

INTEGRATING EMERGING DATA SOURCES INTO OPERATIONAL PRACTICE

APPLYING BIG DATA TOOLS AND TECHNOLOGIES TO USE EMERGING DATA SOURCES

Agencies have the potential to collect, use, and share data from connected and automated vehicles (CAV), connected travelers, and connected infrastructure elements to improve the performance of their traffic management systems and traffic management centers (TMCs) and the facilities they manage. The integration of these techniques and approaches from other industries into these systems offers agencies the ability to explore how they could take advantage of and harness the potential of using and sharing data with connected and automated vehicles, connected travelers, and other sources. The Federal Highway Administration and the ITS Program have developed several resources which provided the basis for the information captured on this fact sheet to support agencies assessing how these big data techniques and approaches could be integrated their systems to collect, use, and share data with the sources of these emerging data.

FIGURE 1. BIG DATA (Source: wordle.net)



HOW CAN BIG DATA TOOLS AND TECHNOLOGIES BE USED BY TRAFFIC MANAGEMENT SYSTEMS AND TMCs?

Projections developed in this project indicate that if a typical agency retained all the information from connected vehicles, connected travelers, and connected infrastructure between now and 2026 the data storage requirements alone would amount to several Petabytes. Such a data store vastly exceeds what any traffic management agency currently stores and processes today. Similarly, the velocity of data capture at a typical agency will increase five-fold from 1TB per day to 5TB per day when connected vehicles make up 50% of the vehicle fleet. Agencies will need to assess what data they should collect, process, store, and share to enhance their traffic management systems and TMCs. Tools and technologies from commercial big data systems including real-time stream processing, NoSQL databases, and cloud-based analytics will be required for traffic management centers to process, store, and analyze these sizes of datasets and velocity of data capture. Similarly, new capabilities of roadside equipment (RSE) will be needed to collect, process, store, and share relevant information with the TMC, other partner agencies, connected and automated vehicles, and connected travelers.

OPPORTUNITIES FOR INTEGRATING EMERGING DATA SOURCES INTO TRAFFIC MANAGEMENT SYSTEMS AND TMCs

Many categories of TMC functions can be enhanced with data from connected vehicles and connected travelers.

Real-time functions	Road Hazard Warnings	Speed Warnings	Intersection Collision Avoidance	Probe Data Collection	Electronic Payments
Near-real-time functions	Ramp Metering	Signal Control & Metering	Lane Management	Personalized Traveler Information	Weather Monitoring
Offline functions	Performance Measurement	Asset Management	Emissions Monitoring		

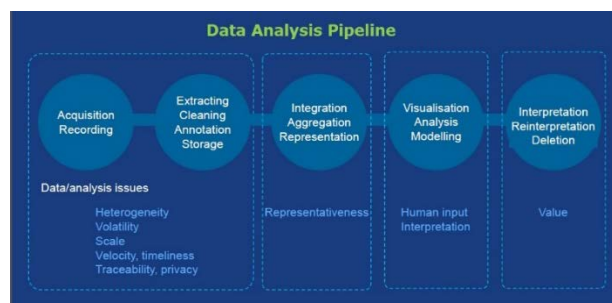
Different approaches for data acquisition, processing, and storage will be needed for real-time, near-real-time, and offline functions of traffic management systems and TMCs. A matrix mapping big data techniques, methods, and tools to the typical functions of a system has been developed to assist agencies assessing what methods may be appropriate to consider. Much of the data processing burden will need to be handled on RSEs and intermediate processors between the field and the TMC when the level of data from connected vehicles and connected travelers becomes significant.

OPPORTUNITIES TO COLLECT, COMPILE, SAVE AND SHARE DATA WITH CONNECTED AND AUTOMATED VEHICLES AND CONNECTED TRAVELERS

Roadside Equipment (RSE) is the field device responsible for collecting, processing, storing, and sharing data from connected and automated vehicles with the TMC. There are currently no defined standards for how an RSE should interface to a TMC and traffic management systems. As more and more connected vehicles arrive in our networks, the quantity and speed at which the raw data is arriving will increase. Summary statistics and various types of aggregations will be sent by the edge processor applications to the TMC for near-real-time and offline functions. A common aggregation function that will serve many TMC functions and traffic management systems is the collection of basic safety message (BSM) data from individual vehicles into trajectories. Significant geofencing (i.e. the process of locating each BSM inside a specific geographic boundary, such as the northbound direction of the freeway), and sorting and selecting among the potentially many types of messages will be need to be done by the RSE to determine which data are relevant to which TMC function.

PLANS AND PROCESSES TO ENHANCE CURRENT TRAFFIC MANAGEMENT SYSTEMS

FIGURE 2. DATA ANALYSIS PIPELINE (Source: International Transport Forum)



There are many opportunities for agencies to integrate big data tools and technologies to collect, process, store, and share data from CAVs and connected travelers for enhancement of traffic management and TMCs. Agencies can gain experience with big data tools and aggregation technologies through pilot and research projects, perhaps focused on data already being collected, such as data from traffic signal systems or freeway management systems. Methodologies for data acquisition, aggregation, and

integration of connected vehicles and connected traveler data with traffic management systems will be developed by early adopters and USDOT. These approaches will be designed and developed in the next 3-5 years. In perhaps ten years or more, new generation traffic management systems will likely become available that directly manipulate the connected vehicle and connected traveler streams.

FOR MORE INFORMATION:

- <https://www.its.dot.gov/v2i/>
- [https://ops.fhwa.dot.gov/Project FAQ: html link here](https://ops.fhwa.dot.gov/Project%20FAQ%3A%20html%20link%20here)
- FHWA-JPO-16-424 Integrating Emerging Data Sources into Operational Practice: State of the Practice Review
- FHWA-JPO-18-625 Opportunities for Integration of Emerging Data for Transportation Systems Management and Operations
- FHWA-JPO-18-626 Capabilities and Limitation of Devices to Collect, Compile, Save and Share CV Related Data

