthrough%20Using%20Third%20Parties%20to%20Deliver%20I2V_For%20Distribution.pdf
CV PFS Using Third Parties to Deliver I2V: Interface Control Document (2020)	https://engineering.virginia.edu/sites/default/files/Connected-Vehicle-PFS/Projects/3rd%20Party%20I2V/Task3_3%20CVPFS-Release_Version_1.0_Interface_Control_Document.pdf
CV PFS Using Third Parties to Deliver I2V: Consensus Recommendation (2020)	https://engineering.virginia.edu/sites/default/files/Connected-Vehicle-PFS/Projects/3rd%20Party%20I2V/Task3-4_CVPFS-Consensus_Recommendation-Release_1.0_05.pdf

References

1. National Academies of Sciences, Engineering, and Medicine. (2021). NCHRP Synthesis 561: Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation. Retrieved November 2022. https://doi.org/10.17226/26094.
2. DriveOhio. (2020). Ohio CV/AV Communications Plan Framework. Retrieved March 2023. https://drive.ohio.gov/wps/wcm/connect/gov/67d7f01b-5621-4861-b8e2-529bef8223c7/31-Comms+Plan-Optimized.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=ROOTWORKSPACE.Z18_K9I401S01H7F40QBNJU3SO1F56-67d7f01b-5621-4861-b8e2-529bef8223c7-nHyE80K.
3. SAE International. (2022). Requirements for Probe Data Collection Applications J2945/C_202206. Retrieved March 2023. https://www.sae.org/standards/content/j2945/c_202206/.
4. Connected Vehicle Pooled Fund Study. (2021). V2I Queue Advisory/Warning Applications: Concept and Design High-Level Design. Retrieved March 2023. https://engineering.virginia.edu/sites/default/files/Connected-Vehicle-PFS/Projects/V2I%20Queue%20Warning/612141-00002_CVPFS_QueueWarning_Task5_HLD_Final.pdf.
5. Crash Avoidance Metrics Partners LLC. (2021). Event-Driven Configurable Messaging (EDCM) Design & Development and Work Zone Queue Advisory/Warning (QA/QW) System. Retrieved March 2023. https://www.campllc.org/event-driven-configurable-messaging-edcm-design-development-and-work-zone-queue-advisory-warning-qa-qw-system/.
6. US United States Department of Transportation (USDOT). (2017). Vehicle-to-Infrastructure Safety Applications Project – Final Report. FHWA-JPO-18-606. Retrieved January 2023. https://rosap.ntl.bts.gov/view/dot/37212.
7. EcoDrive II I2V Connected Vehicle Pilot Project – City Fleet – Signalized Intersection Approach and Departure Optimization Application. Retrieved March 2023. https://tcdocs.ingeniumcanada.org/sites/default/files/2020-05/City of Ottawa - I2V Connected Vehicle Pilot Project - City Fleet - Signalized Intersection Approach and Departure Optimization Application.pdf.
8. National Academies of Sciences, Engineering, and Medicine. (2021). NCHRP Synthesis 952: Guidebook for Managing Data from Emerging Technologies for Transportation. Retrieved November 2022. https://nap.nationalacademies.org/login.php?action=guest&record_id=25844 .
9. United States Department of Transportation (USDOT). (2016). Integrating Emerging Data Sources into Operational Practice: State of the Practice Review. FHWA-JPO-16-424. Retrieved November 2022. https://rosap.ntl.bts.gov/view/dot/35143.
10. Research Report AP-R662B-21. Road Authority Data for Connected and Automated Vehicles Module 2 Guidance for Incidents Data Provision to Connected and Automated Vehicles.
11. National Institute of Standards and Technology (NIST). (2010). Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Doc 800-122. Retrieved November 2022. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf.

12. United States Department of Transportation (USDOT). (2016). CV Pilot Deployment Program Phase 2: Data Privacy Plan: New York City. Retrieved November 2022. https://rosap.ntl.bts.gov/view/dot/32311.

13. United States Department of Transportation (USDOT). (2017). CV Pilot Deployment Program Phase 2: Data Privacy Plan: Tampa. Retrieved November 2022. https://rosap.ntl.bts.gov/view/dot/32034.

14. United States Department of Transportation (USDOT). (2017). CV Pilot Deployment Program Phase 2: Data Privacy Plan: Wyoming. Retrieved November 2022. https://rosap.ntl.bts.gov/view/dot/32295.

15. USDOT Privacy Factsheet. (2019). Accessed March 2023. Available at: https://its.dot.gov/factsheets/pdf/Privacy_factsheet.pdf