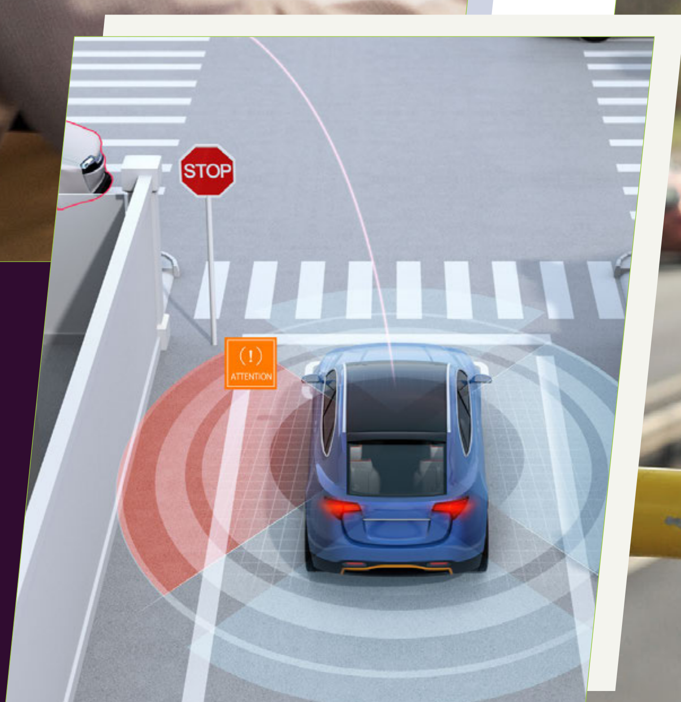
U.S. Department of Transportation
**Federal Highway Administration**

# Considerations of Current and Emerging Transportation Management Center Data




## NOTICE



## QUALITY ASSURANCE STATEMENT

## Technical Report Documentation Page

| | | |
|---|---|---|
| **1. Report No.**<br>FHWA-HOP-18-084 | **2. Government Accession No.** | **3. Recipient's Catalog No.** |
| **4. Title and Subtitle**<br>Considerations of Current and Emerging Transportation Management Center Data | | **5. Report Date**<br>July 2019<br>**6. Performing Organization Code** |
| **7. Authors**<br>Michael L. Pack, Nikola Ivanov, Jocelyn K. Bauer, Elizabeth Birriel | | **8. Performing Organization Report No.** |
| **9. Performing Organization Name and Address**<br>Leidos<br>11251 Roger Bacon Drive<br>Reston, VA 20190<br><br>Michael Pack, LLC, Columbia, Maryland 20145 | | **10. Work Unit No. (TRAIS)**<br>**11. Contract or Grant No.**<br>Contract No.<br>DTFH61-16-D-00053 |
| **12. Sponsoring Agency Name and Address**<br>Federal Highway Administration<br>U.S. Department of Transportation<br>1200 New Jersey Avenue, SE<br>Washington, DC 20590 | | **13. Type of Report and Period Covered**<br>Final Report<br>**14. Sponsoring Agency Code**<br>HOTM |
| **15. Supplementary Notes**<br>Jimmy Chu, FHWA Task Order Contract Officer's Representative | | |
| **16. Abstract**<br>This report examines some of the fundamental aspects of using emerging data from third parties—understanding what is becoming available at the time of this publication, how it is collected, the business models used by the companies that sell it, and possible data use cases. Contract considerations for working with private sector data providers are discussed. The document also highlights business questions an agency can explore to determine the potential value of internal data assets, and summarizes the private sector's desire for agency-owned data. | | |
| **17. Key Words**<br>Emerging data, third-party data, transportation management centers, business models, data sharing, crowdsourced data, connected vehicle data. | | **18. Distribution Statement**<br>No restrictions. |

| | | | |
|---|---|---|---|
| **19. Security Classif. (of this report)**<br>Unclassified. | **20. Security Classif. (of this page)** Unclassified | **21. No of Pages**<br>156 | **22. Price**<br>N/A |

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

## SI* (MODERN METRIC) CONVERSION

| FACTORS APPROXIMATE CONVERSIONS TO SI UNITS | | | | |
|---|---|---|---|---|
| SYMBOL | WHEN YOU KNOW | MULTIPLY BY | TO FIND | SYMBOL |
| **LENGTH** | | | | |
| in. | inches | 25.4 | millimeters | mm |
| ft | feet | 0.305 | meters | m |
| yd | yards | 0.914 | meters | m |
| mi | miles | 1.61 | kilometers | km |
| **AREA** | | | | |
| $in.^2$ | square inches | 645.2 | square millimeters | $mm^2$ |
| $ft^2$ | square feet | 0.093 | square meters | $m^2$ |
| $yd^2$ | square yards | 0.836 | square meters | $m^2$ |
| ac | acres | 0.405 | hectares | ha |
| $mi^2$ | square miles | 2.59 | square kilometers | $km^2$ |
| **VOLUME** | | | | |
| fl oz | fluid ounces | 29.57 | milliliters | mL |
| gal | gallons | 3.785 | liters | L |
| $ft^3$ | cubic feet | 0.028 | cubic meters | $m^3$ |
| $yd^3$ | cubic yards | 0.765 | cubic meters | $m^3$ |
| NOTE: volumes greater than 1,000 L shall be shown in $m^3$ | | | | |
| **MASS** | | | | |
| oz | ounces | 28.35 | grams | g |
| lb | pounds | 0.454 | kilograms | kg |
| T | short tons (2000 lb) | 0.907 | megagrams (or "metric ton") | Mg (or "t") |
| **TEMPERATURE (exact degrees)** | | | | |
| °F | Fahrenheit | 5 (F-32)/9 or (F-32)/1.8 | Celsius | °C |
| **ILLUMINATION** | | | | |
| fc | foot-candles | 10.76 | lux | lx |
| fl | foot-Lamberts | 3.426 | candela/$m^2$ | cd/$m^2$ |
| **FORCE and PRESSURE or STRESS** | | | | |
| lbf | poundforce | 4.45 | newtons | N |
| lbf/$in.^2$ | poundforce per square inch | 6.89 | kilopascals | kPa |

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

## SI* (MODERN METRIC) CONVERSION (continued)

| APPROXIMATE CONVERSIONS TO SI UNITS | | | | |
|---|---|---|---|---|
| SYMBOL | WHEN YOU KNOW | MULTIPLY BY | TO FIND | SYMBOL |
| **LENGTH** | | | | |
| mm | millimeters | 0.039 | inches | in. |
| m | meters | 3.28 | feet | ft |
| m | meters | 1.09 | yards | yd |
| km | kilometers | 0.621 | miles | mi |
| **AREA** | | | | |
| $mm^2$ | square millimeters | 0.0016 | square inches | $in.^2$ |
| $m^2$ | square meters | 10.764 | square feet | $ft^2$ |
| $m^2$ | square meters | 1.195 | square yards | $yd^2$ |
| ha | hectares | 2.47 | acres | ac |
| $km^2$ | square kilometers | 0.386 | square miles | $mi^2$ |
| **VOLUME** | | | | |
| mL | milliliters | 0.034 | fluid ounces | fl oz |
| L | liters | 0.264 | gallons | gal |
| $m^3$ | cubic meters | 35.314 | cubic feet | $ft^3$ |
| $m^3$ | cubic meters | 1.307 | cubic yards | $yd^3$ |
| **MASS** | | | | |
| g | grams | 0.035 | ounces | oz |
| kg | kilograms | 2.202 | pounds | lb |
| Mg (or "t") | megagrams (or "metric ton") | 1.103 | short tons (2000 lb) | T |
| **TEMPERATURE (exact degrees)** | | | | |
| °C | Celsius | 1.8C+32 | Fahrenheit | °F |
| **ILLUMINATION** | | | | |
| lx | lux | 0.0929 | foot-candles | fc |
| $cd/m^2$ | candela/$m^2$ | 0.2919 | foot-lamberts | fl |
| **FORCE and PRESSURE or STRESS** | | | | |
| N | newtons | 0.225 | poundforce | lbf |
| kPa | kilopascals | 0.145 | poundforce per square inch | $lbf/in^2$ |

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

# TABLE OF CONTENTS

## TABLE OF CONTENTS (continued)

# TABLE OF CONTENTS (continued)

# LIST OF FIGURES

## **LIST OF FIGURES** (continued)

# LIST OF TABLES

# LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| AAR | after-action review |
| AASHTO | American Association of State Highway and Transportation Officials |
| ABS | anti-lock braking system |
| API | application programming interface |
| ATMS | advanced traffic management system |
| ATSPM | automated traffic signal performance measure |
| AVL | automated vehicle location |
| BSM | basic safety message |
| CAD | computer-aided dispatch |
| CANBus | controller area network bus |
| CATT | Center for Advanced Transportation Technology |
| CCP | Connected Citizens Program |
| CCTV | closed-circuit television |
| CHART | Coordinated Highways Action Response Team |
| CHP | California Highway Patrol |
| CV | connected vehicle |
| DMS | dynamic message sign |
| DMV | department of motor vehicles |
| DOT | department of transportation |
| DSRC | dedicated short-range communications |
| DSS | decision support system |
| EMDSS | enhanced maintenance and decision support system |
| FDOT | Florida Department of Transportation |
| FHWA | Federal Highway Administration |
| GIS | geographical information system |
| GPS | Global Positioning System |
| GTFS | general transit feed specification |
| HAR | highway advisory radio |
| ICC | I-95 Corridor Coalition |
| ICM | integrated corridor management |
| IGA | intergovernmental agreement |
| INFLO | intelligent network flow optimization |
| IT | information technology |
| ITS | intelligent transportation systems |
| JSON | Javascript object notation |
| LIDAR | light detection and ranging |
| MAC address | media access control address |
| MAP-21 PM3 | Moving Ahead for Progress in the 21st Century third performance measure rule |

# LIST OF ABBREVIATIONS AND ACRONYMS (continued)

| | |
|---|---|
| MATOC | Metropolitan Area Transportation Operations Coordination |
| MAW | motorist advisory and warning |
| MPO | metropolitan planning organization |
| MSA | metropolitan statistical area |
| NPMRDS | National Performance Management Research Data Sharing |
| O-D | origin-destination |
| OEM | original equipment manufacturer |
| OSM | OpenStreetMap |
| O&M | operations and maintenance |
| PANYNJ | Port Authority of New York and New Jersey |
| PFS | pooled fund study |
| RFID | radio frequency identification |
| RFP | request for proposal |
| RGB | Red, Green, Blue |
| RITIS | Regional Integrated Transportation Information System |
| RSU | roadside unit |
| RTC | realtime closure |
| RTIMIS | Realtime Traffic Incident Management Information System |
| RWIS | road weather information system |
| SaaS | software as a service |
| Skyline | Skyline Technology Solutions |
| SPaT | signal phase and timing |
| TMC | transportation management center |
| TNC | transportation network company |
| TOC | traffic operations center |
| Total Traffic | Total Traffic & Weather Network |
| TSMO | transportation systems management and operations |
| UCAR | University Corporation for Atmospheric Research |
| UDOT | Utah Department of Transportation |
| UPS | United Parcel Service |
| USDOT | United States Department of Transportation |
| VDOT | Virginia Department of Transportation |
| VMT | vehicle miles traveled |
| VSP | Virginia State Police |
| XD | extreme definition |
| XML | extensible markup language |

# EXECUTIVE SUMMARY

This report examines the state of the practice and emerging trends in data, business models, and applications for transportation management centers (TMCs). It provides an understanding of what is available, how it is collected, the business models used by the companies that sell it, acceptable uses of the data, and possible data use cases. It also includes contract considerations for working with private sector data. This document further describes a process by which agencies can determine the potential value of internal data assets while also summarizing the private sector's desire for it.

## DATA

The first section of this report serves as a reference for agencies who are looking to expand capabilities, and explains how emerging data sources will impact operations. Several new types of private sector data are discussed including connected vehicle data, realtime trajectory data, realtime turning movement data, high-resolution traffic signal data, and crowdsourced incident and congestion data. The private sector obtains its data largely from connected devices including cell phones, navigation systems, location-based services, satellites, and other devices embedded in vehicles. This collection strategy allows for wide-area coverage without the need to install dedicated road-side equipment. This ultimately reduces maintenance costs while simultaneously producing larger quantities of data, which is often of higher quality than that collected from legacy sensor-based systems. The private sector has yet to implement a way to produce ubiquitous realtime volume datasets, but progress is being made. The need for dedicated volume sensors will likely diminish over the next 5 years.

The data described in this report brings opportunities for new or improved transportation management capabilities, but there are trade-offs for each type of data. For example, crowdsourced data is susceptible to redundancy and unreliable information such as the generation of false-positive events, but it tends to provide broader coverage than individual agencies may be able to afford with limited resources and jurisdictional responsibilities. Crowdsourced data can also be timelier than agency-generated data.

This report also discusses how these emerging data sources may differ from traditional data sets already in use by TMCs. It provides a process by which agencies can go about determining the value of agency-owned data assets, and discusses the implications of monetizing the assets. The majority of the agencies interviewed for this report were unsuccessful at justifying cost recovery, shared revenue, and sponsorship models. Most private sector data providers still have a desire for agency-produced datasets; however, the type of data deemed most valuable has changed. Volume and speed sensor data from public agencies are no longer as valuable because of the growing penetration of probes. The private sector still values planned event and lane closure data, computer-aided dispatch (CAD) data, driving records or department of motor vehicle data, live closed-circuit television feeds, and parking availability.

## BUSINESS MODELS AND PROCUREMENT STRATEGIES

Private sector data providers use a wide variety of business models and delivery mechanisms to provide data to government agencies. The differences between these models can have sweeping impacts on pricing, contracting, and agency use. Providers may price data by how the agency intends to use the data, such as for realtime use only (no archiving), archive only (historic data), and single-purpose use. The data may be sold for use only by that agency or shared with partner agencies. The agency may purchase all rights to use the data in perpetuity. Other business models include exchanging public sector for private sector data and potentially requiring the agency to promote the private sector's products and employ visible attribution.

When procuring data for use in TMCs, agencies should be prepared to address key policy considerations in selecting a data use agreement. The selection of usage terms by a public agency typically has long-term impacts on its capabilities and costs. These include:

- Who can use the data within the agency?
- Who can the data be shared with and in what format – raw, aggregated, anonymized?
- What, if any, attribution is required when publishing data and data products?
- What type of processing and publications are allowed – research, reports, public traveler information, operational use only, etc.?
- What is the expected level of accuracy and reliability of data?
- How is the data validated?

When working with partner agencies to share internal data (like public safety CAD), the agency will need to focus their attention more on reintegration strategies that make the data more usable within the agency's existing advanced traffic management system platform, reduce duplication, reduce operator fatigue/overload, and ensure the agency staff (or systems) can properly interpret the information within the CAD messages.

## POLICY AND CONTRACT CONSIDERATIONS

The final sections of the report focus on the importance of negotiating acceptable use terms and conditions and understanding an agency's internal needs across different departments. Prior to the procurement process, agencies should communicate frequently across departments and with any relevant partners to understand the needs and potential future uses of the data. Sample data sharing agreements are provided for comparison. The report also talks about the pros and cons of leveraging the private sector for hosting and managing data as compared to developing internal data management capabilities. It is advisable to build reasonable and fair terms that determine what happens when or if data is deemed to be of poor quality, missing, etc.; on what performance metrics the data and contractor will be evaluated; and the implications of underperformance.

Success in reaching consensus on acceptable use terms and conditions is a result of a meaningful effort to understand agency internal needs across different departments, potential uses and interactions with partner agencies, private sector partners, academia, and the public. Prior to the procurement process, agencies should frequently communicate both internally and with any relevant partners to understand the needs and potential future uses of the data.

# CHAPTER 1. **INTRODUCTION**

## PURPOSE

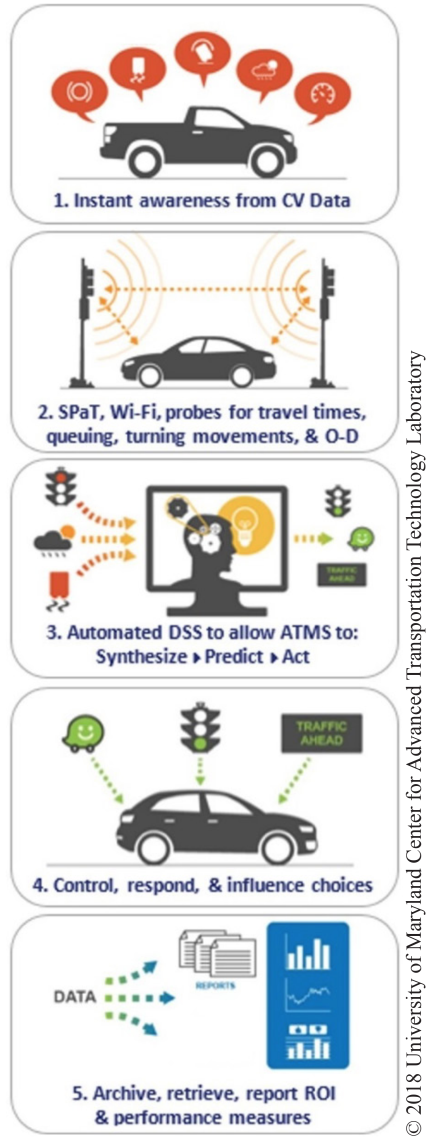
This report documents a Transportation Management Center Pooled Fund Study research project and provides a detailed reference on the concepts, business models, methods, processes, techniques, and other issues related to realtime traffic data collection and dissemination. The audience for this report is transportation management center (TMC) managers and their public sector partners. The project identified, researched, and synthesized an assessment of the following data collection and dissemination factors:

- Private sector business models.
- Public-private partnerships and agreements.
- Private sector data that can supplement traffic data collected by public agencies.
- Public agency data with a focus on value and models when sharing with the private sector.
- Emerging data sources such as crowdsourcing and connected vehicles (CVs).

## BACKGROUND

TMCs constantly search for new and innovative data sources that can improve guidance for operations decisions, provide better predictive capabilities, and enhance the safety and mobility of travelers. Figure 1 shows one of many approaches agencies could use in the future to integrate new, private sector data and agency-deployed sensors to enhance operations. The process depicted describes the following activities:

- Data acquisition, shown in boxes 1 and 2, incorporates the collection of CV data from third parties followed by Wi-Fi re-identification, signal phase and timing (SPaT), and other emerging data products (see chapter 2).
- Data could then be integrated into an advanced transportation management system (ATMS), decision support system (DSS), etc., which interprets and synthesizes the data to predict future conditions, provides greater insights, etc., as shown in box 3.




O-D = origin-destination
ROI = return on investment
CV = connected vehicle
SPat = signal phase and timing
DSS = decision support system
ATMS = advanced transportation management system

**Figure 1. Graphic. How agency traffic management centers integrate private sector data.**

**Source: University of Maryland, Center for Advanced Transportation Technology Laboratory.**

- The resulting system responses could include traffic control, response to changing conditions or incidents, feedback loops, and other actions that influence driver choices on routing and departure times, as represented by box 4.
- Lastly, a data archive for measuring outcomes, reporting, and process improvement is shown in box 5.

The diagram, while convincing, glosses over many of the critical details. The workflow seems simple enough, but many factors can derail efforts to procure and utilize new data or technologies, all of which can heavily influence operations program success, agency budgets, and the public.

This report looks at some basic aspects of using emerging data from third parties—understanding what is available, how it is collected, the business models used by the companies that sell it, acceptable use of the data, and possible data use cases. Details of these aspects are rarely reported in the public domain, making location of reliable data difficult for agencies. This limits agencies' ability to define successful operations strategies and make sound investment decisions.

## SOURCES OF INFORMATION AND METHODOLOGY

The extensive literature review and the gathering of key information for this study came largely from in-person or phone interviews with private sector data providers, systems integrators, and the staff of departments of transportation, universities, and non-profit coalitions. The very nature of this report has the potential to expose key business and agency practices that are both proprietary and sensitive. Most of the private sector and agency interviewees declined the offer of attribution and requested that specific details concerning the exact formats of their data remain confidential.

# CHAPTER 2. EMERGING DATA SOURCES

The private sector is constantly evolving and innovating—trying to create the next data product or service that will give businesses a competitive advantage. While it is difficult to predict the future, this section includes data that is becoming available now or will likely become available in the next several years. Not every private sector provider wants to reveal its anticipated products or schedules, but table 1 provides a list of newer and emerging data sources that not all agencies have used yet.

**Table 1. Newer and emerging data sources.**

| Data | Potential Planning/Decisionmaking Use |
|---|---|
| Crowdsourced[1] Incident and Congestion Data | Crowdsourced incident and event data can help agencies quickly identify all sorts of safety and congestion issues on the roadway without having to invest in additional sensors, cameras, or other costly data from third parties. Preliminary studies have shown Waze crowdsourced data to be comparable in speed and quality to public safety computer-aided dispatch data, and is particularly useful in rural areas where intelligent transportation system coverage is usually limited. Agencies are beginning to leverage this data as a supplemental event/incident data sourced. |
| Roadside BSM Collection | Some modern roadside equipment can now collect basic safety messages (BSM) sent from equipped vehicles. While the percentage of vehicles equipped to broadcast BSM is currently very low, this data has the ability to eventually be highly useful at signalized intersections, work zones, and other locations where split-second decisions have safety implications. Planners, policymakers, and engineers can mine this data (with other datasets) to understand driver behavior and to identify roadway safety improvements. |
| Realtime Trajectory Data | Trajectory data is similar to origin-destination (O-D) data, except that it also provides waypoints (breadcrumb trails). For any trip, data is being continually relayed back to the operations center or advanced traffic management system platform that tells operators which routes travelers are taking to get to their destinations, how fast they are moving, and if that route is normal or abnormal for that type of vehicle, time of day, etc. Operators can use the data to conduct after-action reviews for significant events to understand the impacts of events and road closures. When archived, this data can augment or completely replace traditional trip and O-D studies used by planners and modelers. |
| Crowdsourced Mapping Data | Directional, navigable mapping data created and edited by the public can be used to supplement (and sometimes improve upon) State centerline files. When used in a planning environment, this data can more quickly update maps—often faster than commercial providers that typically update their maps quarterly with significant lag between when new roads are completed and when added to the system. |

**Table 1. Newer and emerging data sources. (continued)**

| Data | Potential Planning/Decisionmaking Use |
|---|---|
| Probe-Based Speed Data | Speed and travel time data from vehicles using navigation systems aids in the study of congestion trends, identification of problem locations, before and after evaluations, and project prioritization. |
| High-Resolution Map Data and Other Asset Management Systems | Companies are now touting their high-resolution mapping data, which includes extremely detailed information (down to the centimeter accuracy) about where curbs are located, how high the curbs are, where road markings are located, and where and what is on a road sign. Planners and policy makers use this data to understand infrastructure degradation, life cycle, and maintenance needs over time. |
| Wi-Fi/Bluetooth Re-identification | Installation of Wi-Fi and Bluetooth equipment at intersections or other decision points helps to better understand travel times and travel patterns on key corridors and arterials. The resulting data is a point-to-point travel time between the two sensors that is often more accurate than traditional sensors and/or probe data. |
| Credit Card Transactions (from Point-of-Sale Vendors) | Point-of-sale vendors are starting to sell data related to where credit card transactions are taking place, the commodities purchased, by whom, and more. This data can be used in realtime by transportation management centers (TMCs) to understand travel patterns, deviations from normal travel patterns, and during and after emergencies (like hurricanes and snow storms) as a surrogate measure of power failures and when businesses are open or closed. Planners can also archive the data and use it as a surrogate for origin-destination studies, trip analytics, and activity-based models. |
| Connected Vehicle Data from Telematics Providers | This data includes direct-from-vehicle measures and warnings, such as heavy-breaking events, traction-control engagement, wiper use, emissions data, temperature data, rollover and/or collision data, and even seatbelt use (in commercial vehicles). This data is available today from telematics providers in over 5 million vehicles in the United States. This data can be used by TMCs to supplement existing incident detection systems, to alert operators about adverse weather conditions such as slippery roads, and as a pre-event warning system. The data can then be archived and used by planners for many activities, including safety studies, congestion analytics, performance measures, emissions studies, and as a surrogate for ground-based weather data. |
| Realtime Turning Movement Data | When operators respond to incidents or lane closures, realtime turning movement data can help them understand what percentage of vehicles are taking alternate routes and the effectiveness that different communication strategies have in changing travel behavior. When archived, this data supplements trip surveys and can inform long-term planning decisions. |
| High-Resolution Signal Data | Some signal manufacturers are beginning to equip signal systems, communications, and logging equipment that collect and distribute highly precise signal phase and timing (SPaT) data, actuator data, and more. High-frequency collection of this information can facilitate signal-retiming efforts, determine where congestion is occurring, and provide better understanding of the signal retiming impacts. |

**Table 1. Newer and emerging data sources. (continued)**

| Data | Potential Planning/Decisionmaking Use |
|---|---|
| Air Quality Sensors | Cities are working with the private sector to deploy air quality sensors to allow management of traffic to minimize congestion, increase safety, and improve air quality. |
| Roadway Weather Predictions | While basic National Weather Service prediction data has been available for many years, these predictions cover wide areas and tend to focus more on air temperature and precipitation averaged over a region. Several companies now offer ground-based (i.e., at the street level) 48-hour weather predictions that are updated every hour. This information can be used to optimize winter weather response operations; improve snow event readiness; reduce staffing, fuel, and chemical costs; pinpoint treatment applications; and generally keep the roads safer and less congested. |
| Computer Aided Dispatch Data | Computer-aided dispatch data is generated by public safety agencies as part of their call intake and dispatch operations. This data can be used in realtime by TMCs to improve incident awareness or to inform them of incidents they would not have been apprised of otherwise. This allows TMCs to respond to incidents more efficiently and effectively and reduces clearance time and the potential for secondary incidents. |

1 Crowdsourced incident and congestion data is any data willingly and intentionally generated and reported by the public who receives something in return for their contributions. For the purposes of this document, it refers to travelers reporting incidents and congestion in realtime as they come across it using smartphone apps or similar technology.

Examples of companies that supply the data in table 1 are included in table 2 along with the sources of their data.

**Table 2. Where the private sector obtains its data.**

| Data Provider | Sources Of Data |
|---|---|
| Speed Data | Global Positioning System (GPS)-equipped vehicles and cell phones. Fleets of trucking companies with telematics. Cell phone navigation users. Embedded vehicle navigation systems. Apps installed on cell phones. Privately owned sensors installed on the agency right-of-way. |
| Crowdsourced Congestion and Incident Data | Waze application and Google/Android navigation app users. Public data inputs. |
| Fleet Management Data | Telematics equipment installed on vehicles rolling off the assembly line and aftermarket installations of equipment. |
| Origin-Destination Data | Location-based services on cell-phone applications. Credit card point-of-sale machines. Breadcrumb trails from navigation apps. Telematics providers. |
| Crowdsourced Map Data | Crowdsourced from a community of public users. |
| Ride-Sharing Location Data | Their own applications installed on driver and rider cell phones. |
| Mapping Data | Vehicles equipped with 360-degree cameras, GPS light detection and ranging (LiDAR), and other sensors and data collection equipment. |

**Table 2. Where the private sector obtains its data. (continued)**

| Data Provider | Sources Of Data |
|---|---|
| Parking Data | Third-party parking management systems, vehicle fleets with cameras/sensors. |
| Incident Data | Curated from their own traffic management systems, media data feeds, agency data feeds, computer-aided dispatch (CAD) feeds, crowdsourced, etc. |
| Road Weather Data | Specialized and hyperlocal weather conditions and predictions are produced from a mix of satellites, ground-based radar, ground sensors, and more. Specialized algorithms and data processing then produce more accurate ground-based weather predictions. Operators use this to understand the weather as expected and then experienced by the driver. |
| Law Enforcement or Emergency Services CAD | Third parties (like Motorola, Hexagon, TriTech, etc.) typically provide software CAD solutions to the law enforcement community, and the data within the CAD is entered directly by dispatchers using a keyboard, mouse, and/or touch screen. |

## DIFFERENCES AND SIMILARITIES IN COLLECTION, PROCESSING, AND AGGREGATING DATA

The public sector has historically relied on deploying physical infrastructure on public rights of way to collect data to support transportation operations. Traditional intelligent transportation systems (ITS) equipment such as sensors, vehicle counters, and cameras allowed agencies to get a better understanding of system conditions. Agencies were responsible for the procurement and installation of these sensors along with ongoing maintenance. This was both a blessing and a curse. Agencies were accustomed to procuring physical things—things that could be owned, tagged, inventoried, and held. Once installed, the agency was fully responsible for care and maintenance, including routine calibration, physical repair, power, communications, etc. If maintained and calibrated, these sensors and devices could provide relatively high-quality data, but only in areas where they were deployed.

The private sector, with rare exception,[1] has been limited in its ability to deploy its own equipment on the public right-of-way. Seeing opportunities in the Internet of Things and wireless technologies, the private sector instead focused on these emerging technologies and the opportunities they presented to potentially add value to agency operations.

As wireless and mobile technology continued to develop and smartphone penetration skyrocketed, the private sector was able to develop new ways of obtaining data of a quality similar to that of public agencies, but on a larger scale, covering most of the roads in the Nation, and doing so without deploying much (if any) infrastructure.

Table 3 summarizes the operations-related uses for each data type and how public sector agencies could obtain the data or an approximation of that data without purchasing it directly from a third party.

1 Eric Lipton, "U.S. System for Tracking Traffic Flow Is Faulted," *New York Times*, December 13, 2009. Available at https://www.nytimes.com/2009/12/14/us/14traffic.html, last accessed March 17, 2019.

**Table 3. Public sector alternatives to obtaining third-party data.**

| Data | Data Uses and Public Sector Data Alternatives |
| --- | --- |
| **Credit Card Transactions** | **Uses:** For origin-destination (O-D) studies, understanding if businesses are opened/closed, and for determining which businesses have electricity.<br><br>**Public Sector Alternative:** Agencies would need to conduct surveys or review census data for O-D studies. Power failure data can be collected from utilities if they are willing to share that information. |
| **Third-Party Connected Vehicle Data** | **Uses:** For event detection, warning of potential safety hazards, weather conditions, and more.<br><br>**Public Sector Alternative:** While basic safety message (BSM) data can be obtained from equipped vehicles, agencies still need to install dedicated roadside equipment to collect and send the BSM data back to an operations center. This means that agencies are still responsible for maintaining physical devices and that data can only be collected in specific locations where roadside units are installed. Additionally, the BSM data is fairly limited in scope compared to what third parties can pull directly from the vehicle at any location on the roadway and transmit back to a central location via a cellular connection. |
| **Realtime Trajectory Data** | **Uses:** Realtime route and detour analysis, evacuation monitoring, signal performance measures, O-D analysis, after-action reviews, etc.<br><br>**Public Sector Alternative:** Agencies would need to conduct surveys or review census data for O-D studies. Floating car runs have been used by many agencies to replicate trajectory data; however, floating car runs are not representative of traveler route choices, and they are almost never done in realtime. Aerial photography studies have been purchased by agencies and used for some of these applications; however, aerial studies only cover a limited geography, and the resulting data is not provided to the agency in realtime. |
| **Realtime Turning Movement Data** | **Uses:** Detour and alternate route adherence, traffic signal performance measures.<br><br>**Public Sector Alternative:** The public sector can only collect realtime turning movement data through the deployment of large quantities of sensors (inductive loops, side-fired microwave/radar), license plate recognition, toll-tag readers, high-definition signal controller data, and/or Bluetooth/Wi-Fi re-identification sensors. |
| **Wi-Fi/Bluetooth Re-identification** | **Uses:** Travel times, traveler information, route selection, signal performance measures.<br><br>**Public Sector Alternatives:** Unlike many of the other data sources in this list, Bluetooth re-identification sensors are equipment that agencies can still procure and install by themselves at signals and other locations along the roadside. This data is similar to probe-based speed data, toll-tag speed data, and license plate recognition data. While most companies sell the sensors directly to the agency for installation, some companies also rent equipment or provide Bluetooth data as a service. |

**Table 3. Public sector alternatives to obtaining third-party data. (continued)**

| Data | Data Uses and Public Sector Data Alternatives |
|---|---|
| High-Resolution Signal Data | **Uses:** Automated traffic signal performance measures (ATSPM), signal operations, signal retiming, volume counts, turning movement analysis.<br><br>**Public Sector Alternatives:** With some exceptions, the private sector typically owns and aggregates the data that is collected directly from the traffic controller; however, some companies have leveraged the open source ATSPM modules in proprietary systems that are sold back to agencies with enhanced user interfaces and data management services. Other companies sell aftermarket products that can be installed in signal cabinets that collect and aggregate this data in a cloud for resale back to agencies. |
| Roadside BSM Collection | **Uses:** Intersection safety, work zone safety, event detection, warning of other potential safety hazards, weather conditions, and more.<br><br>**Public Sector Alternatives:** While BSM data can be obtained from equipped vehicles, agencies still need to install dedicated roadside equipment to collect and send the BSM data back to an operations center. This means that agencies are still responsible for maintaining physical devices and that data can only be collected in specific locations where roadside units are installed. |
| Crowdsourced[1] Incident and Congestion Data | **Uses:** Incident/event detection, traveler information.<br><br>**Public Sector Alternatives:** Agencies have historically gathered event/incident information from the monitoring of closed-circuit television (CCTV) feeds, listening to radio systems, interfacing with law enforcement computer-aided dispatch (CAD), and/or service patrols. While many of these existing detection sources are considered superior to crowdsourced data for major events/crashes, smaller events (like disabled vehicles, debris, etc.) are typically better sourced via the crowd and cover a wider geographic area than what a state transportation agency might normally cover. |
| Crowdsourced Mapping Data | **Uses:** Base-mapping for traveler information, asset management/asset identification, routing.<br><br>**Public Sector Alternatives:** Almost every agency is responsible for mapping roads already. However, many agencies only produce centerline map files. For these agencies, some crowdsourced mapping products, though not authoritative, may actually have better map attributes, assets, and even be more up to date. |
| Probe-Based Speed Data | **Uses:** Traveler information, congestion analytics, performance reporting, problem identification, before-and-after studies, etc.<br><br>**Public Sector Alternatives:** Some agencies do collect probe-based speed data through the deployment of toll-tag readers, license plate re-identification, Bluetooth re-identification, etc. Other agencies have attempted to mimic third party probe data by outfitting maintenance vehicles (or transit vehicles) with Global Positioning System equipment. However, no agency is currently able to collect speed/travel time data at the same scale and with the same number of probes as the private sector can. |

**Table 3. Public sector alternatives to obtaining third-party data. (continued)**

| Data | Data Uses and Public Sector Data Alternatives |
|---|---|
| High-Resolution Map Data (LiDAR or Similar) and Other Asset Management Systems | **Uses:** Asset management, precision navigation, identification of maintenance issues.<br><br>**Public Sector Alternatives:** A few agencies have purchased their own light detection and ranging (LiDAR) equipment and are collecting and storing point-cloud data on their own. The private sector, however, is providing services that collect and manage the point cloud data on behalf of the agency, and distill the data down to more manageable asset attribute information, like the location of signs, reflectivity, pavement conditions, etc. |
| Roadway Weather Predictions | **Uses:** Understanding expected road temperature, moisture, snow cover, wind speed, visibility, etc.<br><br>**Public Sector Alternatives:** While many agencies have deployed road weather information system (RWIS) stations, these stations only provide current weather conditions and only directly at the location of the RWIS station. Only a couple of agencies have the luxury of an in-house meteorologist on staff that can help predict weather for the agency. Even these dedicated meteorologists can struggle with producing surface weather predictions for all road segments. |
| CAD Data | **Uses:** Faster incident detection, incident detection on roads that are not normally covered by CCTV or the agency, understanding who has been dispatched to an event, assessing the severity of the event.<br><br>**Public Sector Alternatives:** CAD is typically generated by public agencies; however, it is usually managed and owned by public safety instead of transportation. The closest equivalent to CAD that is already owned by transportation agencies is event and notification data from advanced traffic management systems. |

1 Crowdsourced incident and congestion data is any data willingly and intentionally generated and reported by the public who get something in return for their contributions. For the purposes of this document, it refers to travelers reporting incidents and congestion in realtime as they come across it using smartphone apps or similar technology.

## Processing and Analytics Differences

Public agencies have kept pace with the latest relational database storage and analytics capabilities necessary for the effective usage of existing agency data—like volume and speed sensors, advanced traffic management system (ATMS) event records, etc. For example, a large State agency might deploy 15,000+ traffic flow sensors on their highways that record volume and speed data every 20 seconds. Most agencies can easily handle this level of data storage. Even though this seems like a significant amount of data, it comes nowhere close to the quantity of data produced by the tracking of Global Positioning System (GPS)-equipped cellular phones, light detection and ranging (LiDAR), or other connected vehicle (CV) applications.

As the public sector continued to invest in relational database technologies to store its sensor and ATMS data, the private sector had to fast track its own, newer storage and analytics capabilities to keep up with development and demands of the wireless and mobile markets, which were creating massive streaming datasets that would not fit into existing relational database management systems. This meant the private sector needed to develop big data analysis and storage capabilities well beyond those of most public sector agencies.

Now that public agencies are tapping into these big-data streams, some are finding that their own relational databases are ill-equipped to process, host, and leverage the data to its fullest potential. Agencies are faced with several choices: 1) adopt or develop the same big data capabilities as the private sector and invest in hardware storage or cloud services, 2) outsource the hosting of these datasets to the private sector or other third parties, or 3) not attempt to ingest raw data, but instead leverage application programming interfaces (APIs) or data summary and insights services from the data providers without the need to download and manage large datasets.

The remainder of this chapter will describe these datasets in more detail—giving an overview of potential applications, data providers, pros and cons, etc. Later chapters will discuss business models, the value of agency data, and contracting guidance.

## CROWDSOURCED INCIDENT AND CONGESTION DATA

### Description

Crowdsourced incident and congestion data is any such data willingly and intentionally generated and reported by members of the public who get something in return for their contributions. Crowdsourced transportation data does not include probe-based data. Instead, crowdsourced data refers to travelers reporting incidents and congestion in realtime as they come across it using smartphone apps or similar technology.

### Applications of Crowdsourced Incident and Congestion Data

Many applications can use crowdsourced incident and congestion data in the same way as agency incident and congestion data. For example, crowdsourced data can help manage traffic in realtime by providing TMCs with awareness of new incidents and congested spots. TMCs can dispatch field units to incident scenes more quickly, or implement congestion mitigation strategies in response to congestion reports. In addition, crowdsourced data applies to realtime operations and can also be archived and used in planning and performance management efforts. Table 4 shows a few agencies that are making use of crowdsourced data and applications today.

**Table 4. Sample list of agencies using crowdsourced data.**

| Data | Application | Peer Contact |
|---|---|---|
| Waze Event Data | Early detection of hazards and events. | Massachusetts and Florida Departments of Transportation (DOTs). |
| Waze Navigation Guidance | Agency "pushes" road closures, detour routes, or preferred routes to the crowd to control and influence traffic. | Port Authority of New York and New Jersey. |
| SeeClickFix Data or Other Citizen Reporting Apps | Early detection of maintenance issues. | Utah and New Hampshire DOTs. |
| Twitter Messages | Pre-event planning. Pre-event warnings. Social sentiment analysis. | Metropolitan Area Transportation Operations Coordination Program (National Capital Region), Iowa and D.C. DOTs. |
| 911 Phone Calls | Event detection, responder deployment, and quick clearance support. | California, Virginia, and Wisconsin DOTs. |

## General Attributes

- **Latency:** Generally, crowdsourced incident and congestion data has a low latency. Because there is little to no verification, data can be available the moment travelers enter it. Waze data supports the ability to identify disabled vehicles, debris, smaller incidents, and incidents off State-monitored roadways significantly faster than public agencies or first responders.
- **Details:** Crowdsourced data may be less detailed than data agencies generate because travelers are asked to enter only basic information associated with an incident or congestion in an effort to balance the benefit of useful user-generated information with the safety of the app users (i.e., to avoid distracted driving).
- **Quality and Coverage:** The number of participants in the crowdsourcing effort relates directly to penetration, quality, and coverage of crowdsourced data. Due to this association, crowdsourced data generally seems to be more prevalent and accurate in dense urban areas, highly traveled corridors, and regions with a high level of technology acceptance.

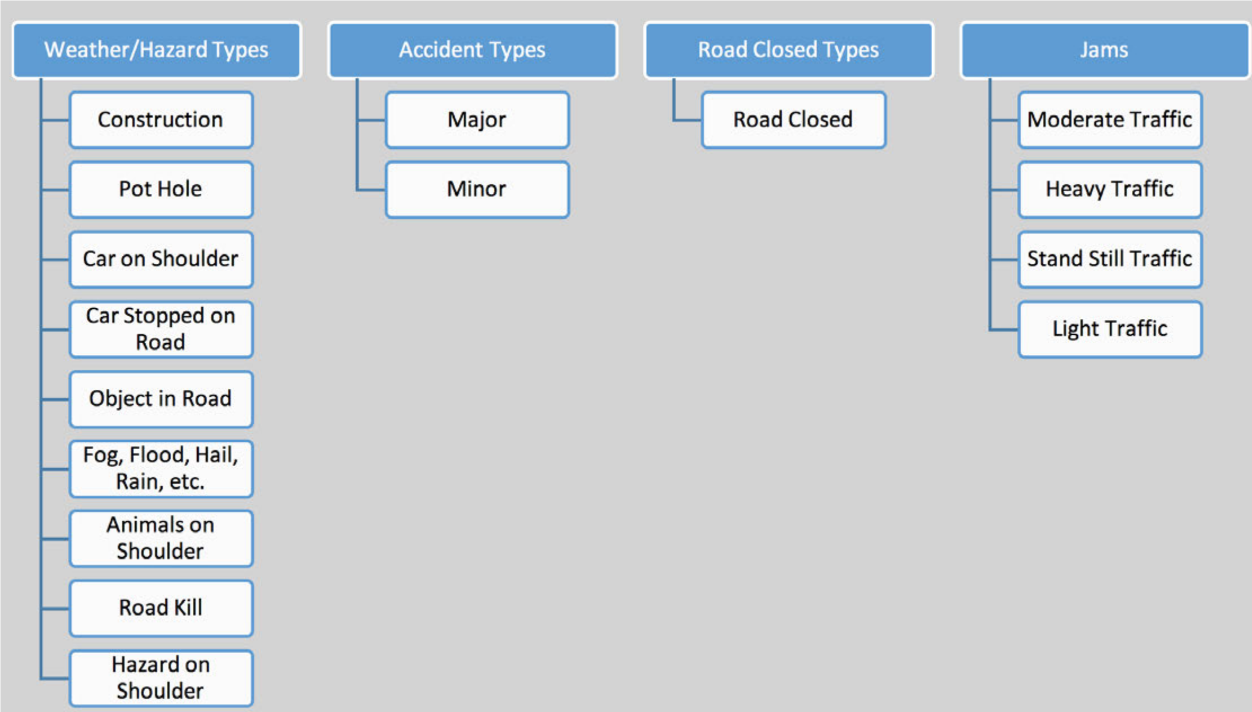



**Figure 2. Diagram. Examples of Waze event types and subtypes.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

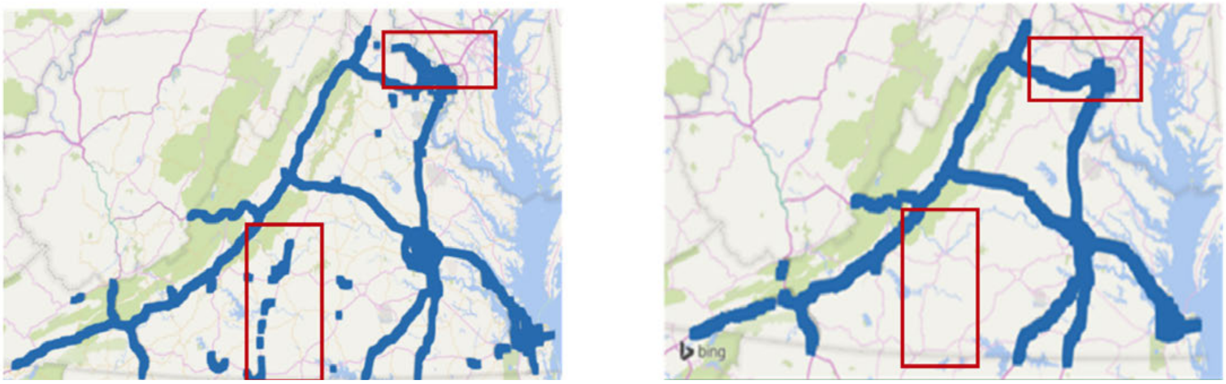
## Data Availability

There are multiple platforms that TMCs can leverage to glean data from the crowd. Some platforms are focused on transportation while others are primarily social networks. Some agencies mine data from social networks by evaluating contents of posted messages and finding key indicators of traffic-related information. For example, some agencies used machine learning algorithms to recognize typical message constructs that may identify an incident. These algorithms can assign a level of confidence to the results by comparing the location information, time, and other attributes across other social networks and agency systems. Other agencies simply monitor social media messages from trusted partners or users (like the media, public safety departments, etc.).

Many States and municipalities are starting to encourage their citizens to use apps that allow citizens (the crowd) to take pictures of maintenance issues (like potholes, damaged lights, signal issues, curb/sidewalk issues, etc.) and send them to the agency (along with geo-located photos, attributes, etc.). A database can manage the citizen reports, which is turning into a "ticketing system" for many smaller agencies.

## Pros and Cons

As shown in figure 3, crowdsourced data tends to provide broader coverage than individual agencies may be able to afford with limited resources and jurisdictional responsibilities. Crowdsourced data may be generated anywhere people travel, encounter issues, and have internet access. This means users can report on interstates, State and county roads, local roads, neighborhoods, etc.





**a) Waze events.** **a) Virginia Department of Transportation events.**

**Figure 3. Screenshots. Red boxes highlight the coverage differences in the Waze data versus the Virginia Department of Transportation data.**

**Source: I-95 Corridor Coalition. n.d. *Closing the Realtime Data Gaps Using Crowd-Sourced Waze Event Data.* Unpublished technical report.**

Crowdsourced data can be timelier than agency-generated data due to the higher probability of travelers coming across an incident or encountering congestion than an agency detecting it via its sensors, cameras, and field patrols. A recent I-95 Corridor Coalition (ICC) study compared the amount of time it takes for events to appear in the Waze data feed compared to the amount of time it takes the agency to identify the incident and put it into their ATMS or 511 platform.[2] Agencies usually detect major collisions on interstates sooner than Waze. However, Waze identifies smaller events, debris, and disabled vehicles before the ATMS—sometimes by 15 minutes or more. These smaller events do have the potential to become larger or cause secondary incidents if not dealt with in a timely manner. Table 5 shows the results of the ICC study, indicating the average length of time Waze reported an event before the DOT ATMS reported it when both reported the same event (includes events in California, Florida, and Virginia).

**Table 5. Comparison of Waze versus department of transportation event reporting.**

| Type of Event | Avg. Time a Waze Event was Reported Before DOT Reporting | Percentage of All Waze Events that were Included in the DOT's ATMS Logs |
|---|---|---|
| Freeways/Ramps Crashes | 3 Minutes | 40 percent |
| Primary/Secondary Crashes | 3 Minutes | 12 percent |
| Freeways/Ramps Disabled Vehicles | 14 Minutes | 37 percent |
| Primary/Secondary Disabled Vehicles | 16 Minutes | 4 percent |

Source: I-95 Corridor Coalition. n.d. *Closing the Realtime Data Gaps Using Crowd-Sourced Waze Event Data.* Unpublished technical report.

ATMS = advanced transportation management system. DOT = department of transportation.

2 I-95 Corridor Coalition, *Closing the Realtime Data Gaps Using Crowdsourced Waze Event Data*. Unpublished technical report.

As a general observation, official entities do not verify crowdsourced data because any user can report anything. Crowdsourced data providers usually have built-in mechanisms for a "self-moderating" community, where other travelers can confirm or reject reports, therefore providing some level of confidence that the reported incident or congestion is real. As previously noted, Waze data feeds provide confidence and reliability scores. Confidence scores range from 0 to 10 and measure how other drivers react to the report (a higher score indicates positive feedback from other Waze users). Reliability scores measure the experience level of the person reporting the event. The more a user contributes and the more they receive positive feedback from other Waze users, the higher their score. Many agencies use a combination of the confidence and reliability scores to determine whether it is worth acting on a particular event. Figure 4 shows how these scores are displayed and filtered in one traffic management system.

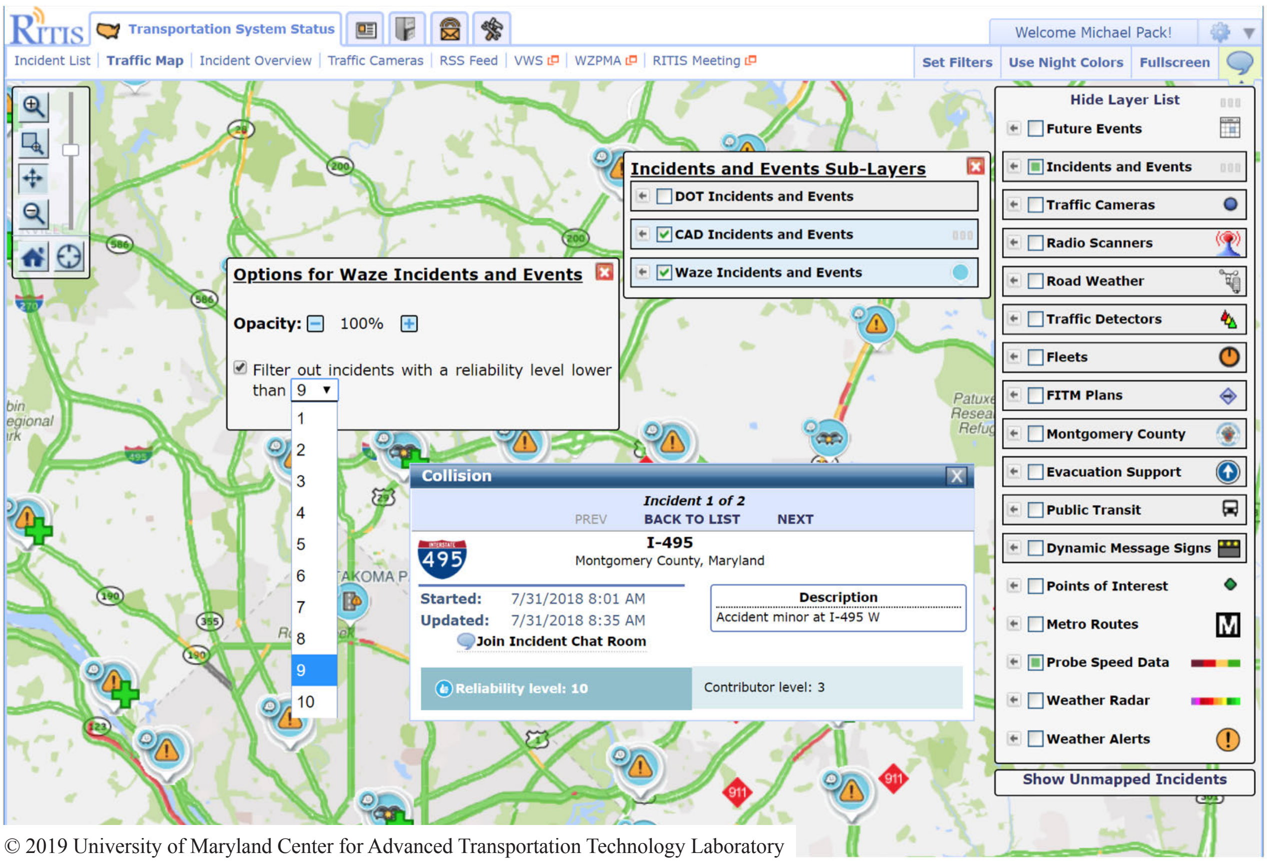





**Figure 4. Screenshot. Many agencies filter out Waze data that have a lower reliability level.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory Regional Integrated Transportation Information System Website.**

## Use Cases for Crowdsourced Incident and Congestion Data

**Use Case: Regional Integrated Transportation Information System (RITIS) Waze data integration.** RITIS ingests realtime Waze data for the entire country. State DOTs and other public agencies use RITIS to fuse data from multiple providers into a single, common operational picture, or as a data aggregator of third-party data feeds like Waze. Formatting Waze and other third-party data into standardized data feeds enables incorporation of data back into a realtime ATMS platform. Figure 5 shows the RITIS map with Waze data overlaid. Users are able to filter the display of Waze events based on the reliability score. Several agencies then access the RITIS data application programming interface and re-ingest the Waze data (along with other third-party incident data feeds) in a single stream.

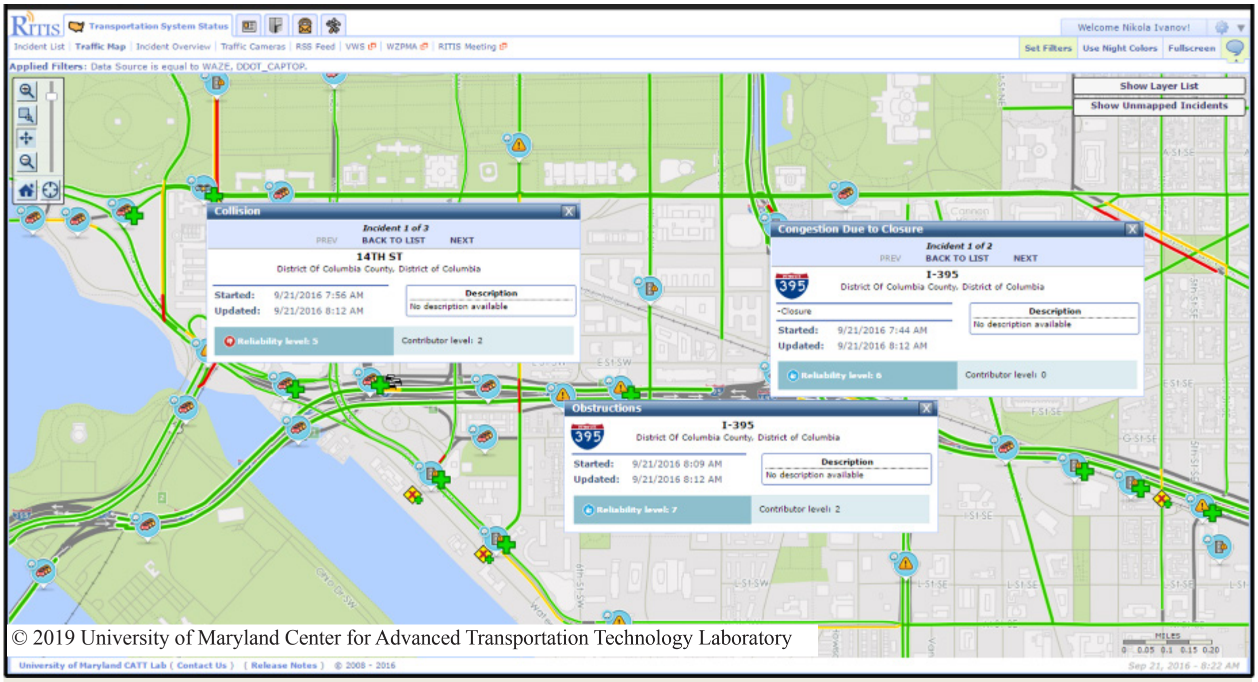



**Figure 5. Screenshot. Realtime Waze data integrated into the Regional Integrated Transportation Information System platform.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

The RITIS platform also archives Waze data at the national level. Operators can use the RITIS archived Waze data analytics module to understand where to stage field responders or plan service routes by analyzing the location of incident hot-spots by time-of-day and day-of-week.

Google-owned Waze provides a smartphone app that allows drivers to report incidents, congestion, and other traffic-related events (figure 6). Over the last several years, Waze has partnered with a number of public agencies (cities, States, and regions) across the world to provide crowdsourced data in exchange for agency-generated planned event information. Waze calls this its Connected Citizens Program (CCP).

Waze provides three primary types of notification data: alerts, jams, and irregularities. Waze provides CCP partners with extensible markup language (XML)/JavaScript Object Notation (JSON) files. Data providers and users must overcome challenges with both data size and quality to transform raw Waze data to reliable information that can help DOTs make data-driven decisions.

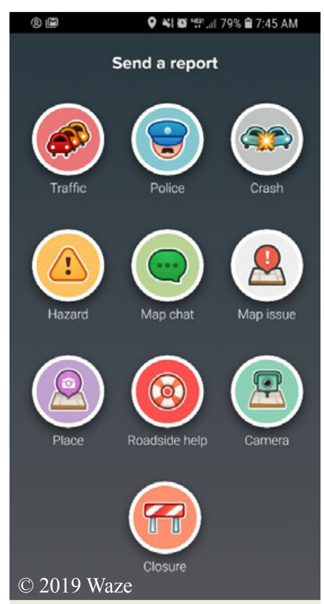



**Figure 6. Screenshot. Waze smartphone app.**

**Source: Waze.**

**Data Size:** Waze generates a substantial number of new notifications on a daily basis. The Center for Advanced Transportation Technology (CATT) Laboratory at the University of Maryland has been ingesting data from Waze for several years in its RITIS realtime data fusion platform. On an average day, CATT Lab's RITIS platform receives approximately 27,000 Waze events across 12 States (approximately 54,000 events if including jams). This is significantly more than the number of ingested, stored, and visualized events across all States,[3] as shown in table 6. For example, Waze could result in a 400-fold increase in the number of events in Massachusetts. As such, ingesting only the desired Waze events by applying filtering will be essential to increasing the usability of the data. As shown in figure 7, the average number of Waze events is highly correlated to the vehicle miles traveled (VMT) in each State.

**Table 6. Waze versus department of transportation events per day (excluding jams).**

| State | Avg. Waze Events per Day | Avg. DOT Events per Day |
|---|---|---|
| CA | 28,3889 | 3,184 |
| DC | 777 | 16 |
| FL | 17,210 | 1,895 |
| IA | 810 | 114 |
| MA | 5,613 | 14 |
| PA | 9,171 | 70 |
| VA | 9,168 | 681 |

Source: University of Maryland Center for Advanced Transportation Technology Laboratory.

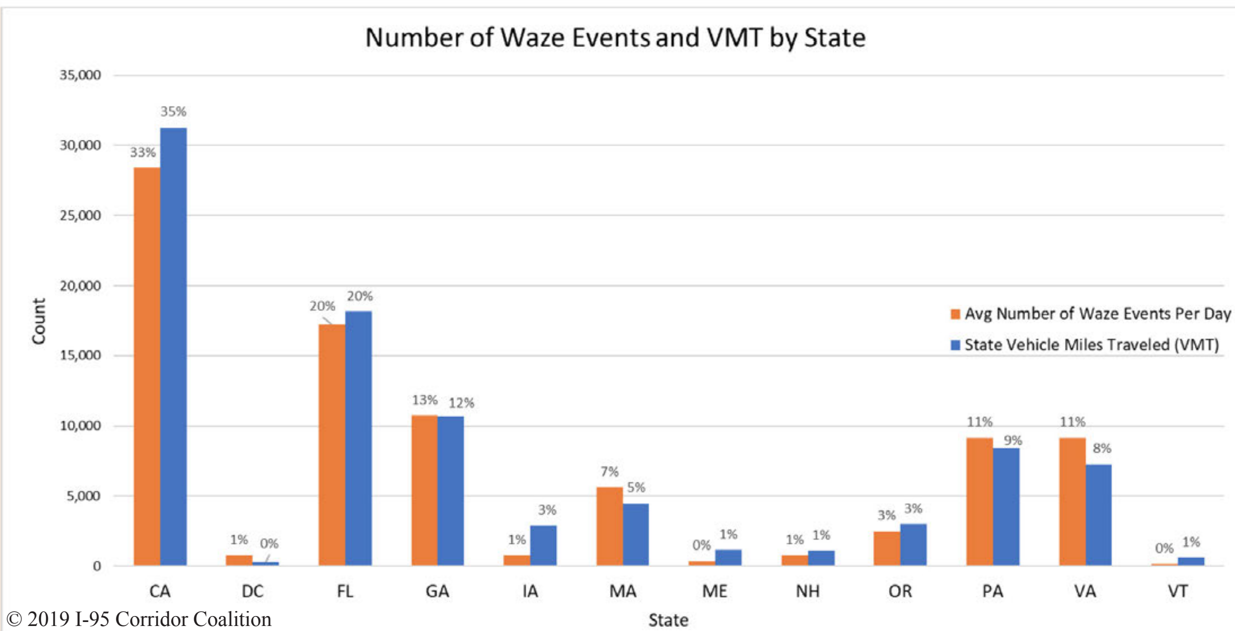



**Figure 7. Chart. Waze events and vehicle miles traveled by State (excluding jams).**

**Source: I-95 Corridor Coalition. n.d. *Closing the Realtime Data Gaps Using Crowd-Sourced Waze Event Data*. Unpublished technical report.**

**Data Quality:** Waze's data quality can be susceptible to redundancy, completeness, and unreliable information such as the generation of false-positive events.

3 Table 6 displays only select States for which the CATT Lab had access to both DOT ATMS and Waze data.

- **Redundancy:** Waze users on occasion report the same event using different event types and report the event at different locations and times. Although Waze has built-in mechanisms to reduce event duplication (e.g., confidence and reliability scores), Waze data still contains duplicative events. The variation in locations and times is likely a result of a delay in entering an event into the Waze application, while the different event types may be a result of the actual event being unclear or a user selecting an incorrect event type. It is essential to cluster duplicative Waze events and consolidate interchangeable event types to increase the quality. Due to the data sharing between DOT TMCs and Waze, replication of DOT events within the system can result in further duplication within the Waze data. These events generally have identical spatial and/or temporal attributes, such as timestamp and location. Although Waze data does not cause quality issues, either exclusion or consolidation of these events should take place after data ingestion.
- **Completeness:** While Waze data frequently covers more roads and geographies than ATMS, there is no guarantee that the public/crowd will report every incident. Therefore, some incidents may go unreported, making the data incomplete. While not as significant an issue as missing data, some incidents can remain open or active for too long in the Waze data feeds after they have officially been cleared. This is true for incidents, disabled vehicles, debris, construction, etc.
- **Reliability:** Waze data can be difficult to validate. False-positive or insignificant events may be either incorrectly reported or unverifiable when a TMC attempts to validate that the event occurred. Common examples include a vehicle pulled over on the shoulder for less than 1 minute, a fender bender with no damage to the vehicles, or a small animal identified as roadkill. Events with short durations and/or lower confidence and reliability scores are more likely to be false-positive events. Authenticated matching procedures for comparing Waze events against information received by TMCs will help organizations automate a method for validating Waze data and better understand elements associated with reliable event reports.

**Example Waze Incident Data:** Waze supports the types and subtypes of user-generated alerts shown in Table 7:

**Table 7. User-generated alert types and subtypes supported by Waze.**

| Alert Type | Alert Subtype |
|---|---|
| **ACCIDENT** | ▪ ACCIDENT_MINOR<br>▪ ACCIDENT_MAJOR<br>▪ NO_SUBTYPE |
| **JAM** | ▪ JAM_MODERATE_TRAFFIC<br>▪ JAM_HEAVY_TRAFFIC<br>▪ JAM_STAND_STILL_TRAFFIC<br>▪ JAM_LIGHT_TRAFFIC<br>▪ NO_SUBTYPE |

**Table 7. User-generated alert types and subtypes supported by Waze. (continued)**

| Alert Type | Alert Subtype |
|---|---|
| WEATHERHAZARD/ HAZARD | ■ HAZARD_ON_ROAD<br>■ HAZARD_ON_SHOULDER<br>■ HAZARD_WEATHER<br>■ HAZARD_ON_ROAD_OBJECT<br>■ HAZARD_ON_ROAD_POT_HOLE<br>■ HAZARD_ON_ROAD_ROAD_KILL<br>■ HAZARD_ON_SHOULDER_CAR_STOPPED<br>■ HAZARD_ON_SHOULDER_ANIMALS<br>■ HAZARD_ON_SHOULDER_MISSING_SIGN<br>■ HAZARD_WEATHER_FOG<br>■ HAZARD_WEATHER_HAIL<br>■ HAZARD_WEATHER_HEAVY_RAIN<br>■ HAZARD_WEATHER_HEAVY_SNOW<br>■ HAZARD_WEATHER_FLOOD<br>■ HAZARD_WEATHER_MONSOON<br>■ HAZARD_WEATHER_TORNADO<br>■ HAZARD_WEATHER_HEAT_WAVE<br>■ HAZARD_WEATHER_HURRICANE<br>■ HAZARD_WEATHER_FREEZING_RAIN<br>■ HAZARD_ON_ROAD_LANE_CLOSED<br>■ HAZARD_ON_ROAD_OIL<br>■ HAZARD_ON_ROAD_ICE<br>■ HAZARD_ON_ROAD_CONSTRUCTION<br>■ HAZARD_ON_ROAD_CAR_STOPPED<br>■ NO_SUBTYPE |
| MISC | ■ NO_SUBTYPE |
| CONSTRUCTION | ■ NO_SUBTYPE |
| ROAD_CLOSED | ■ ROAD_CLOSED_HAZARD<br>■ ROAD_CLOSED_CONSTRUCTION<br>■ ROAD_CLOSED_EVENT<br>■ NO_SUBTYPE |

Source: Waze, Google Developers Web page. Available at: https://developers.google.com/waze/data-feed/overview, last accessed March 17, 2019.

Waze data is in either XML or JSON format.

An example of a weather hazard notification is below in the XML format with a description of each element as shown in table 8.


```
<item>

      <pubDate>Thu Nov 26 14:02:29 +0000 2015</pubDate>

      <georss:point>45.02395420471421 7.670893079148089</georss:point>

      <linqmap:uuid>9fd1ee98-7b56-37e9-a2d4-72e9478dd838</linqmap:uuid>

      <linqmap:magvar>6</linqmap:magvar>

      <linqmap:type>WEATHERHAZARD</linqmap:type>

      <linqmap:subtype>HAZARD_ON_ROAD_CONSTRUCTION</linqmap:subtype>

      <linqmap:reportDescription>

      scambio di carreggiata causa lavori dalle 00:00 del 16 novembre
      2015 alle 23:59 del 21

      gennaio 2016

      </linqmap:reportDescription>

      <linqmap:city>Torino</linqmap:city>

      <linqmap:country>IT</linqmap:country>

      <linqmap:roadType>4</linqmap:roadType>

      <linqmap:reportRating>0</linqmap:reportRating>

      <linqmap:reliability>10</linqmap:reliability>

</item>

<item>

      <pubDate>Thu Nov 26 14:02:26 +0000 2015</pubDate>

      <georss:point>45.02395420471421 7.670893079148089</georss:point>

      <linqmap:uuid>ed06a695-53ee-347c-a6eb-133bf8746880</linqmap:uuid>

      <linqmap:magvar>6</linqmap:magvar>

      <linqmap:type>WEATHERHAZARD</linqmap:type>

      <linqmap:subtype>HAZARD_ON_ROAD_CONSTRUCTION</linqmap:subtype>

      <linqmap:reportDescription>

      chiusura notturna causa lavori di manutenzione dalle 23:00 alle
```

```
    05:30, solo nei giorni

    feriali dalle 23:00 del 9 novembre 2015 alle 05:30 del 5 dicembre
    2015

    &lt;/linqmap:reportDescription&gt;

    &lt;linqmap:city&gt;Torino&lt;/linqmap:city&gt;

    &lt;linqmap:country&gt;IT&lt;/linqmap:country&gt;

    &lt;linqmap:roadType&gt;4&lt;/linqmap:roadType&gt;

    &lt;linqmap:reportRating&gt;0&lt;/linqmap:reportRating&gt;

    &lt;linqmap:reliability&gt;7&lt;/linqmap:reliability&gt;

&lt;/item&gt;
```

**Table 8. Weather hazard event notification elements.**

| Element | Value | Description |
|---|---|---|
| pubDate | Time | Publication date. |
| georss:point | Coordinates | Location per report (Lat long). |
| linqmap:uuid | String | Unique system ID. |
| linqmap:magvar | Integer (0-359) | Event direction (Driver heading at report time. 0 degrees at North, according to the driver's device). |
| linqmap:type | See alert type table | Event type. |
| linqmap:subtype | See alert subtypes table | Event subtype depends on parameter. |
| linqmap:reportDescription | String | Report description (supplied when available). |
| linqmap:street | String | Street name (as is written in database, no canonical form, may be null). |
| linqmap:city | String | City and state name [City, State] in case both are available, [State] if not associated with a city (supplied when available). |
| linqmap:country | String | See two letters codes in http://en.wikipedia.org/wiki/ISO_31661. |
| linqmap:roadType | Integer | Road type (see road types table in the appendix). |
| linqmap:reportRating | Integer | User rank between 1-6 (6 = high ranked user). |
| linqmap:jamUuid | String | If the alert is connected to a jam-jam ID. |
| linqmap:Reliability (new) | 0-10 | How reliable is the report, 10 being most reliable. Based on reporter level and user responses. |

Here is the same example in JSON format:


```
{"country":"IT", "roadType":1, "magvar":258, "subtype":"",
"reportRating":0, "reliability":6, "reportDescription":"blocco del
traffico per alcuni veicoli nella ZTL (Zona Traffico Limitato) Non possono
circolare Veicoli per il trasporto persone Dal lunedì al venerdì, dalle
ore 8 alle ore 19 veicoli benzina Euro",
"location":{"x":7.6800935614336545,"y":44.9991565694201},
"type":"WEATHERHAZARD", "uuid":"39d9dc07bd743b35ba6b833f5cbd1ce1",
"pubMillis":1448546704610},
{"country":"IT", "magvar":0, "subtype":"ROAD_CLOSED_
EVENT", "city":"Nichelino", "street":"Via Fenestrelle",
"reportRating":0, "reliability":9, "reportDescription":"lavori",
"location":{"x":7.627331910061528,"y":45.00419885851123},
"type":"ROAD_CLOSED", "uuid":"1064e72c0d3b332d95c61dcab524aa5c",
"pubMillis":1446918728242},
```


**Example Waze Speed/Congestion Data:** Waze generates traffic jam information by processing the following data sources:

- GPS location points sent from users' phones (users who drive while using the app) and calculations of the actual speed versus average speed (on specific timeslot) and free-flow speed (maximum speed measured on the road segment).
- User-generated reports shared by Waze users who encounter traffic jams. These appear as regular alerts and affect the way we identify and present traffic jams.

Table 9 describes traffic jam parameters, which are also in an XML format.

**Table 9. Traffic jam parameters.**

| Element | Value | Description |
|---|---|---|
| pubDate | Time | Publication date. |
| linqmap:type | String | TRAFFIC_JAM. |
| georss:line | List of longitude and latitude coordinates | Traffic jam line string (supplied when available). |
| linqmap:speed | Float | Current average speed on jammed segments in meter/second. |
| linqmap:length | Integer | Jam length in meters. |
| linqmap:delay | Integer | Delay of jam compared to free flow speed, in seconds (in case of block, 1). |
| linqmap:street | String | Street name (as is written in database, no canonical form (supplied when available). |

| Element | Value | Description |
|---|---|---|
| linqmap:city | String | City and state name [City, State] in case both are available, [State] if not associated with a city (supplied when available). |
| linqmap:country | String | Available on EU (world) server (see two letters codes in http://en.wikipedia.org/wiki/ISO_31661). |
| linqmap:roadType | Integer | Road type (see road types table in the appendix). |
| linqmap:startNode | String | Nearest Junction/street/city to jam start (supplied when available). |
| linqmap:endNode | String | Nearest Junction/street/city to jam end (supplied when available). |
| linqmap:level | 0-5 | Traffic congestion level (0 = free flow 5 = blocked). |
| linqmap:uuid | String | Unique jam identifier. |
| linqmap:turnLine | Coordinates | A set of coordinates of a turn only when the jam is in a turn (supplied when available). |
| linqmap:turnType | String | What kind of turn it is: left, right, exit R or L, continue straight, or NONE (no info) (supplied when available). |
| linqmap:blockingAlertUuid | String | If the jam is connected to a block (see alerts). |

An example of a traffic jam notification is below in the XML format.


```
<item>
      <pubDate>Sun Nov 29 12:57:44 +0000 2015</pubDate>
      <linqmap:uuid>52cf216f-799e-3b62-9b72-5cb6a15e9c67</linqmap:uuid>
      <linqmap:type>Medium</linqmap:type>
      <georss:line>
      40.680629 -74.004695 40.681749 -74.005537 40.682689 -74.005947
      40.683742
      -74.00628 40.684477 -74.006569 40.685214 -74.006994 40.686049
      -74.007391
      40.688904 -74.009512 40.690987 -74.011508 40.700833 -74.015145
      </georss:line>
      <linqmap:speed>15.3629673206283</linqmap:speed>
      <linqmap:length>2433.0</linqmap:length>
      <linqmap:delay>61</linqmap:delay>
      <linqmap:endNode>Hugh L. Carey Tunnel</linqmap:endNode>
      <linqmap:street>Hugh L. Carey Tunnel</linqmap:street>
      <linqmap:city>New York, NY</linqmap:city>
      <linqmap:country>US</linqmap:country>
      <linqmap:roadType>3</linqmap:roadType>
      <linqmap:level>1</linqmap:level>
      <linqmap:turnType>NONE</linqmap:turnType>
</item>
```


**Use Case: Metropolitan Area Transportation Operations Coordination (MATOC) TweetDeck.** Figure 8 is from the MATOC program in the National Capital Region. The MATOC operations center actively monitors Twitter feeds through a software application called TweetDeck (https://tweetdeck.twitter.com/). Through this application, the MATOC operations personnel are able to get information from other agencies, the media, and even the public who may prefer tweeting rather than using Waze. Monitoring these feeds requires additional diligence and practice (figure 9). Operators must configure their TweetDeck to report out on users, hashtags, and other information deemed relevant to the operator. It took several months of effort to refine their settings before the feeds were providing the level of information they desired. They must also tweak their settings every few months to ensure everything continues to run smoothly. Despite the initial effort involved, the operators like the system so much that they dedicated an entire media wall to the feeds so that it is always visible.

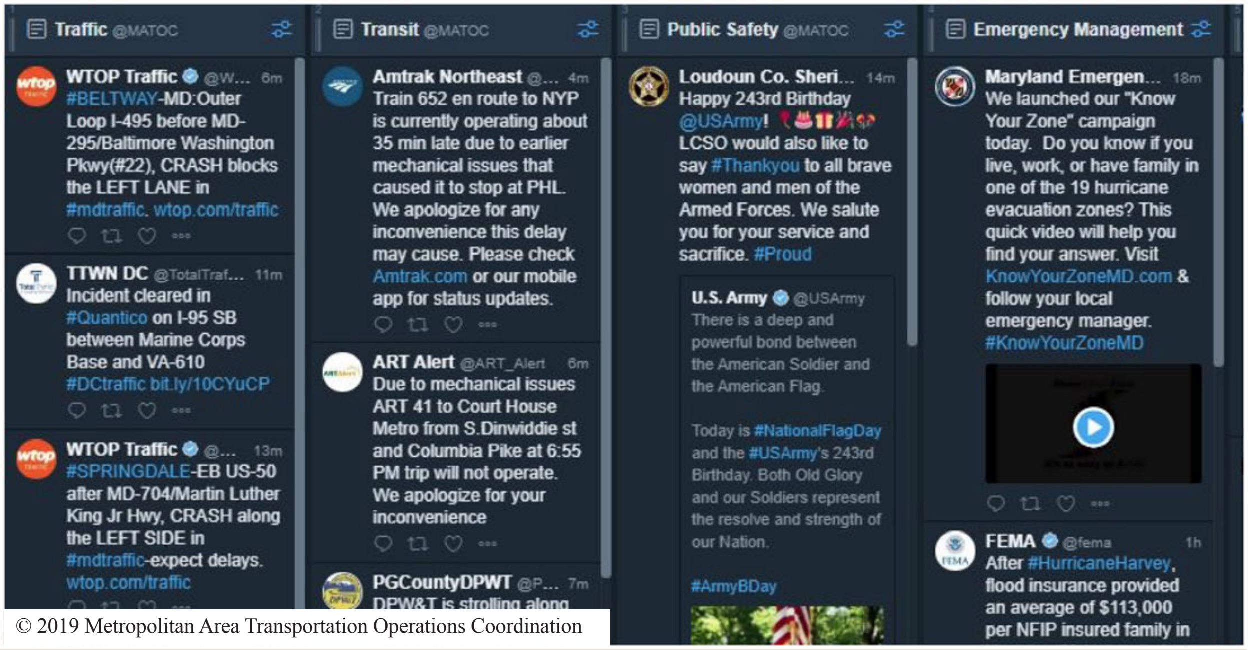




**Figure 8. Screenshot. A small portion of the Metropolitan Area Transportation Operations Coordination Twitter monitoring system.**

**Source: TweetDeck screenshot provided courtesy of the Metropolitan Area Transportation Operations Coordination Program.**

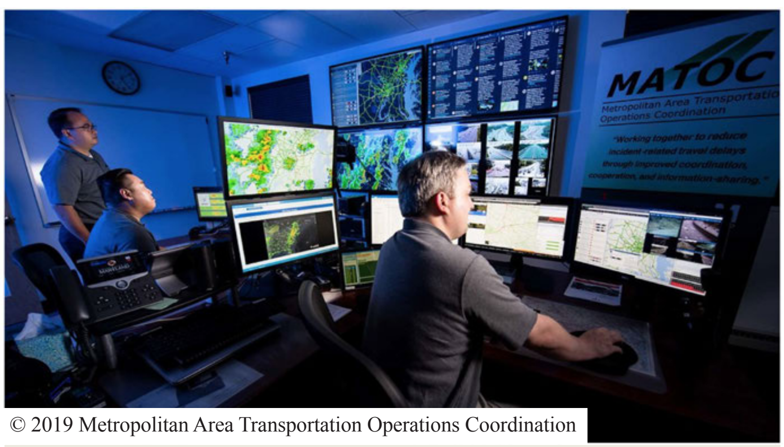




**Figure 9. Photo. The Metropolitan Area Transportation Operations Coordination operations room with TweetDeck prominently displayed on the upper-right media wall panel.**

**Source: Metropolitan Area Transportation Operations Coordination.**

The following provide direct access to Twitter feeds:

- The Twitter API (part of the Twitter Developer Platform).
- Through the Twitter application or website.
- Through a number of third-party Twitter applications (like TweetDeck) and data mining tools.

The Twitter API and examples of data from tweets are available at https://developer.twitter.com/en/docs/basics/getting-started.

Many agencies, however, will not directly integrate Twitter feeds into their ATMS platform, but will use third-party viewers of data for situational awareness. Figure 10 shows an example of a traffic alert from the Denver Police Department.


**Figure 10. Screenshot. The Denver Police Department uses the #Traffic hashtag to alert the public (and other agencies) about traffic events.**

**Source: Twitter/Denver Police.**

**Use Case: Pushing Data back to Crowdsourced Users.** Crowdsourced data can go two ways. Several agencies are now using Waze's CCP to provide agency data back to Waze users. This can come in the form of providing the location, times, etc. of road construction projects, special events, or collisions (figure 11). However, the Port Authority of New York and New Jersey (PANYNJ) is also using Waze to influence routing and driver behavior.

PANYNJ leverages Waze's website for agency partners to create realtime road closures when they need traffic to divert away from specific roads leading up to a terminal. When traffic congestion increases on a particular route, PANYNJ tells Waze about the road closure, which then prohibits Waze user's navigation systems from directing them onto that roadway.

A realtime closure can indicate temporary closure of a segment in one or both directions to all Waze users. When the realtime closure is active, the affected segment will be marked with red-and-white stripes, and Waze will not route any traffic through or onto the segment. Waze instructs users to report a realtime closure only when a road closure is complete in one or both directions.[4]

There are several possible ways to prevent Waze from routing traffic over a segment. Table 10, excerpted from Waze's technical support documentation, explains the options.

During a realtime closure, Waze will not route to a destination on the closed segments, nor partway through the closed segment, even if that is the closest segment to the destination. Instead, it will pick a stop point on the next closest segment. If the closed segment is much longer than the part of the road actually closed to all traffic, this can result in directing users to a nearby street even though they should be able to drive to the destination. For this reason, if a closure is very localized and going to persist for more than 1 week (e.g., a bridge replacement), it might be worth the effort to edit the closed segment.[5]

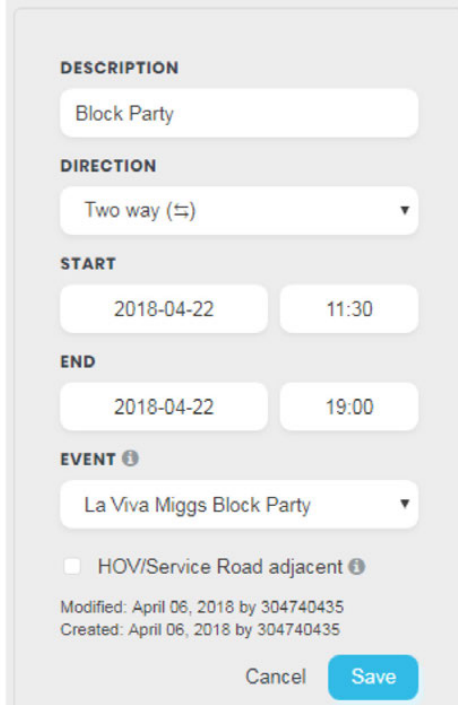





**Figure 11. Screenshot. Example of a crowdsourcing event.**

**Source: Waze.**

If the user is on a closed segment, Waze will find a route that begins in the closed segment and find the shortest distance to a segment with no closure.

**Table 10. Realtime closure options for Waze Connected Citizens Program agency members.**

| Method | Vehicles Affected | Takes Effect | Ends | Traffic Data | Guidance |
|---|---|---|---|---|---|
| Realtime closure | All | Immediate | Expires | Kept | Preferred option for temporary (even long-term) one-way or two-way closure. Visible to drivers. Immediate effect. Automatically removed when it expires. |
| Road direction change | All | Tile update | Permanent | Kept | Use only for permanent change in direction from two-way to one-way. |

4 Wazeopedia, "Realtime closures." Available at: https://wazeopedia.waze.com/wiki/USA/Real_time_closures, last accessed March 17, 2019.

5 Wazeopedia, "Realtime closures." Available at https://wazeopedia.waze.com/wiki/USA/Real_time_closures#edit_the_closed_segment, last accessed March 18, 2019.

**Table 10. Realtime closure options for Waze Connected Citizens Program agency members. (continued)**

| Method | Vehicles Affected | Takes Effect | Ends | Traffic Data | Guidance |
|---|---|---|---|---|---|
| Time-based segment restriction | Some | Tile update | Optionally expires | Kept | Use only where the restrictions (time of day/day of week) are permanent, or where certain vehicle types are allowed or prohibited. Example: No passenger cars on weekdays. |
| Time-based turn restriction | Some | Tile update | Optionally expires | Kept | Use where travel on the segment is allowed, but turns onto the segment are temporarily forbidden or else permanently forbidden at certain times of day, days of week, or for certain vehicle types. Example: No left turn 4:00pm-6:00pm. |
| Permanent turn restriction | All | Tile update | Permanent | Kept | Use where a turn onto the segment should be permanently forbidden for all vehicles. |
| Conversion to a penalized road type | All | Tile update | Permanent | Kept | Penalties make routing less likely, but are not absolute. Vehicles with a destination on the segment will be routed onto the segment. Normally private road is used. |
| Disconnection | All | Tile update | Permanent | Lost | Use only if the disconnection is permanent. All turn data is lost. |
| Deletion | All | Tile update | Permanent | Lost | Use only if the road is permanently closed. All data is lost. |

Source: Wazeopedia, "Realtime closures." Available at https://wazeopedia.waze.com/wiki/USA/Real_time_closures, last accessed March 17, 2019.

**Use Case: Utah DOT Crowdsources Maintenance** The Utah DOT (UDOT) worked with SeeClickFix to customize a solution for citizen reporting of roadway maintenance issues. SeeClickFix is an application that allows citizens to take pictures of maintenance issues and send them to the agency (along with geo-located photos, attributes, etc.). In addition to collecting citizen requests, the solution acts as a service-request management system that has helped to improve agency efficiencies and transparency. The Utah Click'n Fix website (figure 12) even prints out service reports to show how the agency is managing requests over time. When more than one citizen service request comes in for a single issue, consolidating and handling the requests as one saves time on maintenance issues. This consolidated work means that employees no longer have to write 50 emails or make 50 phone calls in response to 50 separate citizen service requests. Utah citizens can report issues through UDOT's smartphone app, which empowers citizens and the DOT to resolve maintenance and safety issues more quickly (figure 13).

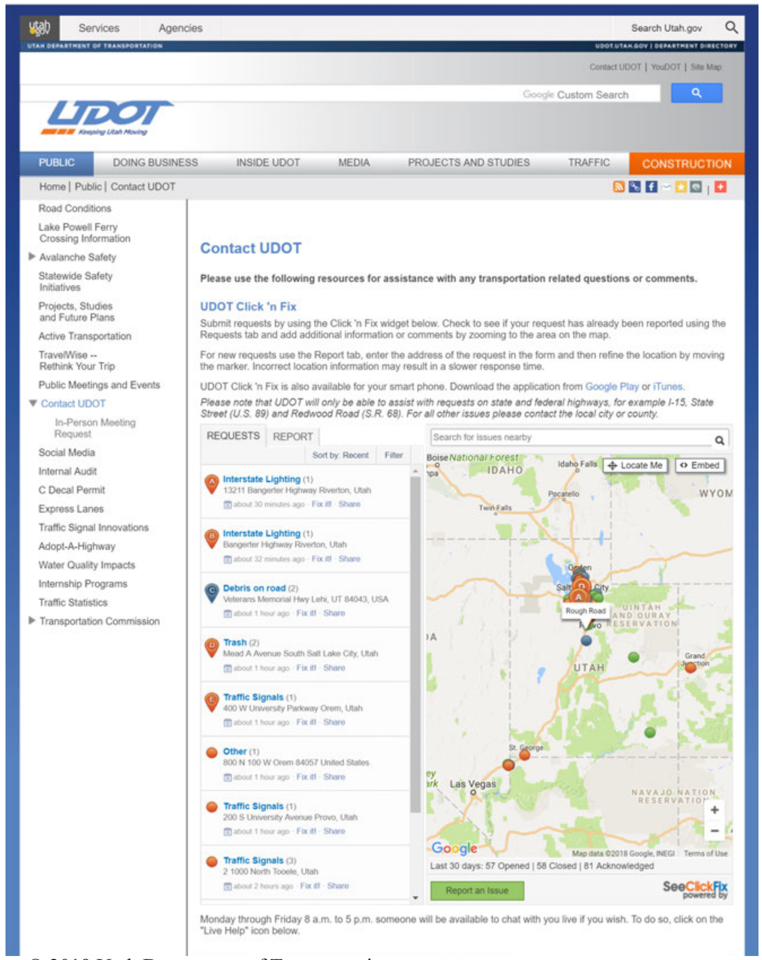





**Figure 12. Screenshot. Utah Department of Transportation's Click'n Fix widget accompanies the Click'n Fix application.**

**Source: Utah Department of Transportation, Contact Website. Available at: https://www.udot.utah.gov/main/f?p=100:pg:0:::T,V:376.**

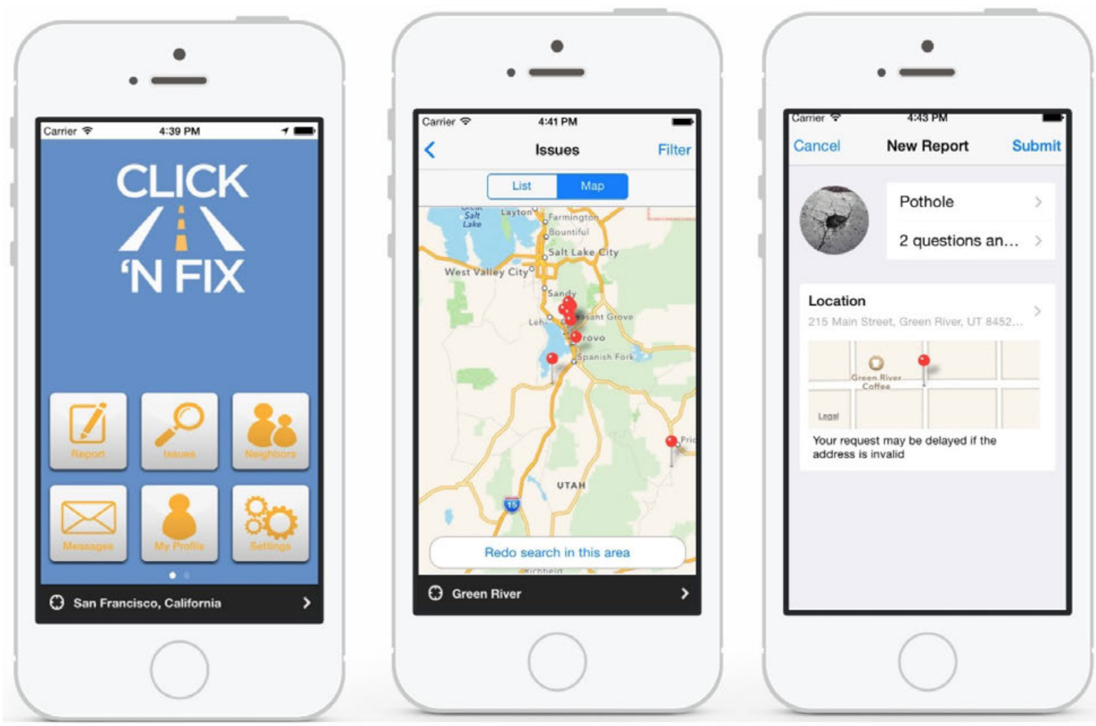





**Figure 13. Illustration. Utah Department of Transportation's smartphone app empowers citizens and the agency to resolve maintenance and safety issues more quickly.**

**Source: Utah Department of Transportation, SeeClickFix Website. Available at: https://seeclickfix.com/pages/case-studies/utah-dot.html.**

## ROADSIDE BASIC SAFETY MESSAGE DATA

### Description

CVs and infrastructure can use special communications protocols, such as dedicated short-range communications (DSRC), 5G (the fifth generation of wireless technology), and others, to exchange basic safety messages (BSM) and infrastructure messages (figure 14). These communications protocols allow vehicles and infrastructure to exchange messages within 1,000 meters approximately 10 times per second primarily to support safety applications, and secondarily to support mobility applications. Under BSM Part 1, vehicles communicate their size, position, speed, heading, acceleration, and brake system status to

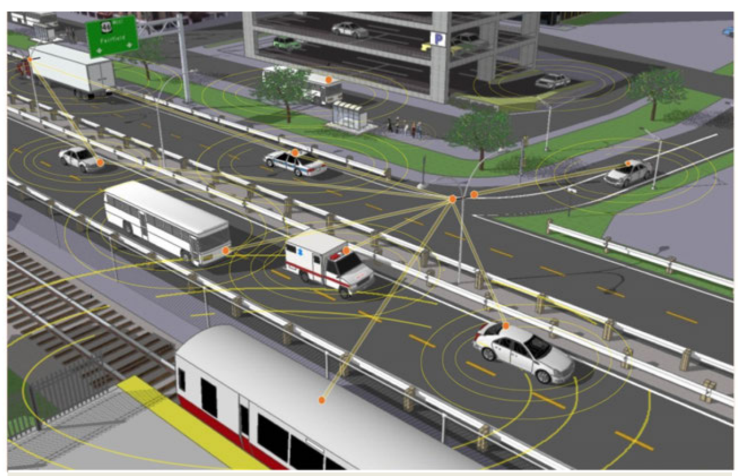


**Figure 14. Illustration. U.S. Department of Transportation connected vehicle and connected infrastructure concept.**

**Source: U.S. Department of Transportation, Intelligent Transportation Systems Joint Program Office. Available at: https://www.its.dot.gov/itspac/october2012/PDF/data_availability.pdf.**

each other and to the infrastructure.[6] BSM Part 2 includes additional data elements, such as weather data and vehicle status data. Similarly, connected infrastructure can transmit its status to vehicles or perform action in response to received BSM data. Although poorly defined, infrastructure message format does include some potential message types that represent digital descriptions of the roadway (e.g., signal phase and timing (SPaT) messages or MAP type messages). In addition to these, infrastructure messages could include information regarding speed limits, especially as they change in work zones and school zones; dynamic message sign information; and other advisories.

## Applications

BSM data has many potential uses, including:[7]

- Vehicle platooning (speed harmonization, cooperative adaptive cruise control, etc.).
- Queue warnings.
- Intelligent traffic signal system.
- Transit signal priority.
- Mobile accessible pedestrian signal system.
- Emergency communication and evacuation.
- Work zone alerts.

## Attributes

BSM is broadcast approximately 10 times per second, which is sufficient for most basic safety applications. However, many roadside unit (RSU) providers also have other data transmission capabilities that can have latency as low as 20 ms and as high as several seconds depending on the application.

## Details

- BSM Part 1 provides only core information needed for immediate safety applications:
  - Vehicle size.
  - Position.
  - Speed.
  - Heading.
  - Acceleration.
  - Brake system status.

6 For more information about BSM Part 1 and 2, see B. Cronin, "Vehicle Based Data and Availability," presentation (n.d.). Intelligent Transportation Systems Joint Program Office, Research and Innovative Technology Administration, USDOT. Available at: https://www.its.dot.gov/itspac/october2012/PDF/data_availability.pdf.

7 For more details, see B. Cronin, "Vehicle Based Data and Availability," presentation (n.d.). Intelligent Transportation Systems Joint Program Office, Research and Innovative Technology Administration, USDOT. Available at: https://www.its.dot.gov/itspac/october2012/PDF/data_availability.pdf.

- BSM Part 2 promises additional data elements, such as:
  - Recent braking.
  - Path prediction.
  - Throttle position.
  - Differential GPS.
  - Stability control.
  - Exterior light status.
  - Wiper status.
  - Ambient temperature.
- Infrastructure messages can contain information such as:
  - Current signal phase and residual time.
  - Signal request and status.
  - Pedestrian status.
  - MAP information – digital representation of the intersection.
  - Air quality.
  - Roadway friction information.
  - Traveler information.

### Data Availability

Current data coverage is limited to pilot sites and test sites, at least in terms of data that is available to TMCs. As connected vehicles and infrastructure deployments proliferate, the coverage will most likely increase as well.

Many vehicle manufacturers and original equipment manufacturers (OEMs) are building out CV functionality. However, it is not clear what data may be available for agencies to consume beyond existing pilot deployments and test sites. OEMs are building their own cloud services or partnering with other third parties that are based on the concept of bundling connected data and selling it to interested parties, including the public sector.

For example, at least one company acts as a neutral third party that provides the service of collecting and transmitting CVs data to interested parties using APIs.

At this time, the United States Department of Transportation (USDOT) has provided access to a number of CV and infrastructure data sets from pilot implementations and test sites through its public data portal (table 11).

**Table 11. Example connected data sets.**

| Example Connected Data Source | Description |
|---|---|
| Wyoming Connected Vehicle (CV) Pilot | Basic safety message (BSM) from Wyoming CV Pilot project. |
| Belle Isle Road Weather Demonstration | Road weather observations collected by several CVs over a period of several months in Belle Isle, Michigan. |
| Minnesota Department of Transportation Mobile Observation | Data from instrumented snowplows and light-duty pickups. |
| Intelligent Network Flow Optimization (INFLO) Prototype | Small-scale demonstration of INFLO prototype system in Seattle, Washington, including 21 vehicles exchanging BSM with roadside units (RSU) and transmitted to the transportation management center. |
| Multi-Modal Intelligent Traffic Signal Systems Study | Study that collected vehicle trajectories by capturing BSM from CV via RSUs. This study also exchanged signal phase and timing message with connected vehicles in the intersection. |

Source: United States Department of Transportation, Intelligent Transportation Systems Joint Program Office.
Available at: https://www.its.dot.gov/data/search.html.

On the infrastructure side, a number of CV communication and network companies provide connected infrastructure functionality; e.g., traffic signals, critical hazard alerts, and roadway weather conditions. Third-party companies have developed applications that utilize infrastructure data to support safety and mobility applications.

## Pros and Cons

CV and infrastructure data allows drivers to be safer through better information regarding vehicle surroundings and environment. It also allows agencies to disseminate important and tailored information directly to drivers and vehicles in a much more integrated fashion than was traditionally possible using dynamic message signs (DMS) or highway advisory radio (HAR).

CV and infrastructure data can be very large and rapid, which provides a challenge for agencies to process it in realtime and use it to guide decisions or manage congestion. It is unclear what level of aggregation (if any) is necessary for data to be useful in each application.

The current level of market penetration for CVs and connected infrastructure is low, therefore, it has minimal use for TMCs focused on realtime operations at this time. However, deployment of this technology will continue across the country with more data becoming available every day.

## Use Cases for Roadside Basic Safety Message Data

**Use Case: Wyoming Connected Vehicle Pilot.** As part of the Wyoming CV Pilot, the University Corporation for Atmospheric Research (UCAR) has implemented a Pikalert system (figure 15) that combines vehicle-based measurements with traditional weather observations to provide alerts to CVs traversing the corridor. Pikalert consists of several components:

- Vehicle data translator.
- Enhanced maintenance and decision support system (EMDSS).
- Motorist advisory and warning (MAW) application.

***Vehicle Data Translator***

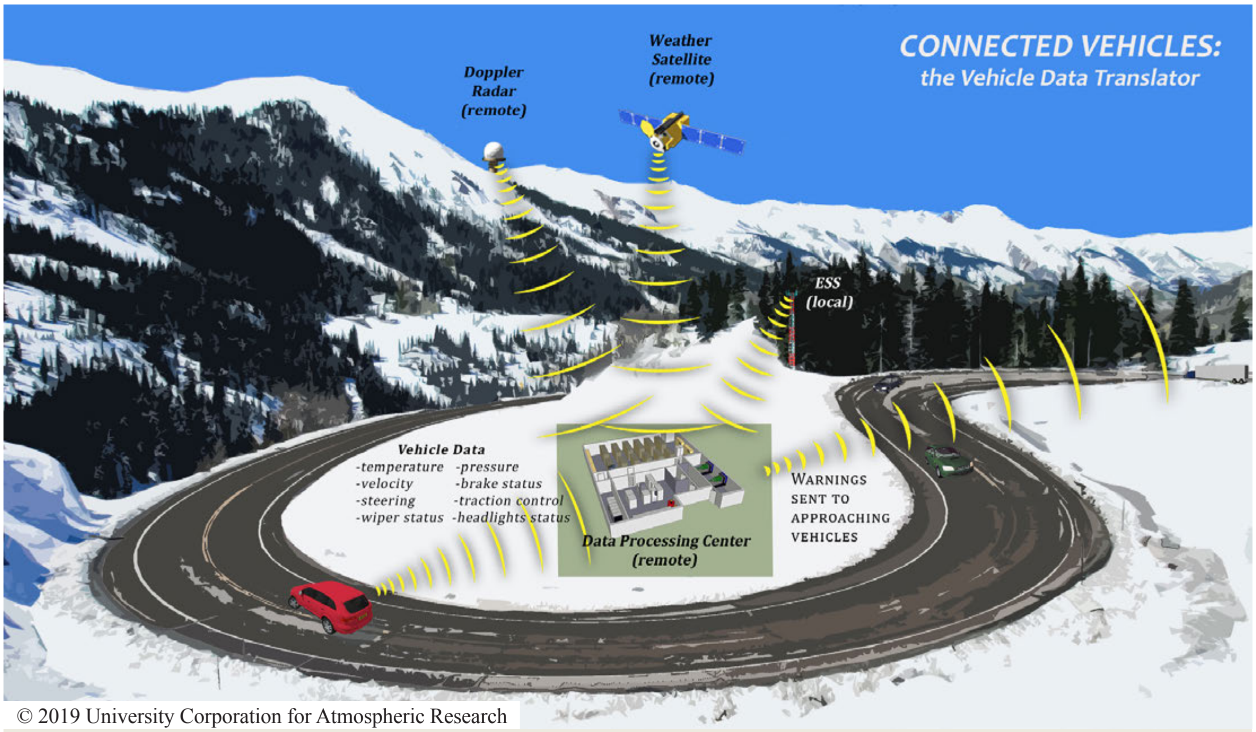


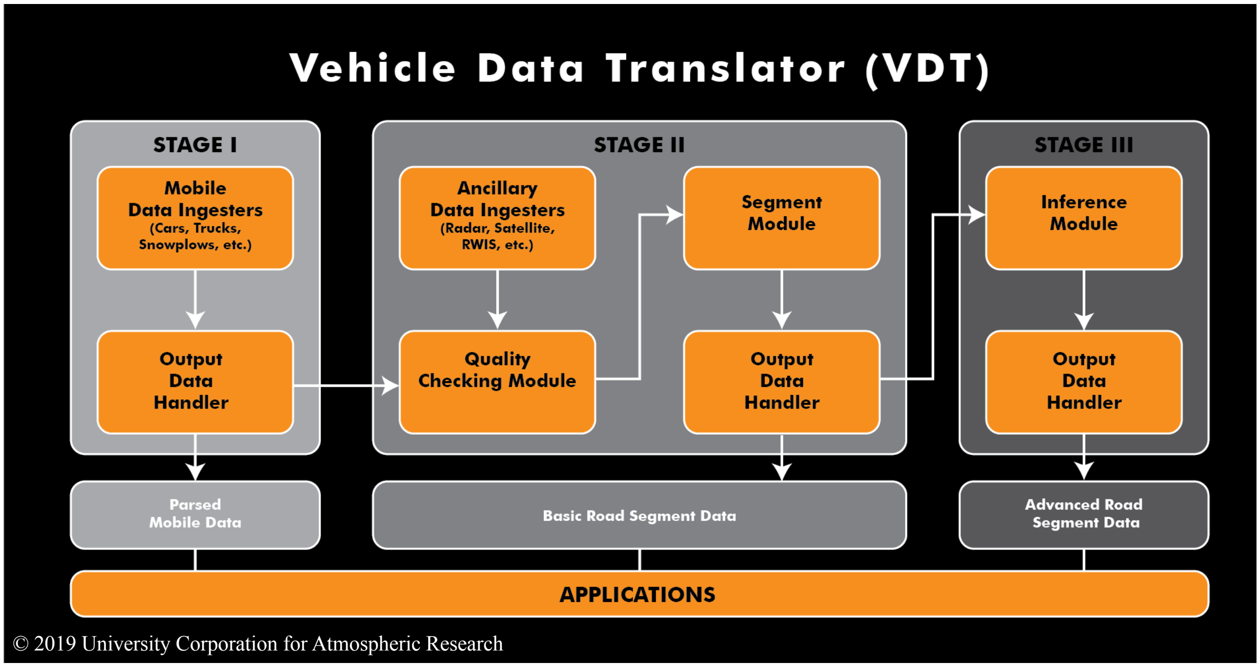

**Figure 15. Illustration. The Pikalert concept.**

**Source: University Corporation for Atmospheric Research, Research Applications Laboratory, "Promoting Vehicle Safety, Mobility, and Environmental Efficiency" Web page. https://ral.ucar.edu/solutions/promoting-vehicle-safety-mobility-and-environmental-efficiency.**

The Vehicle Data Translator consists of three stages:

- ***Stage 1:*** CVs with a controller area network bus (CANBus) and aftermarket sensors provide data from several sources. The data elements include barometric pressure, windshield wipers settings, headlight status, ambient air temperature, speed and heading, adaptive cruise control, location and elevation, hours of operation, anti-lock braking system (ABS) and brake status, stability and traction control, yaw/pitch/roll, accelerometer, steering angle, and differential wheel speed. Data is checked for quality; sorted by time, road segment, and grid cell; and published as parsed mobile data (figure 16).

- ***Stage 2:*** Ancillary data such as that from radar and road weather information system (RWIS) is collected. The quality of this data is checked and published as basic road segment data.
- ***Stage 3:*** Variables are inferred using data from the previous two stages. For example, wiper activity in combination with weather radar, satellite, and temperature data can indicate precipitation type and intensity. Similarly, headlight status, wiper activity, and RWIS information can define visibility measurements, and ABS and traction control and weather radar can indicate pavement conditions.




RWIS = road weather information system.

**Figure 16. Diagram. Vehicle Data Translator Architecture.**

**Source: University Corporation for Atmospheric Research, Research Applications Laboratory.**

***Enhanced Maintenance and Decision Support System.*** EMDSS incorporates CV data into a forecast and decision process. Traditional sensors and equipment-generated data is enhanced and supplemented by CV data that provides more robust coverage and more detailed information about conditions on the entire corridor. This combined information supports maintenance operations and provides the opportunity for a proactive approach in handling adverse weather conditions. This proactivity improves safety and provides an opportunity for more effective use of limited resources by targeting the most critical problem spots before issues arise.

***Motorist Advisory and Warning Application.*** MAW capitalizes on the rich output of vehicle data translators to provide travelers with hyperlocal and near-realtime road weather information, as well as accurate 24-hour forecasts of road weather conditions.

# REALTIME AND ARCHIVED TRAJECTORY DATA

## Description

Trajectory data is time-stamped location data from vehicles, cell-phones, or other GPS-enabled devices moving throughout a network. This is sometimes referred to as "bread-crumb trail" data.

Figure 17 illustrates an example of trajectory data where each red dot represents a ping from the vehicle as it moves from point 0 to point 14.

The system can make data anonymous in different ways. For example:

- Rotation of the unique vehicle identifiers on a set time interval ensures that vehicles cannot be tracked for multiple days in a row.
- Trips may be "clipped" when entering or exiting residential neighborhoods to keep from pinpointing home addresses.

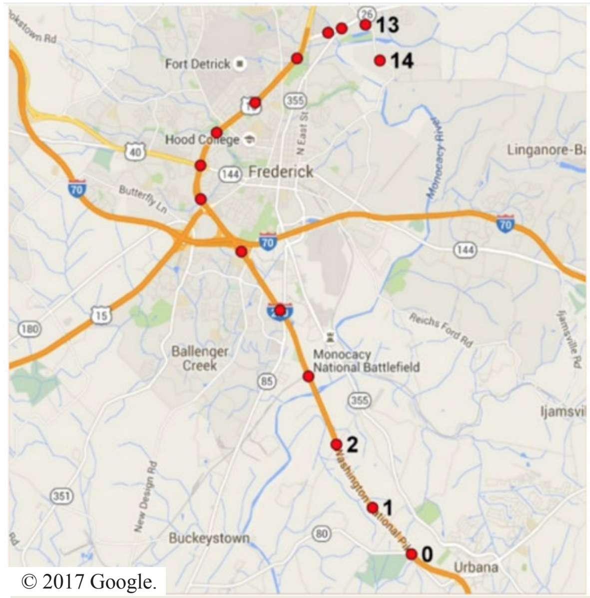



**Figure 17. Map. Example trajectory data for a single trip.**

**Source: Google.**

Because of the privacy protocols, probe vehicle data providers add 1 to 2 days of latency prior to delivery to agencies. However, technology and capabilities are improving each day with anticipation of this data set being available in near-realtime in the next year or two. The data has many application areas for both planning and operations.

## Applications of Realtime Trajectory Data

Trajectory data is a new dataset with a great deal of potential for operations management. Realtime uses of trajectory data include:

- Realtime traffic pattern analysis
  - Operators can look at individual vehicle trips to evaluate corridor demands.
  - In case of an incident or congestion, realtime trajectory data can show the effectiveness of implemented detours or self-detouring patterns.
  - Evaluate impacts of special events that may result in creation of a temporary significant trip origins or destinations and route utilization.
- Multi-modal system utilization
  - Trajectory data can show where and when mode transitions occur and provide operators with ability to influence traveler decisions based on network utilization patterns.

In addition to realtime uses, trajectory data is also valuable in operations planning. For example, planners can use trajectory data as follows:

- Trip patterns between jurisdictions
  - Traditional O-D analysis to identify trip origins and destinations, work versus leisure travel, etc.
  - Waypoint analysis to determine if traffic in a specific area (State, county, traffic analysis zone, business center) originated in the same area, neighboring area, or another more-distant location. This can identify whether certain corridors are mainly local travel or pass-through corridors, etc.
  - Analyze trip clusters to evaluate effectiveness of existing transit service or identifying areas in need of new transit service.
  - Analyze trip patterns that may impact critical freight corridors or ports.

## Attributes

For each individual trip, trajectory data usually includes:

- Unique device/vehicle identifier.
- Unique trip identifier.
- Departure time and location (trip origin).
- Periodic waypoints during the trip, including:
  - Latitude/longitude.
  - Timestamp.
  - Instantaneous speed/heading.
  - Identifier of road segment for waypoint.
- Arrival time and location (trip destination).

Identification of each origin, destination, and waypoint uses latitude/longitude pairs and timestamps. Some providers “snap” these latitude and longitude points to a particular road segment. Collection of waypoints at varying intervals depends on the probe vehicle type and location and can occur from once per second up to once every 5 minutes.

## Data Availability

Trajectory data is relatively new to the market, and only a few companies are providing access to this information. At least one company provides analytics services on top of other third-party data providers and location-based service providers; however, those services may not provide direct access either to individual trips or to the raw trajectory data. Ride-hailing companies are also starting to make limited O-D datasets available to select researchers in limited metropolitan areas. These data sometimes have heavy restrictions on use, limit the O-D analysis to larger geographic zones, and do not provide route choice analysis. As the demand for this data rises and technology improves, other companies with access to “probe-like” data will probably offer similar data sets.

## Pros and Cons

A huge benefit of the trajectory version of O-D data is the ability to construct an entire trip based on origin, destination, and the route taken during the trip. No other data set on the market has the potential to provide such rich information about route and mode choice. Trip data from third parties is more cost effective than travel diaries or roadblock studies. However, there is a 1 to 2 day lag between when the data is collected and when it is available to an agency. This makes the data particularly useful for after-action reviews (AARs) and otherwise understanding the greater impacts of operations decision on the traveler. In the near future, this information will be available in realtime or near realtime, thus making it more applicable to realtime signal control, ramp metering, dynamic routing, and more.

One of the challenges is the size of the data set. For example, data for the relatively small State of Maryland for one year of data includes nearly 100 million individual trip records, and over 7 billion waypoints. Analysis of data of this size requires dedicated information technology (IT) infrastructure and expertise—either in-house or provided by a third party.

Latitude/longitude pairs identify most waypoints, and only a select few providers pre-map the waypoints to TMC-segments or OpenStreetMap (OSM) segments. This means that data users may need to be able to conflate those locations to an underlying map to be able to map the trips to specific roadways and corridors (figure 18). Because latitude/longitude pairs may not be exactly accurate due to GPS errors and noise, in dense urban areas or dense roadway networks, subsequent waypoints may appear to jump from one road to another and back as the probe vehicle traverses the network. Collecting waypoints at longer time intervals makes it more difficult to determine the exact route taken between one ping and another. This presents a challenge when attempting to construct a trip using those waypoints. A benefit of machine learning algorithms is that they allow users to snap the waypoints to the network with a high degree of accuracy. However, implementing machine learning algorithms requires understanding data science, geographical information systems (GIS), and how machine learning works.

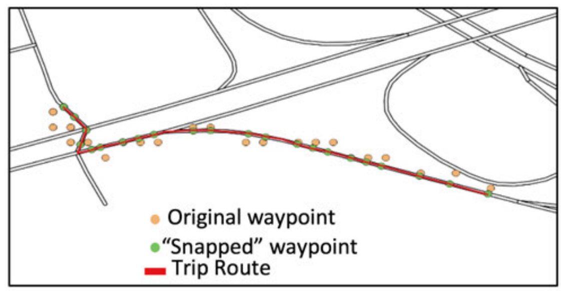





**Figure 18. Illustration. Snapping waypoints to routes can sometimes be a challenge.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

## Use Cases for Realtime and Archived Trajectory Data

**Use Case: Impact of New Tolling on Travelers and Mode Choice.** The Virginia Department of Transportation (VDOT) recently implemented high occupancy tolling on I-66 leading into the District of Columbia. These variable tolls can range from $4 to $47 per trip. Realtime and historic trajectory data can show the impacts of these tolls on route choice. VDOT is using trajectory data to understand which routes motorists took to get into the District prior to the tolling going into effect. They are then evaluating how motorists' routes changed after implementing the tolling. A future application could be to evaluate—in realtime—how changing the rates throughout the day is affecting trips, which arterials are taking the brunt of the extra traffic, and the exact routes that commuters are taking to avoid the tolls.

**Use Case: Impact of Construction on Travelers.** Several DOTs and metropolitan planning organizations (MPOs) leverage O-D (trip) data with trajectories and routing details for realtime operations. Transportation systems management and operations (TSMO) groups can use this data to understand the impacts of work zone management practices, develop different traveler information and communication strategies, establish more effective signal timing plans, and support freight operations. These same agencies can use the trajectory data to support before-and-after studies to show how the finished construction projects have changed route choice.

**Use Case: Maryland DOT Mid-block Signal Timing Analysis.** Trajectory data has the power to support signalized arterial applications. Collection of waypoints throughout a trip enables collection of true travel times within cities among a very large number of routes. Whereas standard probe data provides average trip speeds from intersection to intersection, trajectory data can provide mid-block to mid-block travel times, allowing agencies to quickly understand turning movement travel times and overall signal performance (figure 19).

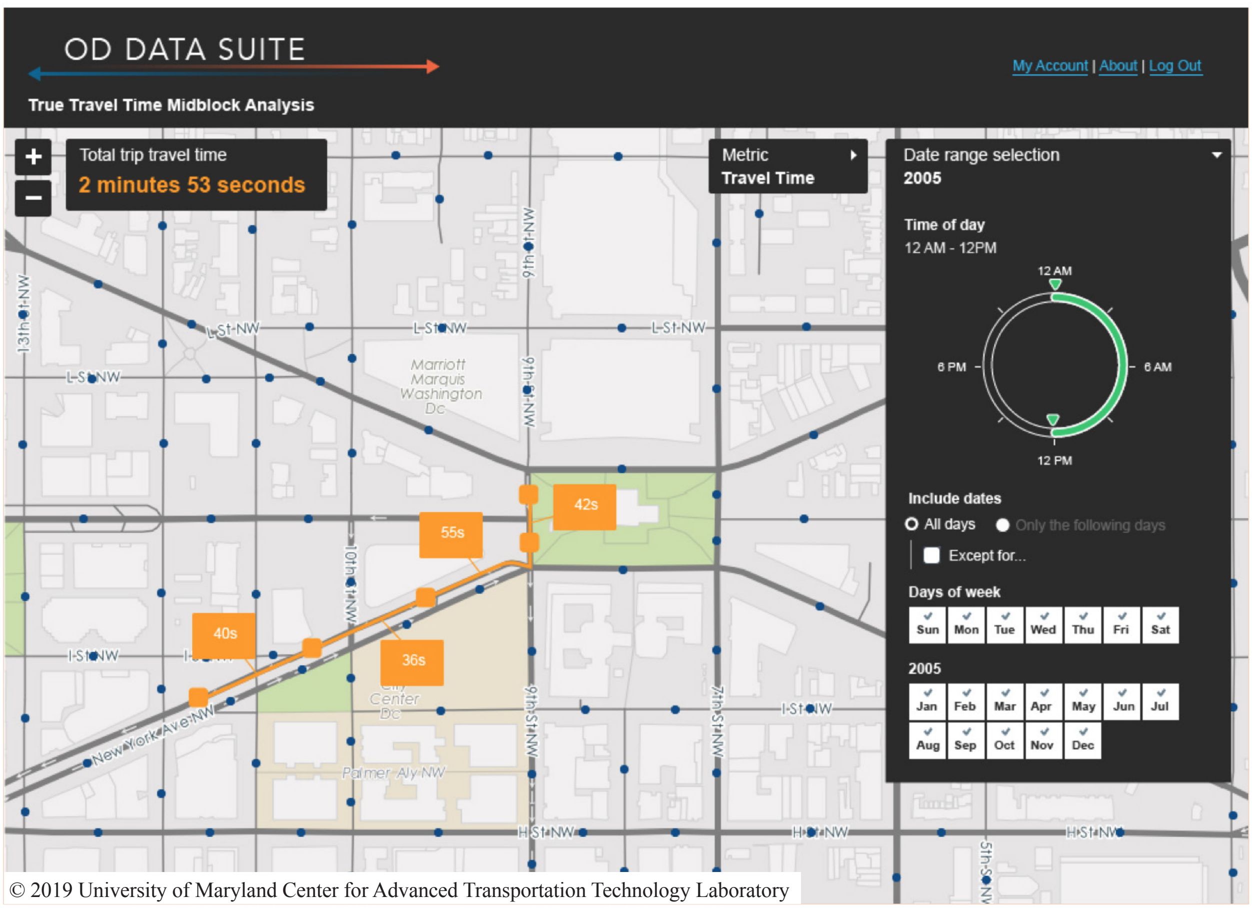



**Figure 19. Illustration. The Center for Advanced Transportation Technology Laboratory's Origin-Destination Data Suite uses trajectory data to conduct midblock analyses.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

The Maryland DOT is funding the development of a mid-block travel-time analysis suite using INRIX trajectory/trips data. This tool will provide a ranked list of all turning movements in a city, allowing signal operators to understand the user delay cost associated with all turning movements in a city (figure 20). The graph in the lower right shows how many trips made it through this particular intersection (left turn) within the first cycle, second cycle, and third cycle.

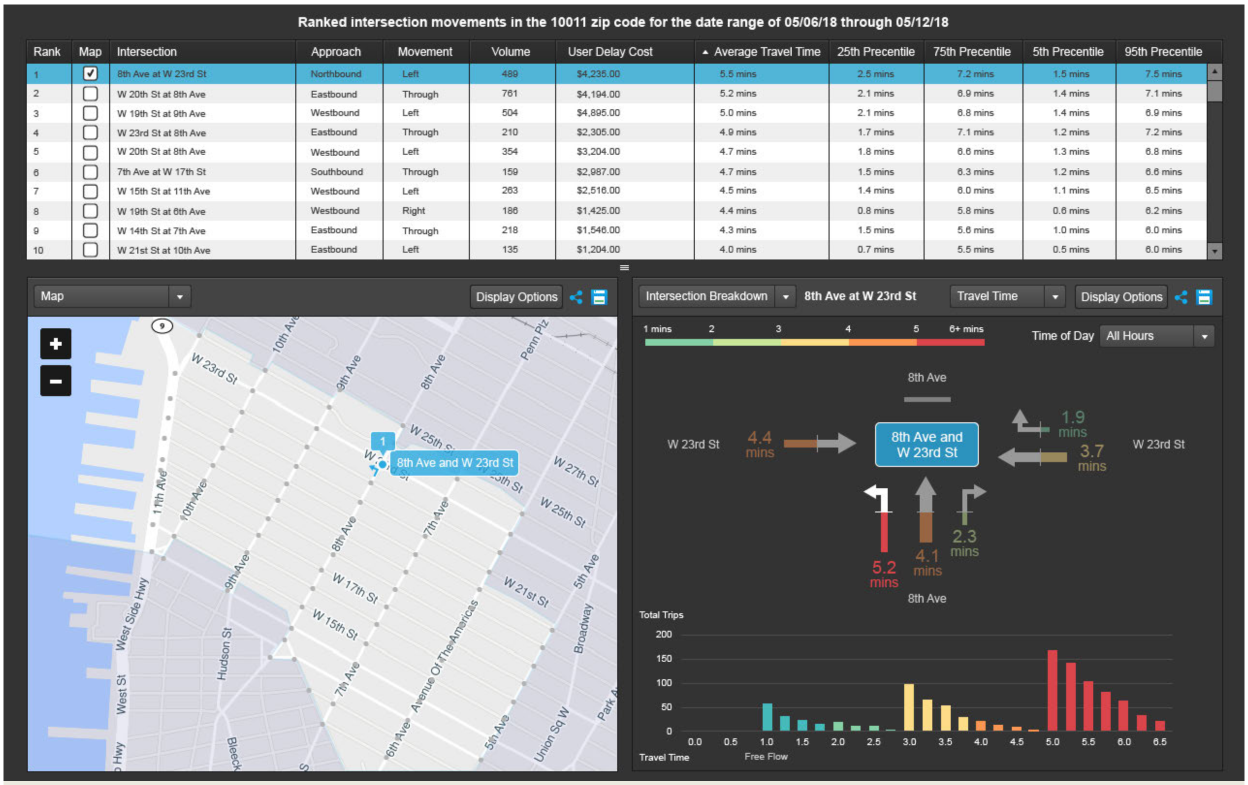





**Figure 20. Screenshot. The Center for Advanced Transportation Technology Laboratory's Origin-Destination Data Suite provides ranked intersection movements by zip code and date range based on trajectory data.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

INRIX trip data comes in several formats with multiple output files that include metadata (table header definitions, source data, device types, waypoint details for each trip, origins, destinations, etc.). While the entirety of the files is too expansive to include in this document, figure 21 is an example output from one of these many files describing trips and the waypoints along a particular trip. This cropped file shows only eight waypoints of a much longer trip. The expanded file shows thousands of additional trips and waypoints.

| TripId | WaypointSequence | CaptureDate | Latitude | Longitude | SegmentId | ZoneName | Frc | DeviceId | RawSpeed |
|---|---|---|---|---|---|---|---|---|---|
| e8ff8a4ca972a2524e5ce7de78c62f9c | 4 | 2018-07-04T00:31:18.000Z | 29.94011 | -90.10623 | 1524482843 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 62 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 5 | 2018-07-04T00:31:27.000Z | 29.94052 | -90.10783 | 1524482843 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 55 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 6 | 2018-07-04T00:31:37.000Z | 29.94133 | -90.10894 | 1524482843 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 35 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 7 | 2018-07-04T00:32:07.000Z | 29.94316 | -90.11223 | 1524486836 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 65 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 8 | 2018-07-04T00:32:17.000Z | 29.94452 | -90.11305 | 1524281492 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 74 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 9 | 2018-07-04T00:33:47.000Z | 29.95298 | -90.11667 | 1524387235 | New Orleans CBD | 2 | ea74443ff04a6b598c39a30bec3f19a4 | 58 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 10 | 2018-07-04T00:35:28.000Z | 29.95584 | -90.12019 | 1524553142 | New Orleans CBD | 3 | ea74443ff04a6b598c39a30bec3f19a4 | 11 |
| e8ff8a4ca972a2524e5ce7de78c62f9c | 11 | 2018-07-04T00:35:38.000Z | 29.95641 | -90.11985 | 1524629560 | New Orleans CBD | 3 | ea74443ff04a6b598c39a30bec3f19a4 | 46 |

**Figure 21. Screenshot. Example data feeds.**

**Source: INRIX.**

INRIX trip data shows actual trajectories and waypoints for individual trips, which allows an agency to conduct true route analysis, travel-time analysis, and more for individual trips. The data is quite voluminous because it contains individual trip data, and third parties have developed a number of tools to help agencies better utilize the data.

**Use Case: Maryland DOT Identifying Travelers Impacted by a Project.** Leveraging trajectory data enables agencies to perform outreach and education prior to major construction projects, detours, etc. Figure 22 shows how the Maryland DOT-funded O-D Analytics suite shows where trips that passed over a very specific road segment originated. The same tool also shows the destination of those same trips.

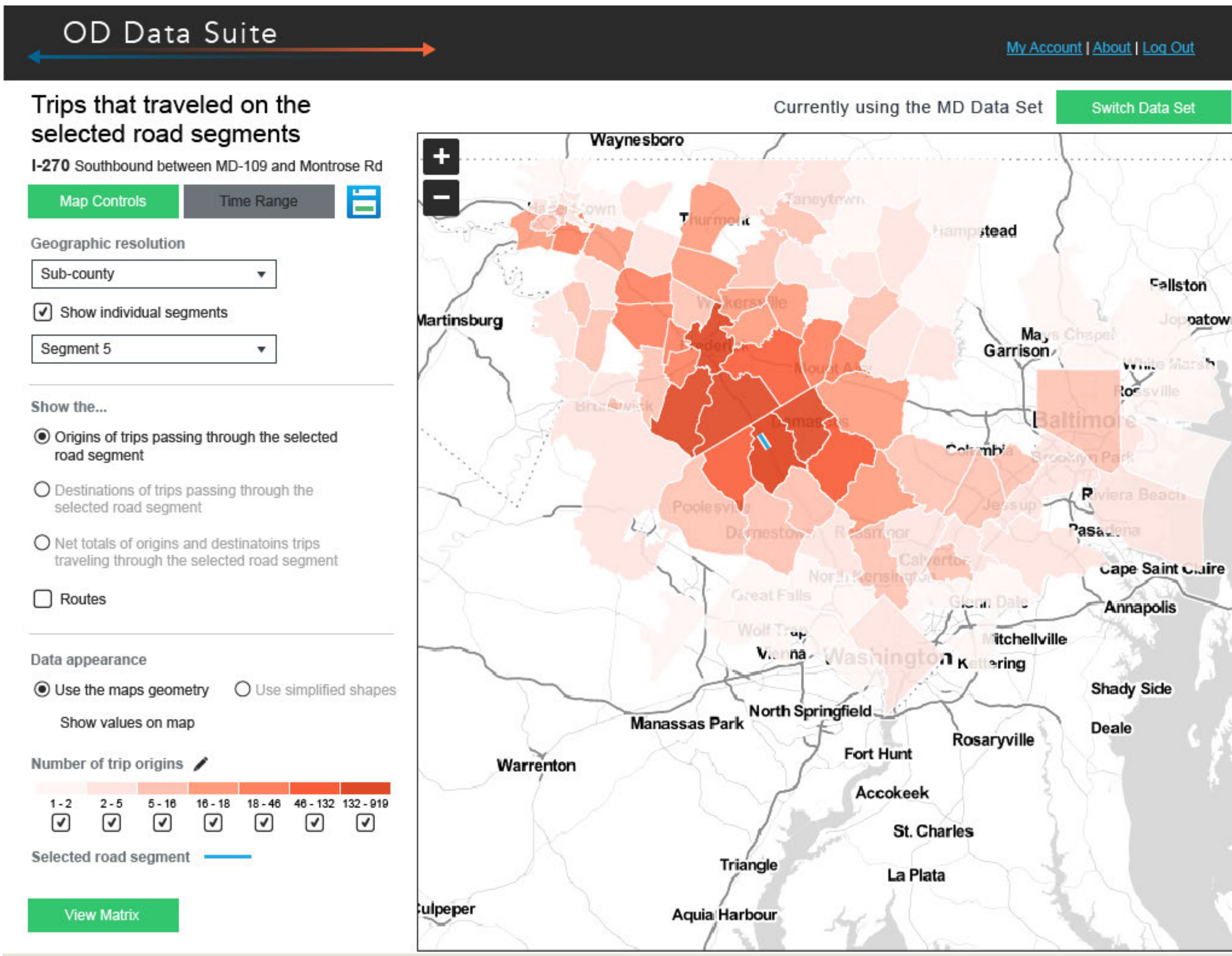





**Figure 22. Screenshot. The Origin-Destination Analytics Suite illustrates the origins for trips that passed over a very specific road segment.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

# CROWDSOURCED MAP DATA

## Description

Traditionally, public and private sector entities generated map data for specific uses. In transportation, most agencies purchase map data from one of the major map data providers. However, over the last several years, crowdsourced maps have become more prevalent and more viable for use in transportation operations. Public apps and websites create and update crowdsourced maps in near realtime or on a small delay if the contributor community needs to verify the changes. Most of the time, these maps are freely available (figure 23). Private sector map providers still use crowdsourcing to improve their maps, but the map data is still proprietary.

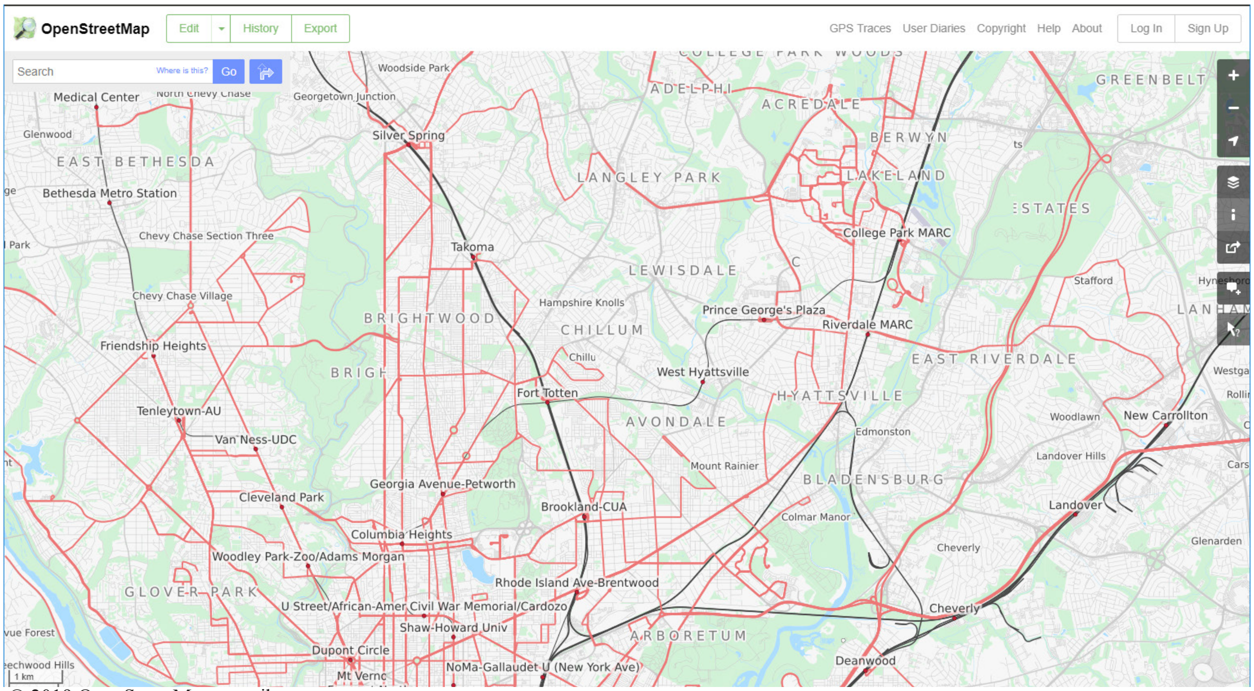





**Figure 23. Screenshot. Example OpenStreetMap.**
**Source: OpenStreetMap.**

## Applications of Crowdsourced Map Data

Crowdsourced map data allows agencies to have access to more accurate and more frequently updated map data than available with traditional maps. It does not necessarily change the way TMCs use maps. Instead, many TMCs use crowdsourced map data in their ATMS or for their traveler information systems.

## Attributes

Crowdsourced map data usually consists of the same or similar data elements as any traditional map, including roadway networks, landmarks, businesses, parks, etc. In general, crowdsourced map data is not a different data product, but rather a differently created data product.

- **Latency and Frequency:** Updates to crowdsourced map data occur more frequently than traditional map data because crowdsourced mapping relies on constant input from customers in the field. Crowdsourced map data can sometimes be updated in realtime, allowing users to attach images, reviews, and other information. If not updated in realtime, data updates usually occur with low latency, which typically is the result of community validation or another type of crowdsourced verification to ensure the information is accurate.
- **Details:** The level of detail in crowdsourced maps varies. For example, raw open source data may contain only basic map data, such as roadway geometries, landmarks, etc. Similarly, another company provides realtime routing; therefore it contains a high level of detail related to roadways and speed limits, although it may be lacking when it comes to landmarks and non-transportation assets and information. Other map providers have a diverse customer base and provide many services, including directions and routing by mode (walking, biking, driving, transit, etc.), landmark and business information, traffic conditions, etc. As a result, their data tends to have a high level of detail in many different contexts; however, not all of it is available for purchase by agencies and TMCs.
- **Quality and Coverage:** Similar to crowdsourced incident and congestion data, quality and coverage of crowdsourced map data relates to the number of participants contributing information. This means that urban and densely populated areas usually result in good data quality and coverage, while less-populated areas may have lower quality and coverage. Due to this, crowdsourced map data is considered a supplemental data set rather than on its own, unless it is for a specialized purpose, such as a routing application.

### Data Availability

There are several crowdsourced map data providers including free, open-source options such as OSM and other private-sector providers. Many of these providers use large user communities that contribute map information, including GPS traces, link creation and identification, landmark and asset information, etc. Others may use technology deployed in vehicles to analyze surroundings and update map information such as speed limit information, restrictions, roadway changes, etc.

While each crowdsourced map data provider offers similar data elements, they collect data in different ways. A large community of global map contributors edits the OSM data. They collect and add map data elements and layers and publish them under the open source license. This approach enables data to be updated frequently and accessible to anyone interested in using the map. Map users are free to enhance the OSM map by adding information to it, with a caveat that users need an open source license to publish added data. While, conceptually, the OSM model appears to be ideal for rich and accurate mapping, the open source caveat often results in users not being willing to contribute their additions to the open source community, since their data is either proprietary, valuable to the contributor, or private in nature.

OSM data is truly free to use. It can be downloaded fully by any agency with few restrictions. While many agencies also think of other maps as being free, one may not actually download the raw basemap data from many common map providers. OSM also includes user-generated map data that goes beyond roads. For example, contributors may have mapped out the locations of nearby items of interest, such as trees, fire hydrants, public restrooms, potholes, etc.

OSM data have some limitations, which may or may not affect traffic operations. Because the OSM system depends on volunteers to produce and edit maps, data quality and consistency can vary from location to location around the world. Metadata can sometimes be lacking, which makes it difficult to know if certain layers are current. The data is also not completely authoritative. While the road network was derived from U.S. Census data, making it fairly trustworthy, other layers (e.g., points of interest, location of trees, etc.) may be less authoritative.

Because of the OSM limitations listed above, a number of third-party map distributors have entered the market to address these limitations. These distributors enhance raw OSM maps and provide additional data elements, layers, and hosted services at a cost.

Most crowdsourced map data providers offer several different ways to obtain and use this data. TMCs are able to purchase raw map data, such as shapefiles and associated attributes, and integrate that data into their native ATMS. Alternatively, TMCs can procure hosted services in forms of tiling services or other map services. These hosted services allow TMCs to embed links to an external map service or to build maps in realtime by querying tiling services hosted elsewhere.

OSM data (figure 24) are available as a topological data structure consisting of four primary elements: nodes, ways, relations, and tags. This data is stored in a primary database that hosts all edits and is the primary source of all OSM data output formats. In addition to raw data, users are able to obtain individual GPS traces submitted by contributors.

Note that figure 24 illustrates the data that is used to create the information on the mapping or image tiles. The map is not just pictures that can readily be integrated into a GIS platform.

```
<?xml version="1.0" encoding="UTF-8"?>
<osm version="0.6" generator="CGImap 0.6.1 (22348 thorn-04.openstreetmap.org)" copyright="OpenStreetMap and contributors" attribution="http://www.openstreetmap.org/copyright" license="http://opendatacommon
 <bounds minlat="38.9500000" minlon="-77.0045600" maxlat="38.9533100" maxlon="-76.9982700"/>
 <node id="49729430" visible="true" version="2" changeset="3046923" timestamp="2009-11-06T10:23:28Z" user="woodpeck_fixbot" uid="147510" lat="38.9527230" lon="-77.0002160"/>
 <node id="737145404" visible="true" version="3" changeset="33125704" timestamp="2015-08-05T11:51:20Z" user="RoadGeek_MD99" uid="475877" lat="38.9503232" lon="-76.9988217">
  <tag k="addr:housenumber" v="500"/>
  <tag k="addr:street" v="Emerson Street Northeast"/>
  <tag k="dataset" v="ABRA locations"/>
  <tag k="name" v="Winebow"/>
  <tag k="shop" v="alcohol"/>
  <tag k="source" v="dcgis"/>
 </node>
 <node id="738189367" visible="true" version="11" changeset="55845086" timestamp="2018-01-29T00:21:45Z" user="cantor34" uid="7496544" lat="38.9518112" lon="-77.0021619">
  <tag k="addr:city" v="Washington"/>
  <tag k="addr:housenumber" v="550"/>
  <tag k="addr:postcode" v="20011"/>
  <tag k="addr:state" v="DC"/>
  <tag k="addr:street" v="Galloway Street Northeast"/>
  <tag k="dc-gis:addr" v="550 GALLOWAY STREET NE, WASHINGTON, DC"/>
  <tag k="dc-gis:dataset" v="MetroStnFullPt"/>
  <tag k="dc-gis:gis_id" v="mstn_017"/>
  <tag k="dc-gis:pubdate" v="2007-04-05"/>
  <tag k="name" v="Fort Totten"/>
  <tag k="network" v="Washington Metro"/>
  <tag k="operator" v="Washington Metropolitan Area Transit Authority"/>
  <tag k="platforms" v="2"/>
  <tag k="public_transport" v="station"/>
  <tag k="railway" v="station"/>
  <tag k="station" v="subway"/>
  <tag k="subway" v="yes"/>
  <tag k="website" v="http://www.wmata.com/rail/station_detail.cfm?station_id=28"/>
  <tag k="wikidata" v="Q5472194"/>
  <tag k="wikipedia" v="en:Fort Totten station"/>
 </node>
 <way id="6052373" visible="true" version="11" changeset="59114581" timestamp="2018-05-20T03:12:10Z" user="njtbusfan" uid="8107451">
  <nd ref="49729414"/>
  <nd ref="49729416"/>
  <nd ref="49729418"/>
  <nd ref="49729420"/>
  <nd ref="49729422"/>
  <nd ref="49729423"/>
  <nd ref="49729425"/>
  <nd ref="49729430"/>
  <nd ref="49729433"/>
  <nd ref="49729434"/>
  <nd ref="49729435"/>
  <tag k="HFCS" v="Collector"/>
  <tag k="highway" v="residential"/>
  <tag k="name" v="Galloway Street Northeast"/>
  <tag k="source:HFCS" v="District of Columbia (DC GIS)"/>
  <tag k="tiger:cfcc" v="A41"/>
  <tag k="tiger:county" v="District of Columbia, DC"/>
  <tag k="tiger:reviewed" v="yes"/>
 </way>
```




**Figure 24. Screenshot. Example of OpenStreetMap data in extensible markup language format.**
**Source: OpenStreetMap. Data is available under the Open Database License.**

## Pros and Cons

The greatest benefit of crowdsourced map data is in its update frequency. Crowdsourced maps do not have to rely on expensive and infrequent satellite sweeps or physical path traversing. Instead, map users can contribute changes and additions to existing maps to keep them updated in near realtime.

As with other crowdsourced data, there is some level of unreliability due to the possibility of users intentionally or unintentionally providing erroneous data. However, these map user communities are usually quick to "self-heal" through change moderation or independent verification by other users.

One provider often prompts its users to answer several questions about locations, businesses, and landmarks (e.g., "is there a wheelchair accessible ramp?") to validate and enhance existing data.

In cases of crowdsourced map data that also requires an open data license, the license may force agencies to share any derived data in the public domain. This might violate other licenses and agreements, and the data might not be wanted. For example, the OSM license requires users who derive data from OSM to publish that data back under Open Data Commons Open Database License (figure 25). This may be undesired if derived data is of sensitive or proprietary nature.

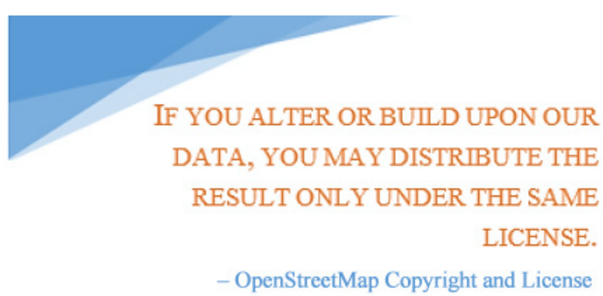



**Figure 25. Screenshot. License rules for OpenStreetMap.**
**Source: OpenStreetMap.**

The above licensing rule should not restrict agencies from using OSM data. Agencies can still create derivative products, such as additional map layers, without having to share those back. It is only when an agency directly edits the OSM files/database that those data need to be re-shared with the community. This subtle difference in understanding the licensing terms allows third parties to leverage OSM data to create their own derivative products and services that they then sell back to agencies and other companies.

For some specific applications in transportation, crowdsourced map data may be not be sufficiently accurate by itself. For example, agency asset-management divisions may need to know details about individual assets that the user community creating the map might not focus on. They could still leverage a crowdsourced map for understanding the road network, but they might have to create their own supplemental layers for certain assets like guardrails, signs, etc. if the user community had not already created those layers.

## Use Cases for Crowdsourced Map Data

**Use Case: Virginia DOT 511 System.** VDOT 511 traveler information system (figure 26) uses MapBox and OSM as their primary map data. VDOT overlays a number of data elements on top of OSM base data, including event locations and details, DMS, color-coded segments based on realtime probe vehicle speed, Waze-reported events, weather conditions, etc.

MapBox builds additional detail on top of OSM and then provides that product and associated services at a cost (figure 27).

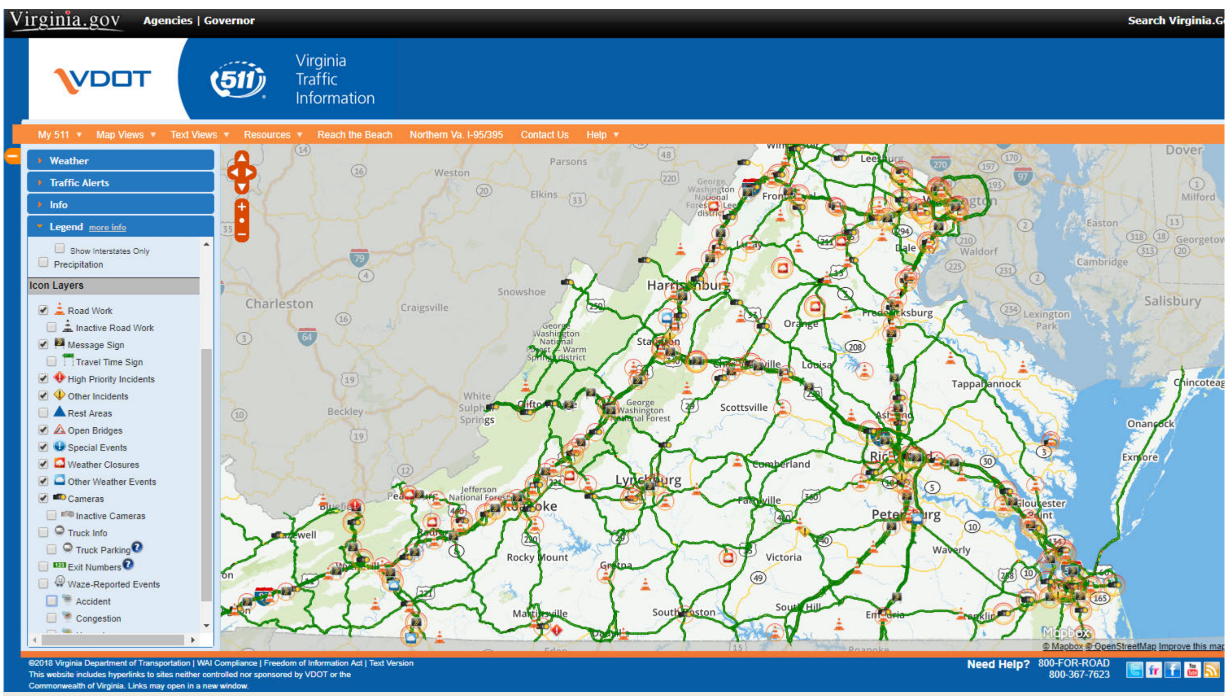





**Figure 26. Screenshot. The Virginia Department of Transportation 511 home page.**

**Source: Virginia Department of Transportation, 511 Website. Available at: http://www.511virginia.org/.**

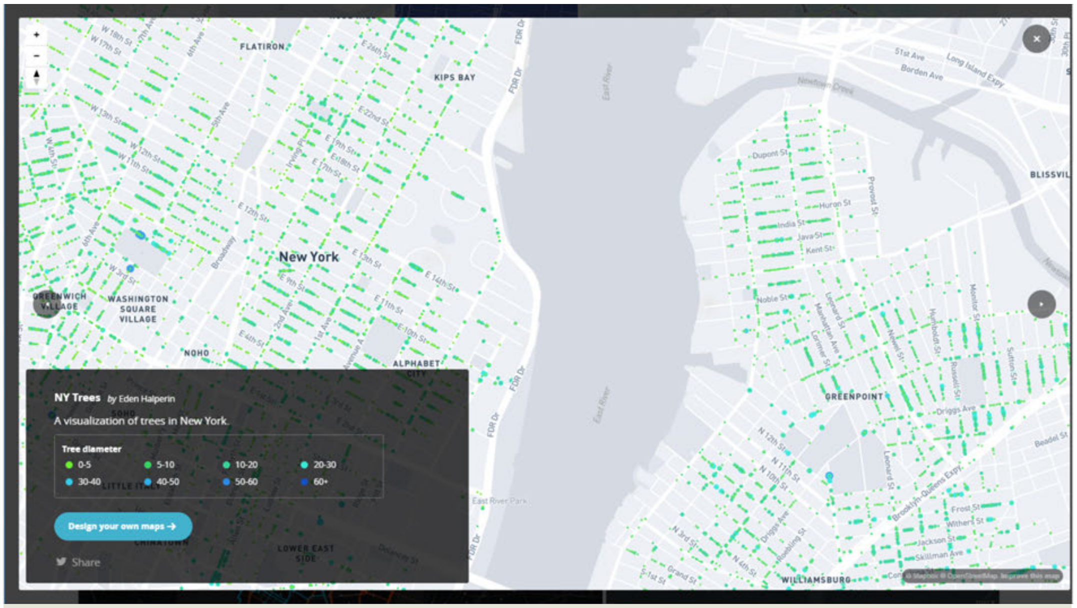





**Figure 27. Screenshot. Example of a MapBox map visualizing trees in New York City.**

**Source: Mapbox, Gallery, Website.**
**Available at: [https://www.mapbox.com/gallery/#map-0](https://www.mapbox.com/gallery/#map-0).**

**Use Case: Port Authority of New York and New Jersey Road Closure Application.** PANYNJ uses the Waze crowdsourced map as part of its online road-closure application to indicate which roads are closed. This helps to keep traffic from being routed onto those roadways. Waze collects crowdsourced map data in addition to collecting crowdsourced congestion and incident data as a basis for their maps (figure 28). Each Waze user can report map errors or contribute new information (e.g., new roads, new infrastructure such as red-light cameras, etc.) using the Waze Map Editor. This makes Waze map very accurate and timely when it comes to the roadway network. However, Waze map is not available for purchase by third parties at this time, and therefore it is of limited use to TMCs.

Waze allows users to use a map editor to suggest changes in existing map data, and if the user is a known map editor in their area, or other users verify the suggested change, that change becomes permanent in the new map. Waze has defined different user levels based on their level of engagement on the platform. Users logged into the Waze app collect points as they drive more miles and as they report events and have those events validated by other users. Classification of Waze users into different map editor levels depends on the amount of driving they do and the number of edits they submit. Initially, users can only make edits within 1 mile around routes they traversed. Trustworthy Waze users gain the ability to provide edits more broadly based on integration and validation of their map edits. In addition to regular Waze users, different geographic regions have dedicated Waze Map Editors verified by Waze as trusted contributors who can validate other editors' contributions and provide guidance.

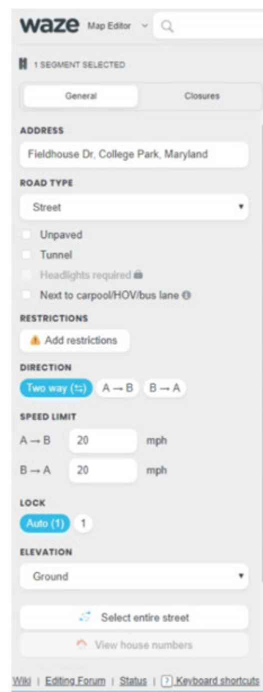





**Figure 28. Screenshot. Inputs panel on the Waze Map Editor.**

**Source: Waze.**

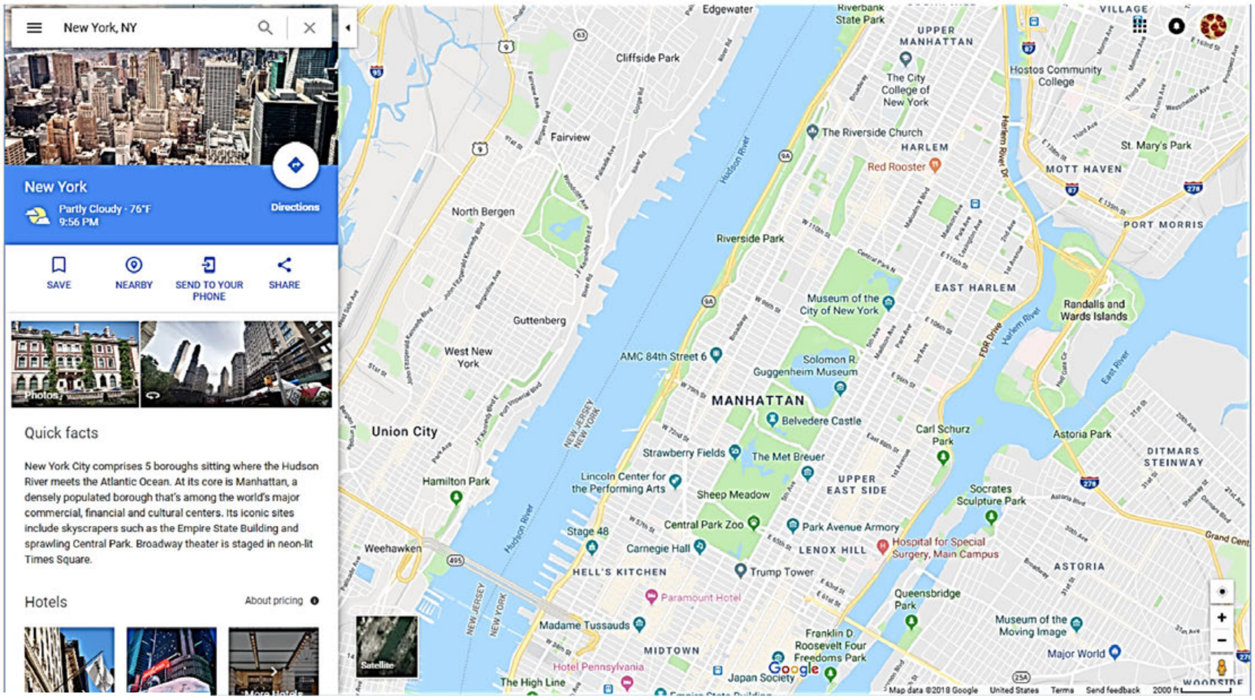
**Use Case: California's Bay Area 511 System.** 511 SF Bay is the multimodal traveler information system managed by a partnership of agencies led by the Metropolitan Transportation Commission, the California Highway Patrol (CHP), and the California Department of Transportation. The system provides traveler information for Bay Area travelers via the web or phone. 511 SF Bay uses Google Maps online to show a range of realtime and static traffic, carpooling, transit, parking, and bicycling information.

**General Transit Feed Specification (GTFS)**

One of the byproducts of Google Maps was the generation of Google Transit Feed Specification (later released as General Transit Feed Specification). GTFS was created as a way for Google to collect agency transit data (schedules, routes, fares, etc.) and display them on the Google Maps. Agencies saw this as a great opportunity to provide transit data to a wider customer base and built their systems to provide GTFS feeds and further enhance Google Maps, opening up new opportunities to provide value-add to their customers.

Google Maps (figure 29) is an example of a proprietary map that uses Google-generated map data in combination with user-contributed data. Google collects location information using their own mobile devices, devices running the Android operating system, and devices running Google location services-powered apps. In addition to this, Google has invested in a fleet of Google StreetView vehicles that traverse roadways (and, recently, trails) and generate images of the locations on the map as part of their StreetView service (figure 30). Users of Google Maps are frequently asked to contribute to map data by validating data collected by Google (such as verifying the location a user navigated to is correct, or verifying that hours of operation for a business are correct), as well as contributing supplemental data such as images of the business or landmark, reviews of services, etc. Due to the significant size of the Google Maps user base, it tends to have one of the more accurate and up-to-date map data sets. Still, updating brand new roads can take a little time—sometimes longer than a DOT would like. While agencies can purchase Google Maps for use in their TMCs, it can be expensive and may limit an agency in what maps can be used for and how they can be modified or enhanced.

The Google Maps platform provides three different products: maps, routes, and places. Each product consists of a number of capabilities available via APIs (figure 31). For example, routing products allows users to generate directions for different modes of transportation, distance matrices that provide travel times and distances for locations, and road traveled during a trip.



© 2018 Google

**Figure 29. Screenshot. Google Map.**
**Source: Google.**

© 2018 Google

**Figure 30. Screenshot. Google StreetView.**
**Source: Google.**

```
{
 "geocoded_waypoints": [{
   "geocoder_status": "OK",
   "place_id": "ChIJyYfhZ79ZwokRMtXcL6CYxkA",
   "types": ["premise"]
  }, {
   "geocoder_status": "OK",
   "partial_match": true,
   "place_id": "ChIJ8YWMWnz4wokRCOVf1CcJCbY",
   "types": ["street_address"]
  }
 ],
 "routes": [{
   "bounds": {
    "northeast": {
     "lat": 40.8171321,
     "lng": -73.99449150000001
    },
    "southwest": {
     "lat": 40.7416627,
     "lng": -74.0728354
    }
   },
   "copyrights": "Map data ©2015 Google",
   "legs": [{
     "distance": {
      "text": "9.7 mi",
      "value": 15653
     },
     "duration": {
      "text": "25 mins",
      "value": 1480
     },
     "end_address": "1 MetLife Stadium Dr, East Rutherford, NJ 07073, USA",
     "end_location": {
      "lat": 40.814505,
      "lng": -74.07272910000002
     },
     "start_address": "75 Ninth Ave, New York, NY 10011, USA",
     "start_location": {
      "lat": 40.7428759,
      "lng": -74.00584719999999
     },
     "steps": [{
       "distance": {
        "text": "440 ft",
        "value": 134
       },
       "duration": {
        "text": "1 min",
        "value": 34
       },
       "end_location": {
        "lat": 40.7422925,
        "lng": -74.004457
       },
       "html_instructions": "Head \u003cb\u003esoutheast\u003c/b\u003e on \u003cb\u003eW 16th St\u003c/b\u003e toward \u003cb\u003eNi
       "polyline": {
        "points": "_rtwFpgubMtBuG"
       },
       "start_location": {
        "lat": 40.7428759,
        "lng": -74.00584719999999
       },
       "travel_mode": "DRIVING"
      }, {
       "distance": {
        "text": "49 ft",
        "value": 15
       },
       "duration": {
        "text": "1 min",
        "value": 29
       },
       "end_location": {
        "lat": 40.7421744,
        "lng": -74.0045361
       },
       "html_instructions": "Turn \u003cb\u003eright\u003c/b\u003e at the 1st cross street onto \u003cb\u003eNinth Ave\u003c/b\u003e",
       "maneuver": "turn-right",
       "polyline": {
        "points": "intwFz~tbMVN"
       },
```




**Figure 31. Screenshot. Example routing result code from Google Maps.**

**Source: Google.**

# PROBE-BASED SPEED DATA

## Description

Over the last several decades, agencies often used floating car data to supplement speed data collected by the traditional static sensors. Use of floating cars was intermittent, tedious, and time consuming, which proved to be only marginally useful. Over the last 10 years, the private sector expanded on this concept by using technology built into vehicles and smartphones to transform a large percentage of travelers into probe vehicles. As the number of probes rose, the accuracy and reliability of probe speed data increased to the point where day-to-day transportation operations and planning efforts realized significant benefits.

Probe vehicle data comes from vehicles and people equipped with embedded GPS devices (in their vehicles or smartphones) and provides speed and travel time information. Planning and operations use this information now that it can be aggregated and anonymized.

Another way to generate probe vehicle data is using toll tags, as the Florida DOT (FDOT) does (figure 32). Toll operators use toll tags to identify and re-identify vehicles as they traverse the toll facility and use that information to calculate speed and travel times between re-identification points.

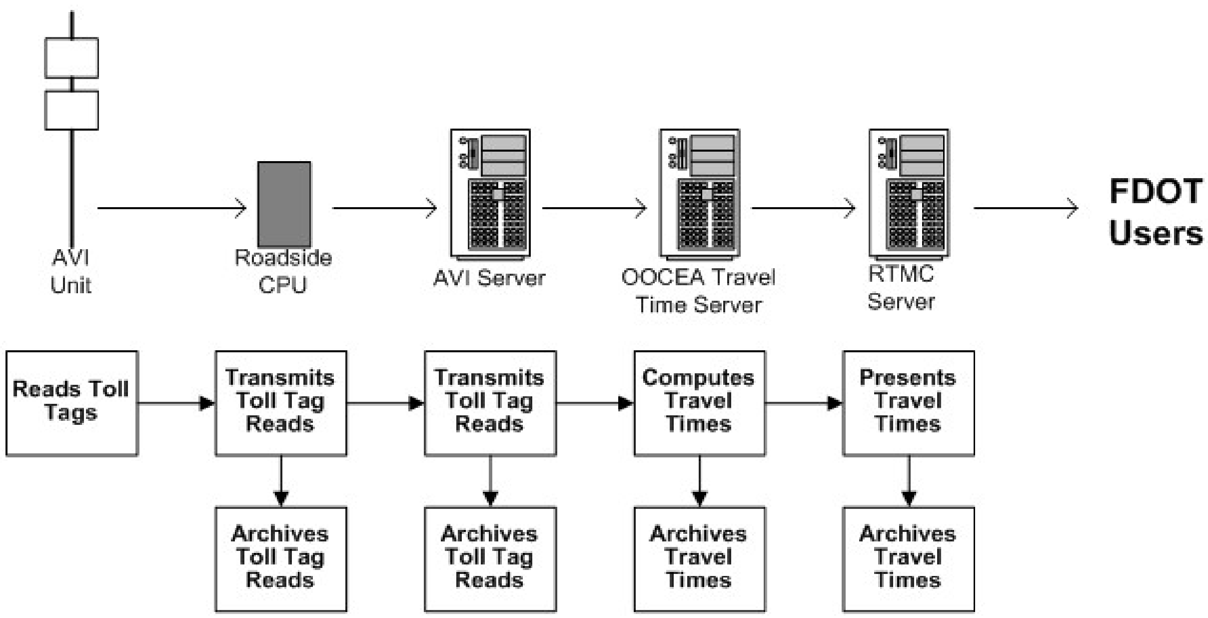



**Figure 32. Diagram. Toll tag travel time calculation.**

**Source: FHWA, iFlorida Model Deployment Final Evaluation Report, 2009. Available at: https://ops.fhwa.dot.gov/publications/fhwahop08050/chap_4.htm.**

## Applications of Probe-Based Speed Data

Probe vehicle data has many applications, including, but not limited to:

- Monitoring realtime congestion.
  - Detecting and identifying incidents.
  - Issuing traveler information.
  - Conducting work zone monitoring and impact analysis activities.
  - Detecting the end of the queue.
  - Comparing realtime speed information to historical trends.
  - Identifying recurring and non-recurring bottlenecks.
- Performance management.
  - Evaluating performance metrics over time: travel time, buffer time, reliability, planning time, and associated indices.
  - Incorporating data into dynamic performance management dashboards.
  - Investigating user delay cost.
  - Meeting Federal performance reporting requirements, including Moving Ahead for Progress in the 21st Century (MAP-21) third performance measure rule (PM3) reporting.
  - Evaluating worst bottlenecks in a region for a period of time.
  - Studying trends, including special event, holiday, and seasonal movements.
  - Exploring the impacts of capital investments – prior to, during, and after completion of the project.
- Planning and Research.
  - Identifying problems.
  - Prioritizing projects.
  - Performing safety analyses.
  - Implementing public participation/information campaigns.
  - Conducting before and after studies.
- Traveler information.
  - Providing realtime travel time information on DMS.
  - Delivering network performance information.
  - Distributing special event and holiday guidance.

## Attributes

Probe vehicle data is different from traditional ITS sensor data because it is link-based (probe) rather than point-based (sensor). This means that calculation of speed and travel time occurs over some distance on the roadway. Each speed/travel time record is associated with a timestamp and geographic link identifier as well as some form of confidence score that provides insight into the reliability of each data record. In the early days of probe vehicle data, providers exclusively used traffic message channel codes to identify geographic links associated with speed and travel time data. However, in recent years, data providers have increased granularity and coverage of their measurements to be able to provide data in sub-segments and outside of the traffic message channel network. Those smaller links vary across providers as they use proprietary technology and aggregation methods in data generation. Links can be as short as a couple hundred feet.

- **Latency:** Probe vehicle data generally has a low latency over a frequency ranging from 10 to 30 seconds. Providers aggregate this data and make it available in feeds, usually within 1 to 2 minutes after collection in the field. This latency is sufficient for most operations and planning purposes.
- **Details:** Probe vehicle data is consistent when it comes to the level of available details. Most providers collect speed, travel time, and quality data per segment of roadway.
- **Quality and Coverage:** Quality and coverage of probe vehicle data has continuously improved over the last decade that it has been available. Providers are constantly adding probes and improving collection and aggregation techniques to reduce latency and increase accuracy and coverage. Because providers collect data using different sources, some differences in data quality in different regions and on different road classes do exist. Some providers may have better quality data on arterials and in urban areas, but may have deficiencies in rural areas. Overall, providers are comparable in terms of quality and coverage, with slight differences being relevant in specialized applications.

To get the best available data and largest possible coverage, agencies sometimes purchase data that may be complementary to each other from multiple providers. Many agencies have developed data use agreements that include quality expectation clauses that ensure that providers adhere to minimum quality requirements and maintain a competitive market. To achieve this, agencies often employ independent validators to analyze data and compare it to ground truth data to ensure it complies with the minimum required accuracy and quality standards.

## Data Availability

Private sector probe vehicle speed data providers work with their partners and service users to collect billions of GPS data points aggregate them into speed and travel time data records based on underlying map segmentation. Transportation Network Companies (TNCs) collect similar data, but are not currently selling this information to public agencies for operations purposes. This data is limited to areas where TNCs operate and have significant penetration, which is limited to large metropolitan areas.

Additionally, many toll facility operators collect probe speed data using toll tag readers. For example, the Florida Turnpike Enterprise collects toll tag data and conflates it to their custom link-based map used in conjunction with their sensor data to monitor congestion and detect any issues related to traffic flow on the toll road.

The Florida Turnpike Enterprise uses toll tag reads from different gantries to determine average link speed on their toll facility. It uses its own roadway segmentation that does not conform to any specific standard, but works well with their internal operations map. Then they calculate speed and travel time across those links based on toll tag identification. The agency collects toll tag data and conflates it to their custom link-based map used in conjunction with their sensor data to monitor congestion and detect any issues related to traffic flow on the toll road.

## Pros and Cons

The primary benefit of probe vehicle data is ubiquitous coverage of the roadway network. In comparison to traditional ITS sensors deployed on a limited subset of the network, probe vehicle data can be collected anywhere there are equipped vehicles or devices. The current state of technology shows that probe vehicle data covers a large and growing percentage of all roadway networks. Figure 33 shows an example from Georgia, where the triangles represent the location of sensors deployed by the Georgia DOT in Atlanta. Roads with color on them are receiving probe-based speed measurements from a private sector probe-based speed data provider.

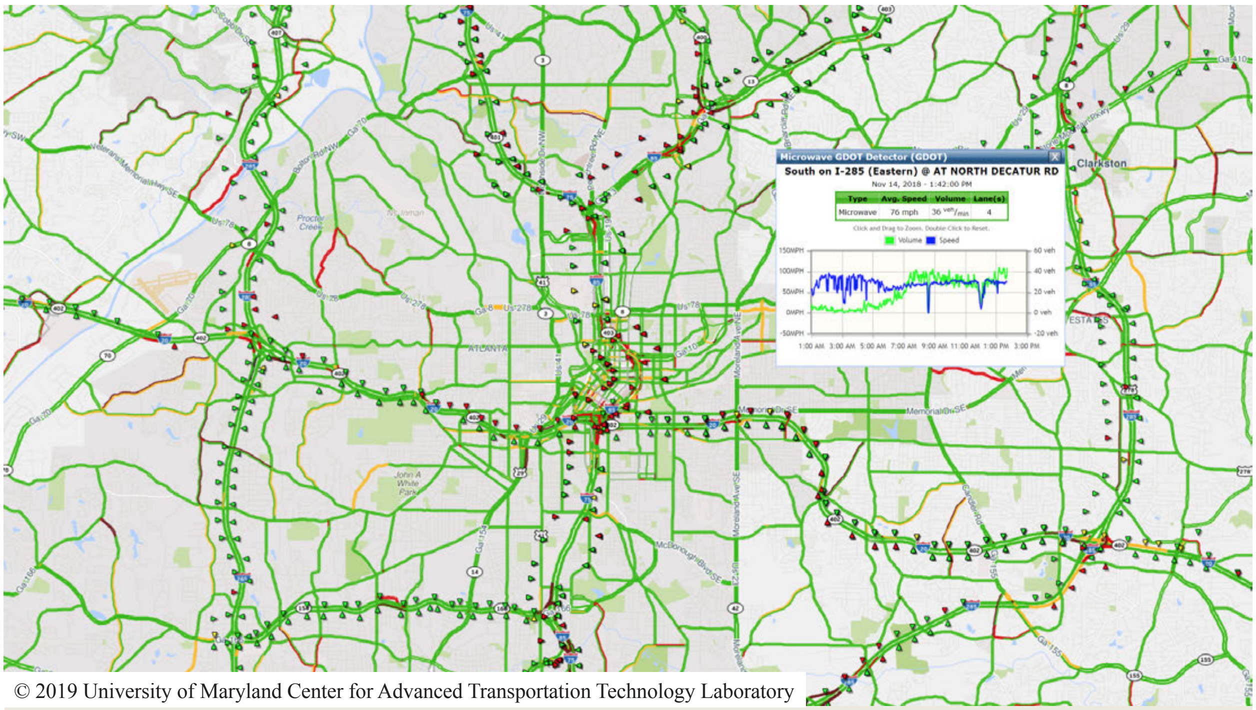





**Figure 33. Screenshot. Probe data provides ubiquitous coverage.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

Probe vehicle data has demonstrated a high level of accuracy and quality on highways, especially in metropolitan areas. However, the quality and accuracy of probe vehicle data on arterial roadways are not as high due to the nature of interrupted flow. However, providers have been working on improving data on the arterial roadway network and have shown some improvements over the last several years.

One of the major challenges associated with probe vehicle data is that it does not have associated volume information. So while speed and travel time information may be accurate, it is not clear how many vehicles may be experiencing that speed and travel time. This information is still mainly generated by the static ITS sensors. In recent months, there has been emerging research providing methods to approximate volumes for probe vehicle data. While not 100 percent accurate, these methods will continue to improve and be a viable option for operations and planning use.

## Use Cases for Probe-Based Speed Data

**Use Case: Capital Investment and Project Selection.** Agencies have a responsibility to identify best uses of limited funds and resources to improve safety and mobility. Previous project selection and capital investment decisions required significant analysis and research influenced by political pressure or vocal groups. With the availability of probe vehicle data, agencies have an opportunity to identify necessary projects based on insights from data.

For example, agencies can use probe vehicle data to generate a list of the most congested spots in a region (municipality, county, State, or multi-State region). Probe vehicle data helps to identify bottlenecks, sets of consecutive roadway segments where speeds drop below a certain threshold and remains below that threshold for some period. Properties of each bottleneck include its average length, duration, intensity, and frequency and pattern of occurrence. Significant and repeating bottlenecks can be good indicators of issues that can be resolved through implementation of improved operational strategies or capital investments. Figure 34 shows bottleneck ranking using INRIX data.

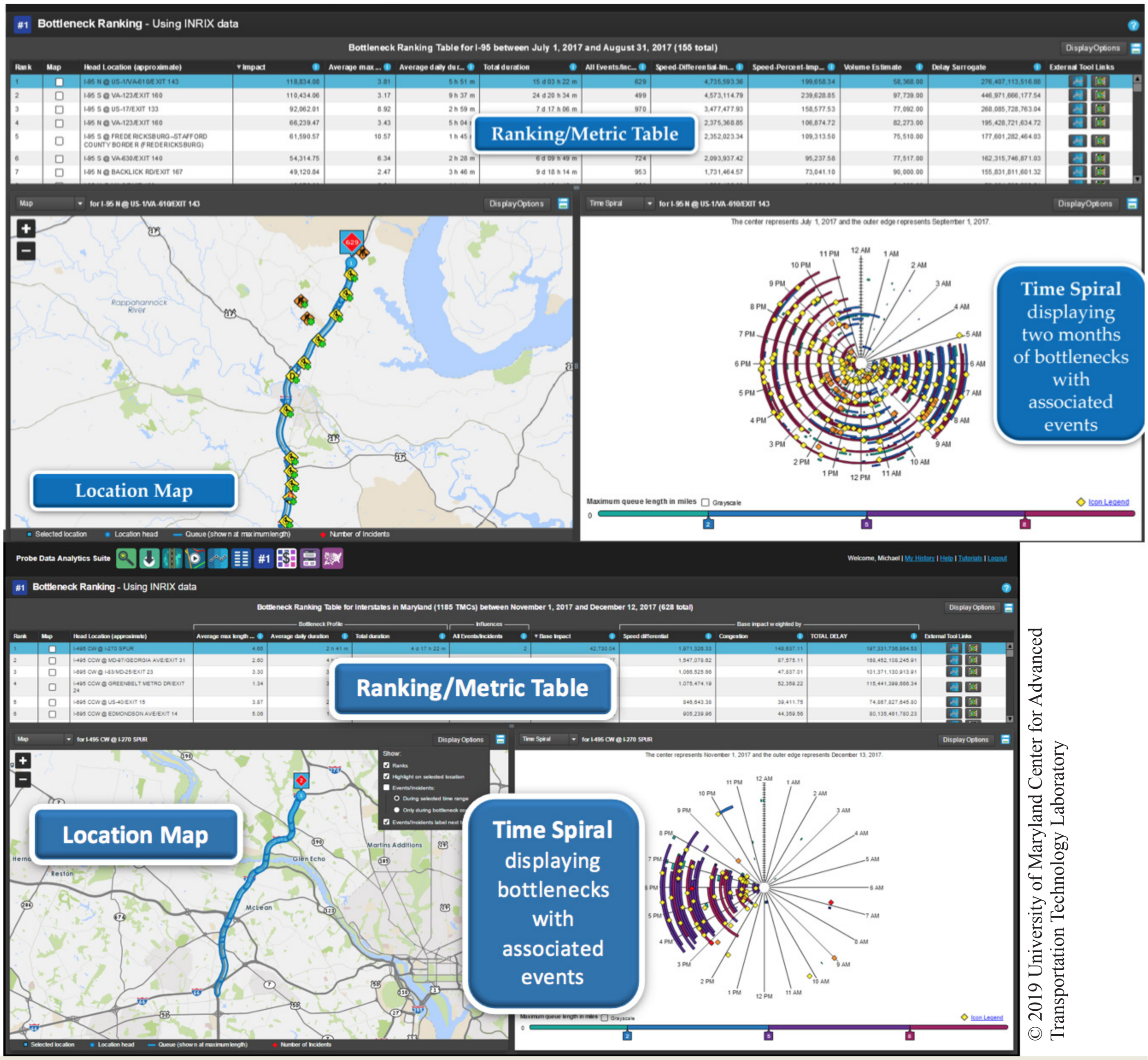



**Figure 34. Screenshot. Bottleneck ranking.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

Once a particular project location is identified, probe vehicle data can be used to analyze conditions prior to, during, and after implementation of the solution to determine the actual impact of the investment (figure 35). Animating speed data over time shows changes in congestion patterns. Probe vehicle speed data can be combined with other data sets, such as traffic volumes, traffic type classifications (commercial or passenger), and estimated value of travelers' time, to determine user delay costs (figure 36) and cost changes resulting from project implementation. Similar analysis can evaluate major event impact.

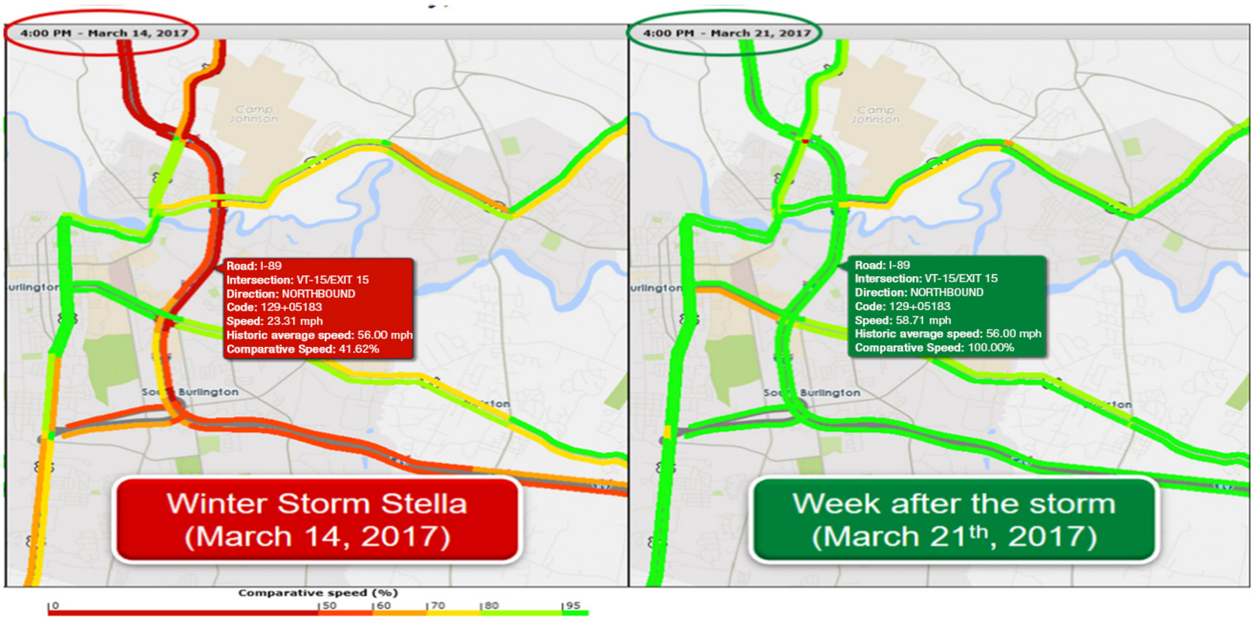





**Figure 35. Screenshot. Trend map shows a comparison of performance before, during, and after a major event.**

**Source: I-95 Corridor Coalition, Real-Time Traffic Incident Management Information System User Group Presentation. Available at: www.i95coalition.org.**

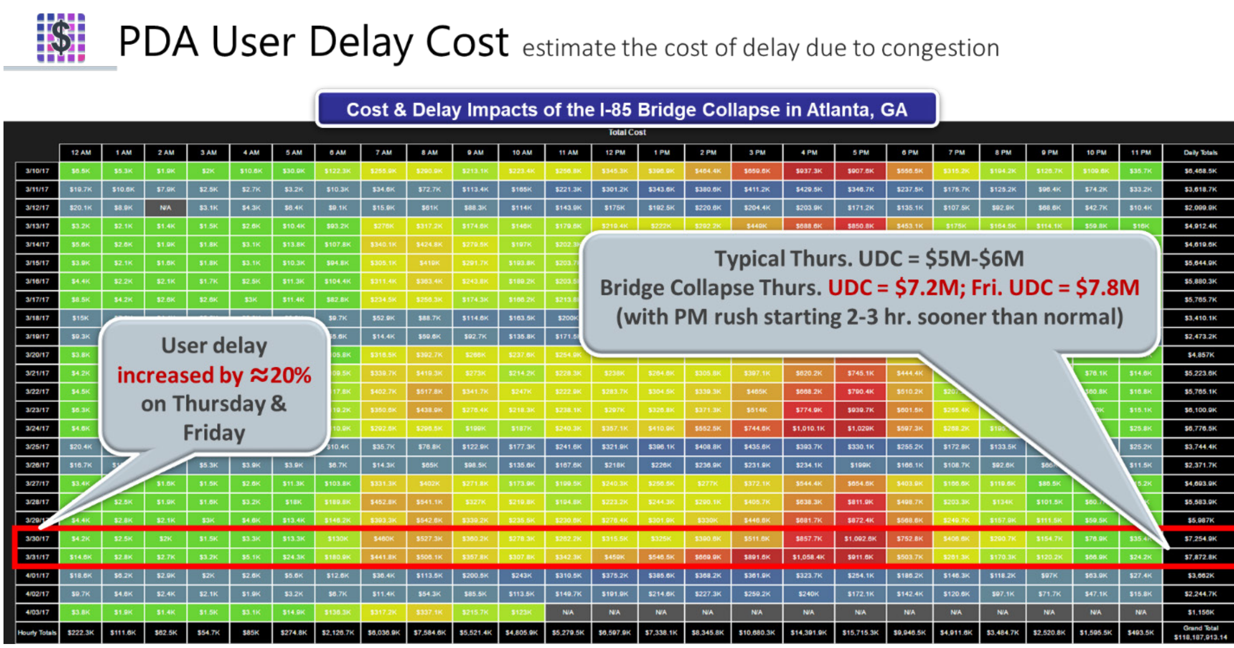





**Figure 36. Screenshot. User delay cost resulting from a major event.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

**Use Case: Performance Management and Reporting.** Probe vehicle data can help agencies evaluate mobility performance of the transportation system at varying geographic levels, from hyperlocal to regional. Probe vehicle speed and travel time data can provide the ability to calculate valuable metrics such as travel time index, buffer time index, and planning time index, which provide insight into system performance in comparison to average conditions, historical conditions, or ideal free-flow conditions. These metrics can help an agency evaluate its operations and identify areas for improvement.

Realtime dynamic performance dashboards use these metrics (figure 37) to identify issues in realtime as well as for inputs into daily, weekly, monthly, quarterly, and annual reports to the legislature, decision makers, and the public. Figure 38 shows an example performance report.

**PA Turnpike Closure (includes NJ)** + Add widget | Select a dashboard...

PA Turnpike Closure Travel Times into NJ (includes NJ)

| Corridor | Differential | Current | Historic | Differential | Current | Historic |
|---|---|---|---|---|---|---|
| US-1 Northbound between I-276/Pennsylvania Tpke and I-95 | ↓8 | 50 mph | 58 mph | ↑1 | 07 min | 06 min |
| I-95 Northbound between US-1/Exit 46 and PA--NJ State Border | ↓1 | 63 mph | 64 mph | 0 | 05 min | 05 min |
| I-95 between PA--NJ State Border and US-1/Exit 67 Northbound | 0 | 63 mph | 63 mph | ↑1 | 10 min | 09 min |
| I-295 between US-1/Exit 67 and Exit 60 Southbound | ↑1 | 66 mph | 65 mph | 0 | 07 min | 07 min |
| I-195 between I-295/Exit 60 and Exit 7 Eastbound | ↓18 | 39 mph | 57 mph | ↑3 | 09 min | 06 min |
| NJ-29 between I-95 and US-1 (TRENTON) (SOUTH) Southbound | ↓6 | 39 mph | 45 mph | ↑1 | 08 min | 07 min |
| NJ-29 between US-1 (TRENTON) (SOUTH) and I-195/I-295/Exit 60 Southbound | ↓13 | 30 mph | 43 mph | ↑3 | 09 min | 06 min |
| US-1 Northbound between I-95 and PA--NJ State Border | ↑5 | 65 mph | 60 mph | 0 | 06 min | 06 min |
| US-1 between PA--NJ State Border and I-295/I-95 Northbound | ↓3 | 46 mph | 49 mph | 0 | 08 min | 08 min |
| 10S - 413 (PA/NJ) - I-95 to US 130 | ↓1 | 23 mph | 24 mph | ↑1 | 09 min | 08 min |
| I-295 between Exit 60 and US-130/Exit 57 Southbound | ↑2 | 67 mph | 65 mph | 0 | 02 min | 02 min |

Using INRIX data — Updated Feb 2, 2017 4:33 PM (21s ago)

PA Turnpike Closure Travel Times into PA (includes NJ)

| Corridor | Differential | Current | Historic | Differential | Current | Historic |
|---|---|---|---|---|---|---|
| US-1 Southbound between I-95 and I-276/Pennsylvania Tpke | ↓43 | 13 mph | 56 mph | ↑23 | 29 min | 06 min |
| I-95 Southbound between PA--NJ State Border and US-1/Exit 46 | 0 | 60 mph | 60 mph | 0 | 05 min | 05 min |
| I-95 between PA--NJ State Border and US-1/Exit 67 Southbound | ↑2 | 59 mph | 57 mph | ↓1 | 10 min | 11 min |
| I-295 between US-1/Exit 67 and Exit 60 Northbound | 0 | 66 mph | 66 mph | 0 | 07 min | 07 min |
| I-195 between I-295/Exit 60 and Exit 7 Westbound | ↓18 | 44 mph | 62 mph | ↑3 | 10 min | 07 min |
| NJ-29 between US-1 (TRENTON) (SOUTH) and I-95 Northbound | ↓2 | 42 mph | 44 mph | ↑1 | 08 min | 07 min |
| NJ-29 between US-1 (TRENTON) (SOUTH) and I-195/I-295/Exit 60 Northbound | ↓3 | 52 mph | 55 mph | 0 | 04 min | 04 min |
| US-1 Southbound between PA--NJ State Border and I-95 | ↓3 | 51 mph | 54 mph | 0 | 07 min | 07 min |
| US-1 between PA--NJ State Border and I-295/I-95 Southbound | ↑4 | 52 mph | 48 mph | ↓1 | 07 min | 08 min |
| 10N - 413 (PA/NJ) - US 130 to I-95 (PA) | ↓5 | 18 mph | 23 mph | ↑2 | 10 min | 08 min |
| I-295 between US-130/Exit 57 and Exit 60 Northbound | ↓1 | 65 mph | 66 mph | 0 | 02 min | 02 min |

Using INRIX data — Updated Feb 2, 2017 4:33 PM (22s ago)

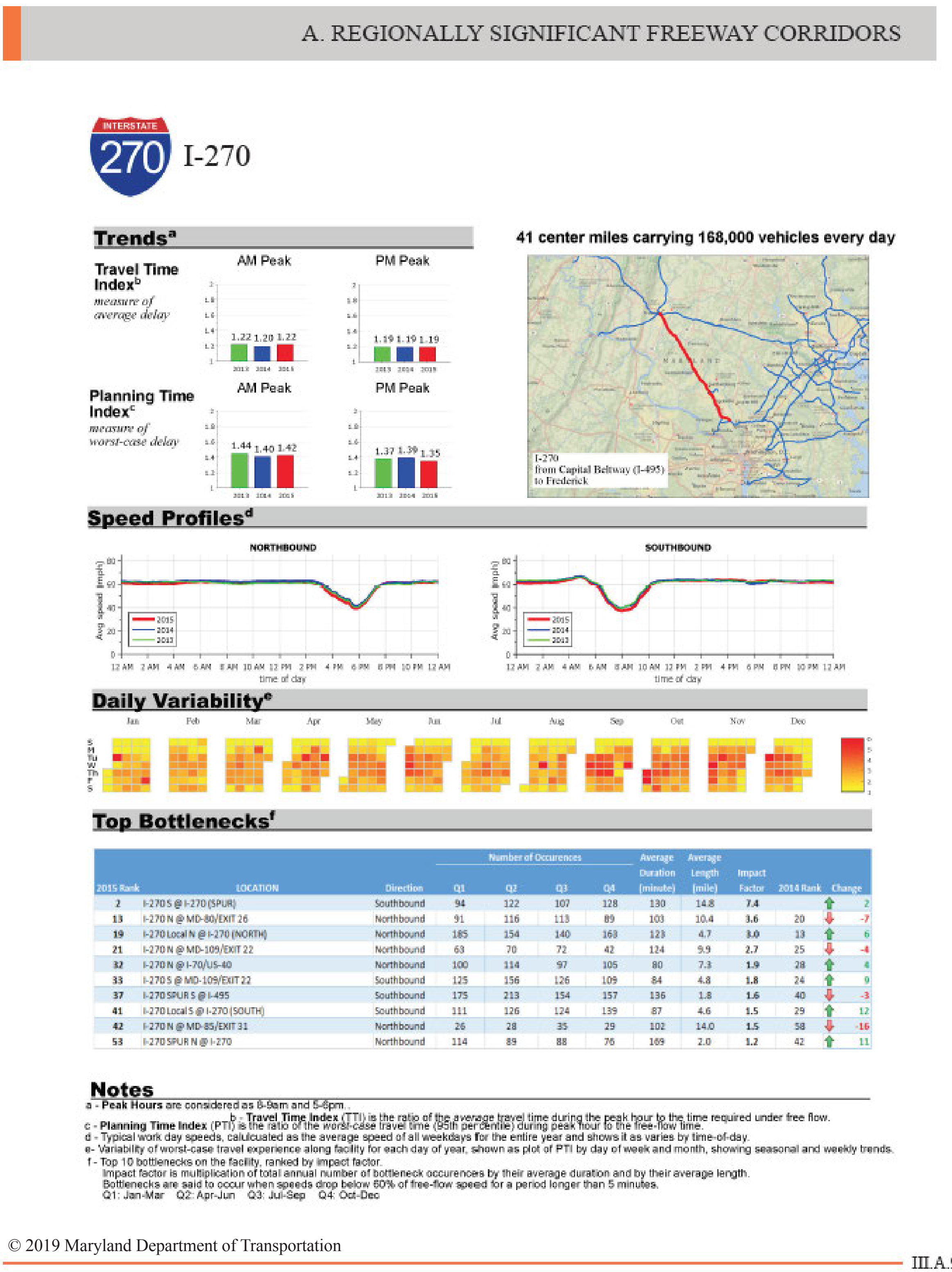
Travel Time US 1/I-95 (PA) to Tpk Exit 7 (NJ)

| Corridor | Differential | Current | Historic | Differential | Current | Historic |
|---|---|---|---|---|---|---|
| US 1 (PA) to Exit 7 (NJ) via 95 | ↓9 | 53 mph | 62 mph | ↑5 | 35 min | 30 min |
| US 1 (PA) to Exit 7 (NJ) via US 1/NJ 29 | ↓17 | 37 mph | 54 mph | ↑11 | 32 min | 21 min |



**Figure 37. Screenshot. Example of a performance dashboard travel time widget.**

**Source: Pennsylvania Turnpike Commission.**



A. REGIONALLY SIGNIFICANT FREEWAY CORRIDORS

## I-270

### Trends[a]

**Travel Time Index[b]**
*measure of average delay*

AM Peak: 2013: 1.22, 2014: 1.20, 2015: 1.22
PM Peak: 2013: 1.19, 2014: 1.19, 2015: 1.19

**Planning Time Index[c]**
*measure of worst-case delay*

AM Peak: 2013: 1.44, 2014: 1.40, 2015: 1.42
PM Peak: 2013: 1.37, 2014: 1.39, 2015: 1.35

**41 center miles carrying 168,000 vehicles every day**

I-270
from Capital Beltway (I-495)
to Frederick

### Speed Profiles[d]

NORTHBOUND
SOUTHBOUND
Avg speed (mph)
2015
2014
2013
12 AM 2 AM 4 AM 6 AM 8 AM 10 AM 12 PM 2 PM 4 PM 6 PM 8 PM 10 PM 12 AM
time of day

### Daily Variability[e]

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
S M Tu W Th F S

### Top Bottlenecks[f]

| 2015 Rank | LOCATION | Direction | Number of Occurences Q1 | Q2 | Q3 | Q4 | Average Duration (minute) | Average Length (mile) | Impact Factor | 2014 Rank | Change |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | I-270 S @ I-270 (SPUR) | Southbound | 94 | 122 | 107 | 128 | 130 | 14.8 | 7.4 | | 2 |
| 13 | I-270 N @ MD-80/EXIT 26 | Northbound | 91 | 116 | 113 | 89 | 103 | 10.4 | 3.6 | 20 | -7 |
| 19 | I-270 Local N @ I-270 (NORTH) | Northbound | 185 | 154 | 140 | 163 | 123 | 4.7 | 3.0 | 13 | 6 |
| 21 | I-270 N @ MD-109/EXIT 22 | Northbound | 63 | 70 | 72 | 42 | 124 | 9.9 | 2.7 | 25 | -4 |
| 32 | I-270 N @ I-70/US-40 | Northbound | 100 | 114 | 97 | 105 | 80 | 7.3 | 1.9 | 28 | 4 |
| 33 | I-270 S @ MD-109/EXIT 22 | Southbound | 125 | 156 | 126 | 109 | 84 | 4.8 | 1.8 | 24 | 9 |
| 37 | I-270 SPUR S @ I-495 | Southbound | 175 | 213 | 154 | 157 | 136 | 1.8 | 1.6 | 40 | -3 |
| 41 | I-270 Local S @ I-270 (SOUTH) | Southbound | 111 | 126 | 124 | 139 | 87 | 4.6 | 1.5 | 29 | 12 |
| 42 | I-270 N @ MD-85/EXIT 31 | Northbound | 26 | 28 | 35 | 29 | 102 | 14.0 | 1.5 | 58 | -16 |
| 53 | I-270 SPUR N @ I-270 | Northbound | 114 | 89 | 88 | 76 | 169 | 2.0 | 1.2 | 42 | 11 |

### Notes

a - **Peak Hours** are considered as 8-9am and 5-6pm..
b - **Travel Time Index** (TTI) is the ratio of the average travel time during the peak hour to the time required under free flow.
c - **Planning Time Index** (PTI) is the ratio of the worst-case travel time (95th percentile) during peak hour to the free-flow time.
d - Typical work day speeds, calulucated as the average speed of all weekdays for the entire year and shows it as varies by time-of-day.
e- Variability of worst-case travel experience along facility for each day of year, shown as plot of PTI by day of week and month, showing seasonal and weekly trends.
f - Top 10 bottlenecks on the facility, ranked by impact factor.
Impact factor is multiplication of total annual number of bottleneck occurences by their average duration and by their average length.
Bottlenecks are said to occur when speeds drop below 60% of free-flow speed for a period longer than 5 minutes.
Q1: Jan-Mar Q2: Apr-Jun Q3: Jul-Sep Q4: Oct-Dec

© 2019 Maryland Department of Transportation

III.A.9

**Figure 38. Screenshot. Example performance report.**

**Source: Maryland Department of Transportation, 2015 Maryland Mobility Report. Available at: https://www.roads.maryland.gov/OPPEN/2015%20mobility%20report%20draft_highres_for%20website1.pdf.**

**Use Case: Traveler Information and Holiday Traffic Forecasting.** Archived probe vehicle data provides insight into travel patterns, especially as they relate to holiday traffic. Agencies can use these historical patterns to forecast impacts and inform the traveling public accordingly. For example, the Baltimore Metropolitan Council uses probe vehicle speed and travel time data to identify best days and times for travel during Thanksgiving week and publishes that information on their website to inform the traveling public (figure 39).

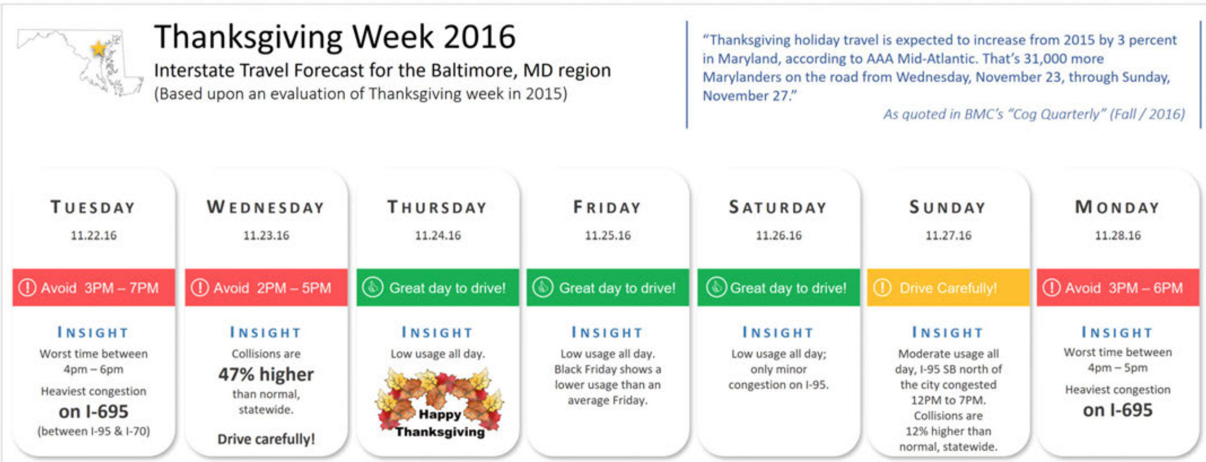





**Figure 39. Screenshot. Example of an Interstate travel forecast for the Baltimore, MD region during the week of Thanksgiving in 2016.**

**Source: Baltimore Metropolitan Council.**

**Use Case: Multiple Uses Across Planning and Operations.** State DOTs, MPOs, and cities are using probe-vehicle speed data for a range of planning and operations purposes including traveler information and travel time generation, before and after studies, project prioritization, and performance measurement.

Maryland DOT, Arizona DOT, and Pennsylvania DOT are examples of organizations that use probe speed data from INRIX for traveler information and travel time generation, before and after studies, project prioritization, and performance measurement.

INRIX primarily works with freight operators, fleet vehicles (such as taxicabs, United Parcel Service, FedEx, etc.), individual OEMs, and their custom app users to collect location data over time. INRIX aggregates individual probe readings per roadway segment to calculate segment-based speed and travel time in near realtime. In addition to speed-readings and travel time calculations, INRIX provides confidence scores that implicitly indicate the number of probes used to generate data and the level of modeling/imputing/archived data used to supplement low probe count. INRIX makes this segment-based speed and travel time data available to TMCs as a realtime data feed, archived raw data dump, or as part of the analytics package that includes both data and a set of data analysis tools.

INRIX provides speed, travel time, and confidence score values per segment of the road at a frequency as low as once per minute. They offer several different roadway segmentation patterns:

- Traffic message channel code.
  - Traffic message channel standard segmentation that splits road network into segments of varying length from tenth of a mile to several miles (figure 40 and figure 41).
- Extreme Definitions (XD).
  - INRIX proprietary segmentation scheme that provides broader and more granular coverage than traffic message channel codes (figure 42 and figure 43). There is no relation between XDs and traffic message channels because they represent different segmentation patterns.

| TMC | Type | RoadNumber | RoadName | FirstName | LinearTMC | Country | State | County | ZIP | Direction | StartLong | StartLat | EndLong | EndLat | Miles |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 125P05269 | P1 | I 95 | | HWY 903/EXIT 168 | 12500118 | USA | NC | HALIFAX | 27870 | NORTHBOUND | -77.67429 | 36.36146 | -77.67092 | 36.36661 | 0.40198 |
| 125+05270 | P1 | I 95 | | HWY 125/EXIT 171 | 12500118 | USA | NC | HALIFAX | 27870 | NORTHBOUND | -77.67092 | 36.36661 | -77.64708 | 36.39645 | 2.45490 |
| 125P05270 | P1 | I 95 | | HWY 125/EXIT 171 | 12500118 | USA | NC | HALIFAX | 27870 | NORTHBOUND | -77.64708 | 36.39645 | -77.64245 | 36.40361 | 0.58853 |
| 125+05271 | P1 | I 95 | | US 158/EXIT 173 | 12500118 | USA | NC | HALIFAX | 27890 | NORTHBOUND | -77.64245 | 36.40361 | -77.63114 | 36.42800 | 1.80312 |
| 125P05271 | P1 | I 95 | | US 158/EXIT 173 | 12500118 | USA | NC | HALIFAX | 27890 | NORTHBOUND | -77.63114 | 36.42800 | -77.62920 | 36.43304 | 0.37657 |
| 125+05272 | P3 | I 95 | | HALIFAX/NORTHAMPTON COUNTY LINE | 12500118 | USA | NC | HALIFAX | 27870 | NORTHBOUND | -77.62920 | 36.43304 | -77.62576 | 36.44408 | 0.86238 |
| 110+05548 | P1 | I 95 | | HWY 46/EXIT 176 | 11000199 | USA | NC | NORTHHAMPTON | 27831 | NORTHBOUND | -77.62470 | 36.44533 | -77.61262 | 36.46598 | 1.62204 |
| 110P05548 | P1 | I 95 | | HWY 46/EXIT 176 | 11000199 | USA | NC | NORTHHAMPTON | 27831 | NORTHBOUND | -77.61262 | 36.46598 | -77.61009 | 36.47082 | 0.36917 |
| 110N05548 | P1 | I 95 | | HWY 46/EXIT 176 | 11000199 | USA | NC | NORTHHAMPTON | 27831 | SOUTHBOUND | -77.60987 | 36.47130 | -77.61276 | 36.46574 | 0.42230 |
| 110-05547 | P3 | I 95 | | NORTHHAMPTON/HALIFAX COUNTY LINE | 11000199 | USA | NC | NORTHHAMPTON | 27831 | SOUTHBOUND | -77.61276 | 36.46574 | -77.62470 | 36.44533 | 1.60203 |
| 125-05271 | P1 | I 95 | | US 158/EXIT 173 | 12500118 | USA | NC | HALIFAX | 27890 | SOUTHBOUND | -77.62576 | 36.44408 | -77.62931 | 36.43268 | 0.89357 |
| 125N05271 | P1 | I 95 | | US 158/EXIT 173 | 12500118 | USA | NC | HALIFAX | 27890 | SOUTHBOUND | -77.62931 | 36.43268 | -77.63110 | 36.42810 | 0.33953 |
| 125-05270 | P1 | I 95 | | HWY 125/EXIT 171 | 12500118 | USA | NC | HALIFAX | 27870 | SOUTHBOUND | -77.63110 | 36.42810 | -77.64165 | 36.40534 | 1.68082 |
| 125N05270 | P1 | I 95 | | HWY 125/EXIT 171 | 12500118 | USA | NC | HALIFAX | 27870 | SOUTHBOUND | -77.64165 | 36.40534 | -77.64375 | 36.40081 | 0.33630 |
| 125-05269 | P1 | I 95 | | HWY 903/EXIT 168 | 12500118 | USA | NC | HALIFAX | 27870 | SOUTHBOUND | -77.64375 | 36.40081 | -77.67107 | 36.36641 | 2.83961 |
| 125N05269 | P1 | I 95 | | HWY 903/EXIT 168 | 12500118 | USA | NC | HALIFAX | 27870 | SOUTHBOUND | -77.67107 | 36.36641 | -77.67437 | 36.36132 | 0.39813 |

**Figure 40. Screenshot. Example of INRIX TMC-based metadata.**
**Source: INRIX.**

```
<TMC code=“114-11503” speed=”14” average=”12” reference=”18” delta=”2”
score=”30” travelTimeMinutes=”0.524” contestionLevel=”2” />
<TMC code=“114-08207” speed=”14” average=”14” reference=”17” delta=”0”
score=”20” travelTimeMinutes=”0.264” contestionLevel=”2” />
```

**Figure 41. Screenshot. Example speed and travel time data.**
**Source: INRIX.**

The confidence score indicates the quality of speed and travel time data. The score can be 10, 20, or 30.

- Ten usually indicates historical data or based on road reference speeds.
- Twenty represents medium confidence based on realtime data across multiple segments and/or based on a combination of expected and realtime data.
- Thirty indicates a high-confidence measurement that signifies realtime data for a specific segment.

Note that the data depicted in figures 42 and 43 is not relevant per se; rather, the purpose of these illustrations is to indicate typical data type and configuration.

| SegID | PreviousSegID | NextSegID | FRC | Miles | Lanes | StartLat | StartLong | EndLat | EndLong | RoadNumber | RoadName | LinearID | Country | State | County | PostalCode |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.524E+09 | 1524480539 | 1.524E+09 | 3 | 0.65037 | | 32.433 | -86.6529 | 32.431 | -86.663 | 14 | Highway 14 W | | United States | Alabama | Autauga | 36003 |
| 1.226E+09 | 1226243897 | 1.226E+09 | 2 | 0.56503 | | 35.747 | -109.62 | 35.742 | -109.61 | 191 | US Highway 191 | | United States | Arizona | Apache | 86505 |
| 1.643E+09 | 1642786923 | 1.643E+09 | 3 | 0.28672 | 3.3 | 33.968 | -118.292 | 33.964 | -118.29 | | S Vermont Ave | | United States | California | Los Angeles | 90044 |
| 1.187E+09 | 1187325927 | 1.187E+09 | 2 | 0.72893 | | 37.112 | -107.867 | 37.122 | -107.87 | 550 | Highway 550 | | United States | Colorado | La Plata | 81303 |
| 1.423E+09 | 1422521929 | 1.423E+09 | 3 | 0.46162 | 1.04 | 28.558 | -80.8227 | 28.551 | -80.823 | | Barna Ave | | United States | Florida | Brevard | 32780 |

**Figure 42. Screenshot. Extreme Definition metadata.**
**Source: INRIX.**


<Segment code="266912847" type="XDS" speed="55" average="63" reference="63" score="30" c-value="57" travelTimeMinutes="0.59" speedBucket="2" />
<Segment code="266912834" type="XDS" speed="21" average="63" reference="63" score="30" c-value="57" travelTimeMinutes="1.574" speedBucket="1" />
<Segment code="266912814" type="XDS" speed="30" average="63" reference="63" score="30" c-value="46" travelTimeMinutes="1.257" speedBucket="1" />
<Segment code="267054847" type="XDS" speed="56" average="63" reference="63" score="30" c-value="100" travelTimeMinutes="0.301" speedBucket="2" />
<Segment code="266939167" type="XDS" speed="61" average="63" reference="63" score="30" c-value="78" travelTimeMinutes="0.497" speedBucket="3" />
<Segment code="266939180" type="XDS" speed="59" average="63" reference="63" score="30" c-value="78" travelTimeMinutes="0.512" speedBucket="3" />
<Segment code="267078763" type="XDS" speed="59" average="62" reference="63" score="30" c-value="72" travelTimeMinutes="0.3" speedBucket="3" />
<Segment code="266939187" type="XDS" speed="62" average="63" reference="63" score="30" c-value="100" travelTimeMinutes="0.798" speedBucket="3" />
<Segment code="267048708" type="XDS" speed="60" average="61" reference="63" score="30" c-value="100" travelTimeMinutes="0.675" speedBucket="3" />
<Segment code="266912631" type="XDS" speed="65" average="63" reference="63" score="30" c-value="100" travelTimeMinutes="0.445" speedBucket="3" />
<Segment code="266983743" type="XDS" speed="61" average="62" reference="63" score="30" c-value="100" travelTimeMinutes="0.305" speedBucket="3" />


**Figure 43. Screenshot. Example speed and travel time data expressed in extreme definition format.**
**Source: INRIX.**

North Carolina DOT, FDOT, and Georgia DOT use probe speed data from HERE for these activities (figure 44).

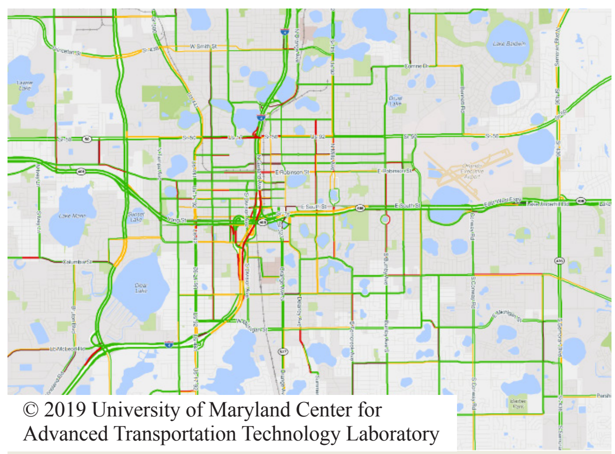



**Figure 44. Screenshot. HERE speed data in Orlando, FL.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

HERE works with smartphone manufacturers and cell service providers, as well as OEMs, to collect location data over time. HERE and INRIX data formats are very similar, with slight differences in how they define data quality measures. HERE also provides data access to TMCs via realtime feeds, archived raw data dumps, and analytics platforms that include both data and analysis tools.

Similar to INRIX, HERE provides speed, travel time, and confidence value per segment of the road at frequencies as low as once per minute (figure 45 and figure 46). HERE data use several segmentation patterns as well:

- Traffic message channel code – same standardized segmentation that INRIX uses.
- Sub-segments:
  - Proprietary HERE segmentation that is more granular than traffic message channel codes, but also dynamically defined. This means that sub-segment measurements are not necessarily available at every polling interval, but are values assigned only when meeting specific speed conditions.
  - Sub-segments can also include per lane data breakdown as well.

HERE probe data also comes with a quality measurement in the form of a confidence factor.

- A confidence factor of 0.0 to 0.5 (including 0.5) represents lower confidence for reference speed measurements.
- A confidence factor of 0.5 to 0.7 (including 0.7) represents medium confidence where measurements are a combination of historic and realtime data.
- A confidence factor of 0.7 to 1.0 (including 1.0) represents high confidence measurements based on realtime data for that specific segment.

In conjunction with the ICC, Maryland DOT is currently testing the use of data from the private


```
<TMC PC="4100" DE="xx" QD="+" LE="1.9"/>
    <CF TY="TR" SP="42.57" SU="42.57" FF="55.0" JF="0.0" CN="0.83">
        <LN NM=”1,2" SU="50.00" JF="0.0"/>
        <LN NM=”3" SU="10.00" JF="8.0"/>
    </CF>
</TMC>
```


**Figure 45. Screenshot. Example HERE traffic message channel per-lane data.**
**Source: HERE.**


```
<FI>
  <TMC PC="4102" DE="2nd St/Exit 1/Exit 2" QD="+" LE="0.66708"/>
  <CF TY="TR" SP="20.33" SU="20.33" FF="54.68" JF="7.60997" CN="0.99" TS="O">
    <SSS>
      <SS LE="0.21537" SP="20.98" SU="20.98" FF="54.68" JF="7.48939" TS="O">
        <LN NM="4" SU="34.41" JF="4.99449"/>
        <LN NM="1,2,3" SU="7.54" JF="9.29383"/>
      </SS>
      <SS LE="0.28697" SP="13.34" SU="13.34" FF="54.68" JF="8.59154" TS="O">
        <LN NM="4" SU="34.41" JF="4.99449"/>
        <LN NM="1,2,3" SU="7.54" JF="9.29383"/>
      </SS>
      <SS LE="0.13160" SP="24.70" SU="24.70" FF="54.68" JF="6.79808" TS="O">
        <LN NM="4" SU="34.41" JF="4.99449"/>
        <LN NM="1,2,3" SU="7.54" JF="9.29383"/>
      </SS>
      <SS LE="0.03313" SP="24.70" SU="24.70" FF="54.68" JF="6.79808" TS="O"/>
    </SSS>
  </CF>
</FI>
```


**Figure 46. Screenshot. Example HERE sub-segment per-lane data.**
**Source: HERE.**

sector vendor, TomTom, for a range of planning and operations functions. TomTom collects probe speed data from its navigation system users, which include both passenger and freight vehicles. Unlike INRIX and HERE, both of which are present globally in some way, TomTom has a strong presence outside of the U.S. and generally larger coverage in Europe. Similar to INRIX and HERE, TomTom collects segment-based speed and travel-time data and provides this data to TMCs as realtime data feeds.

Similar to INRIX and HERE, TomTom provides segment based speed and travel-time at frequencies as low as once per minute (figure 47). TomTom data uses two segmentation patterns as well:

- Traffic message channel code – same standardized segmentation that INRIX and HERE use.
- OpenLR – Open source dynamic location referencing system using binary format.


```xml
<elaboratedData id="L129p04148">
    <basicDataValue
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:type="TravelTimeValue">
        <supplierCalculatedDataQuality>95.0</supplierCalculatedDataQuality>
        <affectedLocation>
            <locationContainedInGroup xsi:type="LocationByReference">
                <predefinedLocationReference>L129p04148</predefinedLocationReference>
            </locationContainedInGroup>
        </affectedLocation>
        <freeFlowSpeed>109.0</freeFlowSpeed>
        <freeFlowTravelTime>73.22</freeFlowTravelTime>
    </basicDataValue>
</elaboratedData>
<elaboratedData id="L129p04149">
    <basicDataValue xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:type="TravelTimeValue">
        <supplierCalculatedDataQuality>95.0</supplierCalculatedDataQuality>
        <affectedLocation>
            <locationContainedInGroup xsi:type="LocationByReference">
                <predefinedLocationReference>L129p04149</predefinedLocationReference>
            </locationContainedInGroup>
        </affectedLocation>
        <freeFlowSpeed>111.0</freeFlowSpeed>
        <freeFlowTravelTime>152.24</freeFlowTravelTime>
    </basicDataValue>
</elaboratedData>
<elaboratedData id="L129p04150">
    <basicDataValue xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:type="TravelTimeValue">
        <supplierCalculatedDataQuality>96.0</supplierCalculatedDataQuality>
        <affectedLocation>
            <locationContainedInGroup xsi:type="LocationByReference">
                <predefinedLocationReference>L129p04150</predefinedLocationReference>
            </locationContainedInGroup>
        </affectedLocation>
        <freeFlowSpeed>109.0</freeFlowSpeed>
        <freeFlowTravelTime>291.07</freeFlowTravelTime>
    </basicDataValue>
</elaboratedData>
<elaboratedData id="L129p04151">
    <basicDataValue xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:type="TravelTimeValue">
        <supplierCalculatedDataQuality>96.0</supplierCalculatedDataQuality>
        <affectedLocation>
            <locationContainedInGroup xsi:type="LocationByReference">
                <predefinedLocationReference>L129p04151</predefinedLocationReference>
            </locationContainedInGroup>
        </affectedLocation>
        <freeFlowSpeed>110.0</freeFlowSpeed>
        <freeFlowTravelTime>201.76</freeFlowTravelTime>
    </basicDataValue>
</elaboratedData>
```


**Figure 47. Screenshot. Example TomTom data.**
**Source: TomTom.**

# HIGH-RESOLUTION ASSET DATA

## Description

The parallel development of higher fidelity sensors, LiDAR, and efficient machine learning algorithms has enabled companies to begin collecting asset data at a high resolution and in realtime. Today, many passenger vehicles are equipped with powerful sensors, and many high-end connected and semi-automated vehicles have high-definition cameras and LiDAR. Companies have found ways to use those sensors and cameras to collect information about a vehicle's environment and use that data to update agency asset-management systems with the latest and greatest information.

For example, a vehicle can use machine learning to process an incoming image from an on-board camera like those found on some higher-end semi-autonomous vehicles and recognize a speed limit sign. Not only does the vehicle recognize that an object is a speed limit sign, but it can also determine the listed speed on the sign and compare it to the stored value (either in a remote or local database) and determine if the sign has changed. The use of these cameras on CVs is a form of crowdsourcing that can be used to support asset management.

Specialized vehicles can measure other asset states, such as pavement cracks, potholes, pavement markings, signs, etc. As vehicles collect data on these assets, they can compare them over time to determine if they are deteriorating at an expected, slower, or faster rate.

## Applications of High-Resolution Asset Data

High-definition asset data can update asset information for maintenance, operations, and planning purposes. For example, detecting the existence of a new sign or change in a sign can provide update information for internal maps and keep the asset inventory up-to-date for maintenance and operation purposes. Similarly, pavement conditions can be stored and analyzed over time to determine the rate of deterioration to prioritize maintenance investments.

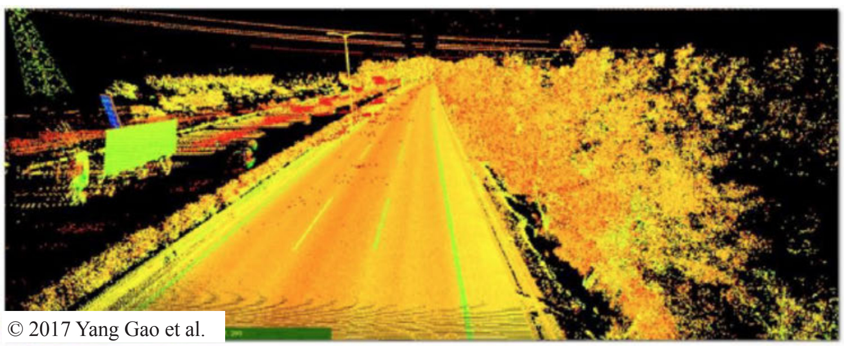
## Attributes

Data attributes vary based on technology and specific asset focus, but generally, data includes location, timestamp, and asset information. Asset information can be available as point clouds, images, deviation measurements, etc.

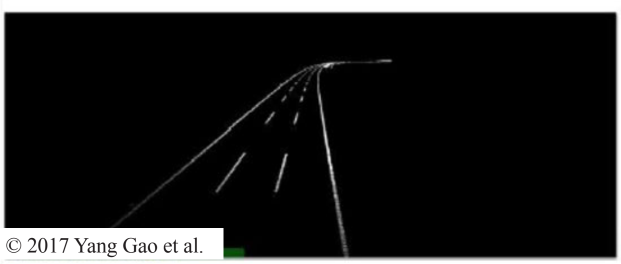
- **Latency:** Certain types of asset collection methods can be close to realtime. For example, camera identification of specific features using image processing and machine learning is quick and is available shortly after. On the other hand, LiDAR or three-dimensional scan data can require extensive processing of raw data to obtain a useful level of information for asset management.
- **Details:** High-resolution asset data generally has a high level of detail. These methods involve collection of high-resolution images or billions of point measurements transformed into very high quality asset information.
- **Quality and Coverage:** Quality of high-resolution data is associated with high cost, which affects the level of coverage possible. For example, LiDAR scans can be expensive and limit the extent of coverage.

**Point Clouds and Asset Mapping.** Point cloud data is sometimes considered a subset of high-resolution asset data. It is the underlying data on which some (but not all) high-resolution asset data is built. Point cloud is becoming a more popular way to identify and provide insights into assets on and adjacent to the road (or the condition of the road itself). Sometimes referred to as LiDAR data or three-dimensional scanner data, point cloud data collects extremely precise x,y,z coordinates that can be combined with timestamps, intensity, Red, Green, Blue (RGB) values, and other attributes to generate three-dimensional views of assets (signs, guard rails, pavement cracking, trees, etc.). A scanned point cloud for even a small area often consists of billions of points. At this scale, many agencies struggle to store and process these point clouds to generate meaningful asset information.

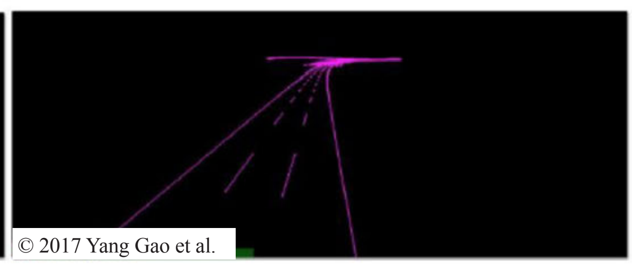
Transformation into mapped assets makes point clouds useful, as shown in figure 48. Companies specialize in extracting meaningful asset information from LiDAR point clouds. For example, one company extracted features such as lanes, guardrails, barriers, signs, street lights, drainage, and other assets and create a mapped asset database used by maintenance and operations at FDOT. Other companies are producing high-resolution maps and developing asset management solutions for DOTs now—all based off of point-cloud data, which they store, transform, and manage for the agency.




**a) Original point cloud data.**




**b) Point cluster of pavement markings.**




**c) Vectorized pavement markings.**

**Figure 48. Illustrations. Examples of point clouds and Asset Mapping.**

**Source: Yang Gao et al. (2017). Automatic extraction of pavement markings on streets from point cloud data of mobile LiDAR. Measurement Science and Technology. 28 085203. Available at [https://doi.org/10.1088/1361-6501/aa76a3](https://doi.org/10.1088/1361-6501/aa76a3).**

## Data Availability

A number of data providers offer high-resolution asset data using new technology and artificial intelligence. Some example providers include the following:

- A private sector data provider has marketed a new capability to detect signs using vehicle camera systems and use machine learning algorithms to determine the sign type and contents to inform the vehicle as well as update internal maps to account for the latest asset information.[8]
- Iowa State University has developed a method using end-to-end deep learning to identify specific objects with a focus on automated vehicles (figure 49). They used convolutional neural networks to identify and classify objects in near realtime using vehicle cameras. This method can recognize and inventory static assets either using crowdsourced data from the public or specially equipped agency vehicles.
- Another private-sector firm uses a system that consists of a series of displacement lasers mounted on a custom-designed semi-trailer that collects pavement structural response at highway speeds. These measurements provide pavement conditions in response to loads to determine if pavement is overloaded or if there is a need for overlay.
- One vendor has developed laser scanners that can be used to quickly (1 million points per second) generate high-resolution point clouds and create measurements, profiles, sections, contours, volumes, etc.[9]

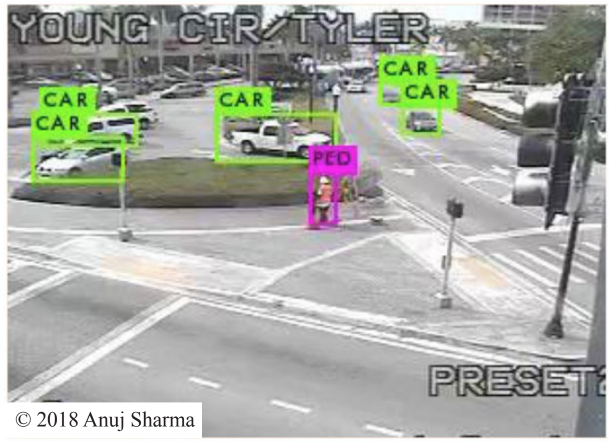





**Figure 49. Illustration. Deep learning object identification presented.**

**Source: Anuj Sharma, presentation to the 2018 TRB Annual Meeting Session P18-21443.**

## Pros and Cons

The primary benefit of the new technology and data collection methods for high-definition asset data is that data can be available in near realtime. Previously, asset data collection required significant effort and cost, leading to infrequent use. Now asset data collection occurs at a much higher frequency and by multiple vehicles, making the data available in nearly realtime.

Another challenge in using sensors, LiDAR, and cameras to collect asset data is that there could be situations in which collected data is inaccurate. For example, a database may claim that there is a speed limit sign at a particular location, but as the vehicles pass that location, a truck in an adjacent lane obstructs the sensor's view of the speed limit sign. In this situation, the sensor, camera or LiDAR may claim removal of the speed limit sign since it was not visible.

8 For more information, see https://www.here.com/en/products-services/here-automotive-suite/connected-vehicle-services/here-road-signs.

9 For more information, see https://geospatial.trimble.com/products-and-solutions/laser-scanning-solutions.

Similarly, some sensors, and many cameras, do not perform as well in adverse weather conditions such as rain, snow, and fog. In this case, data collected by a mobile sensor may be inaccurate or unreliable.

While data is valuable, it may be difficult to handle in its raw format. For example, until processed, point clouds or contours may not be useful or actionable for an agency to provide valuable information. Agencies rarely have the resources or expertise to process massive raw point cloud data sets, so they have to rely on third parties to process that data and package it for use in operations and maintenance.

## WI-FI AND BLUETOOTH RE-IDENTIFICATION DATA

### Description

This technology provides true, measured travel times and O-D data between two locations. Wi-Fi and Bluetooth re-identification data became a valuable source of actionable information as smartphone penetration increased to the point where almost every driver, passenger, bicyclist, and pedestrian has a smartphone that is equipped with Wi-Fi and/or Bluetooth. In recent years, many vehicles also have Wi-Fi or Bluetooth transceivers as well. Wi-Fi and Bluetooth data is primarily location data collected by static Wi-Fi and Bluetooth scanners that can identify Wi-Fi or Bluetooth devices at different locations and use that information to infer speed and travel time information, and if deployed extensively throughout a network, can also be used for travel pattern and O-D analysis.

### Applications of Wi-Fi and Bluetooth Re-identification Data

Some of the key applications of this data set include:

- Travel time calculation.
- Realtime queue warning and information.
- Traffic signal optimization.
- O-D analysis.
- Work zone management.

### Attributes

Most Bluetooth and Wi-Fi data consists of a media access control address (MAC address), a unique device address that identifies it as it moves through the network. For example, a vehicle is detected by one sensor that records the vehicle's MAC address. As that vehicle passes a second sensor, the MAC address is captured again. The distance between the two devices divided by the time between re-identification provides speed and travel times.

### Data Availability

Several private companies provide this data. Some providers offer Bluetooth roadside sensors that detect Bluetooth devices in realtime to approximate travel time. Another company provides Bluetooth devices that can be used to detect passing vehicles and pedestrians to obtain O-D, travel time, and trip pattern information. At least one company offers remote configurable devices that identify both Bluetooth and Wi-Fi devices and can determine travel patterns in various weather and traffic conditions. Another provides small portable devices as well as complete pole-mount sensor solutions for temporary or permanent data collection.

### Pros and Cons

One of the primary benefits of Bluetooth and Wi-Fi data is that it provides direct insight into individual trips between each sensor. Sensors are relatively inexpensive and lightweight, so deployment to address a core data need in an area is easy to accomplish.

The down side of these sensors is that they still require power and communications back to the agency. They are a physical device that needs to be physically deployed and maintained. Data is only collected where sensors have been deployed, making them significantly more costly to deploy on a wide scale when compared to probe data. Data can only be collected from vehicles and Wi-Fi/Bluetooth-equipped devices that are turned on and actively broadcasting their signal. Data filtering needs to be built into the device or the agency to remove anomaly data from vehicles that pull over into retail areas and do not return to the road network for an extended time. Finally, Wi-Fi pollution and general interference can result in distortion or data drops.

## CREDIT CARD TRANSACTION DATA

The All Hazards Consortium and the Department of Homeland Security use data from credit card point-of-sale machines to detect which businesses are open or closed following significant events (like a hurricane, snowstorm, or terrorist attack). The credit card swipe data serves as a surrogate measure of power outage information. Using the crowdsourced credit card swipes is an easy way to globally detect whether services are being restored properly in an area—which helps traffic operations, first responders, and maintenance crews understand where to focus their response and recovery efforts.

### Applications of Credit Card Transaction Data

The All Hazards Consortium's FLEET OPEN/CLOSED Service is designed to help locate open places of business that provide fuel, food, medications and medical supplies, retail stores, and hotel rooms during a prolonged power outage within a city, county, State, region, or across the United States. The basic benefit will include the ability to find fuel, food, pharmacies, and hotel locations during power outages. This information is for commercial use only and is not available to the public, ensuring that non-operational users do not overload the system. Besides this capability, the system can be used to support collaboration with electric companies regarding details on outage areas and where critical infrastructure exists that can be coordinated within the response coordination cycle with the electric sector liaison or contact person.

The service is available 24 hours per day with electric updates made multiple times every day. This provides the latest information in near realtime. Navigating this site is simple and enables users to zoom in on a particular region, State, or city. This information is encrypted, secure, protected, and restricted to users only.

### Attributes

There is no personally identifiable information in these data feeds. The data only has the number of transactions and time of each transaction. There is also positioning information (addresses and business names) that tie the transactions back to a specific location.

### Pros and Cons

The data is only useful during major events that affect power in a region. Therefore, the agency is paying for data not used on a regular basis. The agency must also manipulate the data into a surrogate for power outages. For example, one can keep a record of transactions for all businesses over time, and then provide warnings to users when zero transactions are occurring during a specific time when businesses should be open and operational.

The data provider requires additional levels of security and credentialing to be placed on users of this data. For example, the provider may require a specific level of encryption and a radio frequency identification (RFID)[10] card for users who want to see the data. This is a relatively minor technical barrier, but can be an operational barrier if agencies need to purchase new computer equipment with card readers and RFID badges for employees.

## CONNECTED VEHICLE DATA FROM THIRD PARTIES

### Description

Many vehicle manufacturers assembled millions of vehicles with embedded CV technology (figure 50). Data from these providers can include wiper use, headlight use, heavy braking, traction control, fuel consumption, emissions, travel speed, acceleration and deceleration, and more. Example applications include using wiper activation data for rain and micro-level weather predictions, using heavy braking events for end-of-queue or debris detection, or using traction control engagement events for detecting slippery or icy road conditions. These predictions and detections allow agencies to pre-position maintenance and response crews to mitigate the impacts of events.




**Figure 50. Illustration. Third-party, connected-vehicle events from the Washington, D.C. region.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

10 Radio frequency identification uses electromagnetic or electrostatic coupling in the radio frequency portion of the electromagnetic spectrum to uniquely identify an object, animal, or person. This technology can be used to automatically identify and track tags attached to objects, such as cards.

## Data Availability

Telematics providers have been making this data available to consumers and freight logisticians for years. However, some mapping and location data companies are on the cusp of making this data available to DOTs as part of their regular incident data feeds. Millions of these vehicles are operating on the Nation's roadways—more vehicles than in any CV Pilot in the United States today—and are reporting events precisely as they happen, exactly where they happen, without any need for installed roadside infrastructure. The data comes to a central server via wireless, cellular technologies embedded in vehicles. The data is anonymized, cleaned, and then provided as an incident feed to agencies and automobile manufacturers.

## Applications of Connected Vehicles Data

**Surface Weather Event Detection.** When a driver turns on windshield wipers, this becomes a surrogate for the arrival of rain and (usually) the beginning of a period with wet and slippery pavement. Plotting exactly where and when dozens of vehicles turn on windshield wipers allows not just knowledge of where it is raining, but realtime calculation of the movement of weather at the micro-level; this will enable predictions about where it will hit next. Furthermore, reporting heavy braking, traction control engagement, travel speed, acceleration, and deceleration to TMCs without latency allows smart networks to react almost reflexively as major flow disruptions occur (e.g., to warn vehicles to slow down early).

**Early Event Notifications.** Air bag deployments, rollover indicators, and extreme deceleration can immediately alert agencies of a potential incident on the roadway in advance of other detection technologies (like a 911 call, service patrol identification, etc.).

## Attributes

Data attributes vary from provider to provider, but look very similar to agency incident data. There is usually an event type, location, time stamp, etc.

## Pros and Cons

Because this data is automatically collected over any road network that has access to cellular data service, the DOT can quickly gain enhanced situational awareness on parts of their road network that would otherwise lack coverage. CV data can also alert the DOT of potentially dangerous events and surface conditions long before a scene becomes a true incident. This could allow the agency to be proactive instead of reactive—for example, in responding to slippery surface conditions noticed by traction control engagements in multiple vehicles. Alternatively, the data could potentially be voluminous and distracting to operators—much like crowdsourced data. Therefore, some additional filtering and automated integration with the ATMS is generally recommended.

## REALTIME TURNING MOVEMENT DATA

### Description

Realtime turning movement data describes how many vehicles are taking a particular exit, turn, or are traveling through an intersection. Historically, this data was collected only via deployment of an extensive sensor network on exit ramps and at intersection approaches and exits. However, third parties are now directly measuring or estimating turning movements from a mix of probe data and other sensing technologies.

### Applications for Realtime Turning Movement Data

Realtime turning movement data allows transportation agencies to determine driver compliance rates with routing advice. With the combination of high-resolution signal data and the commercial vehicle CV data, a transportation agency can realize greater benefits from integrated corridor management (ICM) concepts that link freeway operations and signalized arterial operations. Freeway and arterial operations will be able to coordinate better through shared information about network flows, rerouting, and the impacts of incidents and events. Agencies could suggest alternate routes and actually measure what percentage of vehicles are taking their advice (or making up their own paths).

### Attributes

The attributes vary between data providers, but one provider gives the percentage of traffic taking Exit A versus B at each intersection or turning left versus right versus going straight on arterials.

### Pros and Cons

The main drawback is likely that realtime turning movement data is new and the larger transportation community has not fully tested and validated it. Agencies that purchase this data will be able to derive new operational insights, but, since some of the data may be faulty, they need to perform some level of validation to ensure realistic decisions. Despite this caution, there is significant promise for this type of data.

## HIGH-RESOLUTION SIGNAL DATA

Detailed SPaT plus sensor actuation data has a resolution of 1/10 of a second (10 Hz). This means that all of the fundamental metrics needed to understand how each signal is operating and how it is loaded with traffic will be available to the decision support system in realtime. A new class of small and relatively inexpensive “aggregator” devices makes it possible to interface with the controller inside each cabinet. These devices not only bundle sensor data streams and ship them to the TMC but also receive and implement commands coming from the TMC. This capability provides a low-cost connection for decision support systems (DSS) to control precisely the phases and timings of large networks of signals at once. (These devices work with traffic signals of any age or technology, with or without a hardwired communication network, giving it broad usefulness for deployments later nationwide.)

## Description

Traditionally, snapshot view retiming methods managed signalized corridors that required manual data collection and modeling. As a result, daily counts were not representative of changing conditions throughout the day. This gave rise to signal timing that was not responsive to conditions.

New data collection technology built into controllers generated high-resolution signal data recorded over 150 different controller events at a very high frequency. This high-resolution data allows operators to evaluate signal system performance in realtime and to provide strong support for accurate optimization of the system to manage traffic flow.

## Applications of High-Resolution Signal Data

The primary application of high-resolution signal data is in measuring the performance of a signal system in realtime to be able to optimize the system and adapt to the changing conditions in an automated manner. This realtime optimization takes the place of traditional signal operation modeling and estimation. From a planning perspective, archived high-resolution signal data can provide insights in seasonal variations and general changes over time. Similarly, this enables the use of high-resolution data to evaluate queues, turn movement counts, and volume to capacity ratios. In addition to mobility improvement applications, high-resolution signal data also provides safety benefits. It allows operators to detect and record red light and right-turn-on-red violations. Examples of the performance measures supported include:

- **Percentage of Arrivals on Green.** Whereas re-identification provides effective metrics for a corridor containing several signals, high-resolution controller data provides detailed information on the performance of each signalized intersection. In addition to mining the high-resolution data for streamlining maintenance, the sensor actuation data combined with SPaT can be used to judge the health of signal coordination by measuring the percentage of vehicles that arrive on the green phase. The Purdue Coordination Diagram can cascade to all signals in a corridor and provides an effective visual of the quality of progression.
- **Capacity Utilization at Intersections.** The inability of a vehicle to clear a light on a single cycle evokes driver dissatisfaction and complaint calls. Techniques to monitor capacity utilization, namely the frequency of occurrence of split failures, are available through high-resolution data. These metrics indicate when a movement on a phase is over capacity and requires more green, or if multiple phases are failing simultaneously, it indicates the need for capital improvement or demand management.
- **Red-Light and Back-of-Queue Collision Reduction.** Waze and others in the United States are already piloting efforts to share realtime SPaT data with third-party traveler information providers and navigation systems to alert users of impending red lights and green lights.

## Attributes

High resolution signal data includes more than 150 different elements broken into active phase events, active pedestrian phase events, barrier and ring events, phase control events, overlap events, detector events, preemption events, coordination events, and cabinet/system events. All events are timestamped.

Each collection of events comprises a detailed list of signal and detector states. Collection of controller events at 10Hz frequency means that operators have visibility into the state of the signal and intersection at a very fine granularity.

### Pros and Cons

The primary benefit of high-resolution signal data is that it provides many data points at a high frequency, which allows operators to understand the realtime conditions at the intersection and quickly react to or even predict congestion and poor signal performance. Agencies have identified a number of previously unattainable performance measures calculated using this data.

Additionally, high-resolution signal controllers are inexpensive and can be phased into existing systems, allowing agencies to budget for and deploy high-resolution systems more quickly and affordably. The down side of this data set is that, at such a high frequency, it requires appropriate expertise and infrastructure to store and process.

## AIR QUALITY SENSOR DATA

### Description

Air quality sensors are capable of identifying primary pollutants, generated directly by source, as well as secondary pollutants resulting from chemical reactions in surrounding air. The sensors collect a variety of measurements, including nitrogen dioxide, ozone, carbon monoxide, hydrogen sulfide, sulfur dioxide, and others. These sensors often can collect additional data elements such as sound and vibration, temperature, and in some cases light intensity. These low-cost sensors provide a hyperlocal view of air quality that can change from block to block and intersection to intersection in a dense urban area.

### Applications of Air Quality Sensor Data

Air quality sensor data has several uses, especially in urban areas. For example:

- Evaluating the level of air pollutants and providing operators with the ability to direct non-motorized travelers away from poor air quality areas to areas with better air quality, thereby encouraging or discouraging bicycle and pedestrian activity in particular areas.
- Determining air quality at intersections for traffic signal system control. Air quality can become an input variable in timing plans to limit the level of pollution from too many vehicles or heavy-pollutant vehicles idling for too long at a signalized intersection. In effect, air quality-influenced traffic signals can indirectly address congested intersections while also considering the environmental impact factor of vehicles in the congestion.
- Sound and vibration measurements used in conjunction with air quality measures to identify heavy vehicles moving through an urban area. This can help determine and direct urban freight movement patterns over delivery areas and inform effective parking utilization strategies.

### Attributes

Sensors record timestamps and air concentration measures for a variety of pollutants, chemicals, temperature, pressure, and in some cases sound and vibration, as well as light intensity.

### Data Availability

Many different vendors manufacture and sell air quality sensors. Several organizations provide localized air quality data. For example, the City of Chicago is equipped with over 100 air quality sensors as part of the Array of Things[11] collaborative effort between the Urban Center for Computation and Data (joint initiative of Argonne National Laboratory and University of Chicago) and the City of Chicago along with a number of other universities across the country. As part of this project, the city and University of Chicago have published the collected data for public consumption.

### Pros and Cons

Air quality sensors are affordable, but they also provide hyper-local measurements, which means that in order to get a good aerial or regional insight into air quality, an agency would need to deploy a network of air quality sensors. In the case of transportation operations and planning applications, air quality sensors can provide valuable supplemental information, but are rarely sufficient on their own.

## ROADWAY WEATHER PREDICTIONS

### Description

While basic National Weather Service prediction data has been available for many years, these predictions cover wide areas and tend to focus more on air temperature and precipitation averaged over a region. Several companies now offer ground-based (i.e., at street level) 48-hour weather predictions that are updated every hour. This information can be used to optimize winter weather response operations; improve snow event readiness; reduce staffing, fuel, and chemical costs; pinpoint treatment applications; and generally keep the roads safer and less congested. Farmers have utilized the weather prediction services of some of these companies for years, and now, seeing a similar operational need, they have branched out into the transportation domain. These companies are able to produce specialized and hyperlocal weather conditions and predictions from a mix of satellites, ground-based radar, ground-sensors, and most recently using connected vehicles' sensors (often from agency-owned plow and maintenance vehicles), and more. Specialized algorithms and data processing then produce more accurate ground-based weather predictions. Operators use this to understand the weather as expected and then experienced by the driver.

11 For more information, visit the "Array of Things" web page at https://arrayofthings.github.io/. The Array of Things is a collaborative effort among scientists, universities, local government, and communities in Chicago to collect realtime data on the city's environment, infrastructure, and activity for research and public use.

## Application of Weather Predictions

Weather can quickly and dramatically impact safety and mobility on roads. It can impact visibility. High winds can blow trucks off of roads and bridges. Icy roads can significantly impact vehicle performance. Table 12 shows the impacts on roads, traffic, and operations decisions.

**Table 12. Weather impacts on roads, traffic, and operational decisions.**

| Road Weather Variables | Roadway Impacts | Traffic Flow Impacts | Operational Impacts |
|---|---|---|---|
| Air Temperature and Humidity | N/A | N/A | ▪ Road treatment strategy (e.g., snow and ice control).<br>▪ Construction planning (e.g., paving and striping). |
| Wind Speed | ▪ Visibility distance (due to blowing snow, dust)<br>▪ Lane obstruction (due to wind-blown snow, debris) | ▪ Traffic speed<br>▪ Travel time delay<br>▪ Accident risk | ▪ Vehicle performance (e.g., stability).<br>▪ Access control (e.g., restrict vehicle type, close road).<br>▪ Evacuation decision support. |
| Precipitation (Type, Rate, Start/End Times) | ▪ Visibility distance<br>▪ Pavement friction<br>▪ Lane obstruction | ▪ Roadway capacity<br>▪ Traffic speed<br>▪ Travel time delay<br>▪ Accident risk | ▪ Vehicle performance (e.g., traction).<br>▪ Driver capabilities/ behavior.<br>▪ Road treatment strategy.<br>▪ Traffic signal timing.<br>▪ Speed limit control.<br>▪ Evacuation decision support.<br>▪ Institutional coordination. |
| Fog | ▪ Visibility distance | ▪ Traffic speed<br>▪ Speed variance<br>▪ Travel time delay<br>▪ Accident risk | ▪ Driver capabilities/ behavior.<br>▪ Road treatment strategy.<br>▪ Access control.<br>▪ Speed limit control. |
| Pavement Temperature | ▪ Infrastructure damage | N/A | ▪ Road treatment strategy. |

**Table 12. Weather impacts on roads, traffic, and operational decisions. (continued)**

| Road Weather Variables | Roadway Impacts | Traffic Flow Impacts | Operational Impacts |
|---|---|---|---|
| Pavement Condition | ▪ Pavement friction<br>▪ Infrastructure damage | ▪ Roadway capacity<br>▪ Traffic speed<br>▪ Travel time delay<br>▪ Accident risk | ▪ Vehicle performance.<br>▪ Driver capabilities/behavior (e.g., route choice).<br>▪ Road treatment strategy.<br>▪ Traffic signal timing.<br>▪ Speed limit control. |
| Water Level | ▪ Lane submersion | ▪ Traffic speed<br>▪ Travel time delay<br>▪ Accident risk | ▪ Access control.<br>▪ Evacuation decision support.<br>▪ Institutional coordination. |

Source: Federal Highway Administration. Road Weather Management Program. How do Weather Events Impact Roads? Available at: https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm, last accessed March 22, 2019.

Having accurate weather predictions on surface conditions can help agencies alert motorists of unsafe driving conditions, allow for pretreatment, and inform ongoing operational strategies that can dramatically affect safety and mobility.

## Attributes

While each data and service provider is different, common data elements and attributes of road-weather prediction providers can include:

- Pavement and/or bridge temperatures.
- Air temperature.
- Wind speed and direction.
- Visibility.
- Precipitation type, rate, and/or accumulations.
- Salinity content of the moisture on the roadway.
- Ice/frost warnings.
- Comparison of realtime conditions to predicted conditions.

## Data Availability

This data is available from the private sector. One company has invested heavily in its weather solutions, which provide pavement and other weather predictions and data for maintenance and operations response. Originally developed for the agricultural industry, these weather solutions come in many forms, including a web-based client, mobile application, and data feeds. This company has even integrated some of its weather prediction products into other products and solutions, like signal systems applications and analytics.

Another company offers a solution that leverages existing DOT (or other) closed circuit television (CCTV) cameras and image-processing technologies. Leveraging thousands of already-deployed cameras, the company processes images from these cameras to detect precipitation, cloud cover, fog, and other environmental conditions, then merges that information with other National Weather Service predictions and observations. Web-based analytics provide alerts to operators when specific conditions are detected.

A few agencies have invested in in-house, full-time meteorologists. UDOT is one example, maintaining four full-time meteorologists in their Operations center 24 hours per day, 7 days per week. To reduce costs, these staff are allowed to provide meteorological services to other parts of the country in addition to their Utah geographies.

Coverage varies by vendor. At least one provider leverages existing CCTV infrastructure, and therefore, certain attributes and alerts are more focused on areas where cameras have already been deployed. Another provider generally covers the entire country and all roads.

### Pros and Cons

Realtime and predicted road weather information can be extremely beneficial to agencies operating in areas where winter weather, flooding, or other events can dramatically affect mobility and safety. The disadvantages to this data are relatively minimal and include cost, geographical resolution, and operator impacts. For example, while these services are higher resolution than most National Weather Service predictions, they are still limited to 1-kilometer square resolutions. This type of resolution is generally good enough for most maintenance applications. Also, high-quality road weather data can be relatively expensive; however, the benefit-cost can be significant, reducing maintenance costs, excess materials used, and fuel consumption and optimizing resource management. It can also have a significant impact on mobility, traveler information, and safety.

### Use Cases for Roadway Weather Predictions

**Use Case: Indiana DOT.** The Indiana DOT uses the Iteris ClearPath weather prediction software to help better manage the deployment of their snow plows and sand/salt trucks. With approximately 1,000 trucks in its fleet, the agency estimates that they can save over $750,000 in material costs alone ($750 per load of salt) if they can reduce their deployments by one trip. The use of more accurate weather predictions can help achieve this on multiple occasions throughout the year. This type of cost savings (which is conservative in that it only estimates the costs of materials, not operator time, fuel, equipment costs, or the reduction in crashes and safety impacts) quickly justifies the cost of the weather prediction service.

**Use Case: UDOT.** UDOT has taken a different approach to predicting road weather conditions. As noted above, beginning in 2002, the agency employed its own full-time, in-house meteorologist in their traffic operations center (TOC). Unlike most weather reporting, the UDOT team does not provide a percentage chance of precipitation. Instead, they predict whether it will or will not snow, rain, etc. on different sections of roadways.

Established under the TOC's Traffic Management Division, UDOT's weather program has two main components:[12]

- Four staff meteorologists stationed in the TOC provide year-round weather support.
- An ITS component, which manages approximately 70 RWIS stations and expert systems such as bridge spray systems, high wind alerts, and fog warnings.

The meteorologists provide services like:

- Forensic meteorology services (e.g., risk management, filling requests for data made through the Utah Government Records and Access and Management Act (GRAMA)).
- Forecasting services to UDOT maintenance, construction, and TOC personnel.
- RWIS and weather training courses.

Year-round, long-term weather forecasts are provided and used mainly for planning materials, staffing, and equipment. They provide pre-storm, during-storm, and post-storm weather forecasts to the maintenance managers, area supervisors, and local garages. They also provide forecast services for road rehabilitation and avalanche safety. Construction engineers and contractors also receive weather forecasts from the meteorologists for new construction and renovation projects.

A 2008 research study was conducted to estimate the cost-effectiveness of the program on winter maintenance costs.[13] The study estimated the value and additional saving potential of the UDOT weather service to be 11–25 percent and 4–10 percent of the UDOT winter maintenance costs, which include the costs of labor and materials. Based on this program's cost, the benefit-cost ratio was calculated at more than 11:1.

In addition to maintenance and construction, the meteorologists provide weather forecasts to TOC divisions including:

- Signal systems.
- ATMS.
- Incident management team.
- Traffic management.
- Department of public safety.

They issue weather forecasts twice a day or more frequently as needed when conditions are changing rapidly.

**Additional Use Cases:** The Federal Highway Administration's (FHWA's) Road Weather Management Program has documented 27 additional use cases for road weather operations.[14]

12 Zhirui Ye. 2009. *Evaluation of the Utah DOT Weather Operations/RWIS Program on Traffic Operations*. Iowa Department of Transportation: Ames, IA. Available at: https://westerntransportationinstitute.org/wp-content/uploads/2016/08/4W2324_UDOT_Phase_II_Report.pdf, last accessed March 18, 2019.

13 Strong, C. and Shi, X. 2008. "Benefit–Cost Analysis of Weather Information for Winter Maintenance: A Case Study," *Transportation Research Record: Journal of the Transportation Research Board* 2055:119-127.

14 Federal Highway Administration, Road Weather Management Best Practices. Available at: https://ops.fhwa.dot.gov/weather/best_practices/1024x768/transform2.asp?xslname=casestudies_title.xslt&xmlname=casestudies.xml.

## COMPUTER-AIDED DISPATCH DATA

### Description

First responders and public safety agencies use CAD systems to identify the location of an emergency, the type of emergency, and then dispatch units to the scene. Generally, a 911 operator or a dispatcher who receives a call will enter basic information into a computerized system that pushes that information out to responders' mobile data terminals. The system can also be used by dispatchers to visualize field unit assignments to different calls. Some CAD systems are integrated with automated vehicle location (AVL) systems as well as telecommunications systems to allow for quicker information sharing and communications.

Traditionally, CAD systems have been closed secure systems used exclusively by public safety personnel to manage field operations. However, most traffic incidents are first reported via cell phone calls to a 911 center. For example, up to 88 percent of incidents in Virginia were first discovered by the Virginia State Police. This means that public safety agencies are often aware of a traffic incident well before a TOC. By the time the DOT becomes aware of an incident, either via a courtesy or procedural call from the partner public safety agency, or by detecting it using CCTV or other ITS equipment, it could be 5–10 minutes or even longer after the incident occurred. That initial time after the incident is critical in achieving efficient and effective medical response and quick clearance.

Over the last 15 years, many agencies have tried to integrate information from CAD systems into TMC operations to provide quicker incident detection and more efficient response. Some initial efforts focused on co-location of public safety and TMC personnel, but even in those instances, the communication and information sharing were inconsistent and limited. In the more recent past, agencies have begun integrating CAD systems data directly with the ATMS in a TOC to automate data exchange and reduce latency between when an incident occurs and when the TMC becomes aware of it. However, even these system integration efforts have encountered challenges related to information overload, sensitivity, and security.

### Applications of Computer-Aided Dispatch Data

CAD data is mainly used to identify incidents more quickly and to identify incidents that a TMC would not otherwise be aware of either due to lack of coverage or jurisdiction. TMCs have integrated CAD data into their systems is several different ways:

- TMC/Public Safety Co-Location
  - Instead of direct system integration, co-location of staff sometimes allows TMC operators to interact directly with public safety staff and exchange information.
  - This type of information exchange is highly dependent on interpersonal relationships and space configuration, and therefore varies in effectiveness.
- Read-Only Integration
  - In a read-only arrangement, TMCs may have a scrolling screen that just shows CAD messages as they arrive.

    - Alternatively, ATMS may flash an alert when a CAD message arrives and allow the operator to create an ATMS event based off of CAD data and manage it separately.
    - This is the most common type of integration and it exposes two key challenges:
        - Operators can be overwhelmed by alerts and start ignoring them, especially if there are a lot of false-positives.
        - Once a TMC operator creates a related event in the ATMS, the two records continue to diverge as CAD data changes, and the operator may or may not include those changes.
- Two-Way Integration
    - Two-way integration is the least common and most complex way to integrate CAD data. This method includes two-way data exchange between ATMS and CAD systems.
    - Generally, CAD events arrive to the TMC and then TMC operators are able to modify and add information to the event, which then gets sent back to the public safety CAD system.
    - While this type of integration may provide most flexibility and the richest data set, it can also be complex to implement and cause many jurisdictional issues with operators from different agencies entering conflicting data or overwriting each other's changes.

## Attributes

**Latency:** Information is input into a CAD system in realtime as it occurs. As a call arrives to the 911 system, the dispatcher inputs relevant data into the system. This data is then pushed out to relevant field personnel and partners. However, information flowing to partners has to traverse several networks and firewalls between agencies. In addition to time for information to traverse the path between systems, some agencies and CAD systems have limitations when it comes to sharing mechanisms that prevent asynchronous information sharing. These systems may rely on some type of timer to dump and consume data, maybe once a minute or once every several minutes.

**Details:** CAD data varies in level of detail and structure depending on agency and vendor. Generally, CAD data includes basic information including incident type, timestamp, and dispatched unit information. Some CAD data feeds also contain information about location, actions taken during the incident, caller information and status, etc.

An example incident record coming from a CAD system to a TMC looks like this:


```
<Incident>
      <Authorization></Authorization>
      <IncidentId>121150898573862915</IncidentId>
      <TrackingNumber>2018-00494892</TrackingNumber>
      <IncidentCreatedTimestamp>2018-11-02T23:01:38</
IncidentCreatedTimestamp>
      <IncidentEffectiveTimestamp>2018-11-02T23:01:47</
IncidentEffectiveTimestamp>
      <IncidentLastModifiedTimestamp>2018-11-02T23:01:47</
IncidentLastModifiedTimestamp>
      <IncidentAddress>123 EXAMPLE HWY </IncidentAddress>
```

```
    <IncidentStreetPrefix></IncidentStreetPrefix>
    <IncidentHouseNum>123</IncidentHouseNum>
    <IncidentStreetName>EXAMPLE</IncidentStreetName>
    <IncidentStreetSuffix>HWY</IncidentStreetSuffix>
    <IncidentPostDirectional></IncidentPostDirectional>
    <Building></Building>
    <Apartment></Apartment>
    <Community>EXAMPLE TOWN</Community>
    <ExtendedLocation>EXAMPLE BUSINESS NAME</ExtendedLocation>
    <IncidentLatitude>38.89512195241237</IncidentLatitude>
    <IncidentLongitude>-77.03649619898232</IncidentLongitude>
    <IncidentLocationAreaName>962</IncidentLocationAreaName>
    <MsagIntersection>EX ST,EXAMPLE TOWN</MsagIntersection>
    <MsagIntersection>US EXAMPLE HWY,EXAMPLE TOWN</MsagIntersection>
    <MsagZip></MsagZip>
    <MsagState>MD</MsagState>
    <MsagOddEven>EVEN</MsagOddEven>
    <MsagC1>962</MsagC1>
    <MsagC2>1107902</MsagC2>
    <MsagC3>500126 &amp; 502333</MsagC3>
    <MsagC4>502214 &amp; 503724</MsagC4>
    <MsagC5></MsagC5>
    <MsagC6></MsagC6>
    <MsagC7></MsagC7>
    <MsagC8>EXAMPLE TOWN</MsagC8>
    <MsagC9></MsagC9>
    <MsagC10>EXAMPLE COUNTY</MsagC10>
    <MsagC11>6, NORTHERN</MsagC11>
    <MsagC12>96-NORTH</MsagC12>
    <MsagC13>COUNTY</MsagC13>
    <MsagC14>12-02-15</MsagC14>
    <MsagC15></MsagC15>
    <MsagC16></MsagC16>
    <MsagC17></MsagC17>
    <MsagC18></MsagC18>
    <MsagC19></MsagC19>
    <MsagC20></MsagC20>
    <IncidentEventRefName>71 PREMISE CHECKS</IncidentEventRefName>
    <IncidentEventRefDescription>71 PREMISE CHECKS</
IncidentEventRefDescription>
    <IncidentPriorityRefName>3</IncidentPriorityRefName>
    <CurrentIncidentStatusRefName>CLOSED</
CurrentIncidentStatusRefName>
    <IncidentCreatorUserName>CAD0001</IncidentCreatorUserName>
    <IncidentFirstDispatchTimestamp>2018-11-02T23:01:36</
IncidentFirstDispatchTimestamp>
```

```
      <IncidentFirstEnrouteTimestamp>2018-11-02T23:01:37</
IncidentFirstEnrouteTimestamp>
      <IncidentFirstOnSceneTimestamp>2018-11-02T23:01:38</
IncidentFirstOnSceneTimestamp>
      <IncidentCloseTimestamp>2018-11-02T23:01:47</
IncidentCloseTimestamp>
      <IncidentDisposition>
            <IncidentDispositionRefName>C</IncidentDispositionRefName>
            <IncidentDispositionRefDescription>No further action</
IncidentDispositionRefDescription>
            <IncidentDispositionUserName>CAD0001</
IncidentDispositionUserName>
            <IncidentDispositionRemarks></IncidentDispositionRemarks>
            <IncidentDispositionEffectiveTimestamp>2018-11-02T23:01:47</
IncidentDispositionEffectiveTimestamp>
      </IncidentDisposition>
      <CallForService>
            <CallerName></CallerName>
            <CallerPhoneNumber></CallerPhoneNumber>
            <CallerCallBackNumber></CallerCallBackNumber>
            <CallSourceRefName>Offc</CallSourceRefName>
            <CallStatusRefName>CLOSED</CallStatusRefName>
            <CallerAddress></CallerAddress>
            <CallerStreetPrefix></CallerStreetPrefix>
            <CallerHouseNum></CallerHouseNum>
            <CallerStreetName></CallerStreetName>
            <CallerStreetSuffix></CallerStreetSuffix>
            <CallerPostDirectional></CallerPostDirectional>
            <CallerBuilding></CallerBuilding>
            <CallerApartment></CallerApartment>
            <CallerCommunity></CallerCommunity>
            <CallerExtendedLocation></CallerExtendedLocation>
            <CallerLatitude></CallerLatitude>
            <CallerLongitude></CallerLongitude>
            <CallerLanguage></CallerLanguage>
            <CallerLanguage></CallerLanguage>
            <CallForServiceEffectiveTimestamp>2018-11-02T23:01:47</
CallForServiceEffectiveTimestamp>
            <CallForServiceLastModifiedTimestamp>2018-11-02T23:01:47</
CallForServiceLastModifiedTimestamp>
      </CallForService>
      <ServiceRequest>
            <ServiceRequestPriorityRefName>1</
ServiceRequestPriorityRefName>
```

```
            <ServiceRequestStatusRefName>CLOSED</
ServiceRequestStatusRefName>
            <ServiceRequestUnitOrgName>96 EXAMPLE TOWN</
ServiceRequestUnitOrgName>
            <ServiceRequestUnitOrgPath>CAD##96 EXAMPLE TOWN</
ServiceRequestUnitOrgPath>
            <ServiceRequestUnitOrgEventRefName></
ServiceRequestUnitOrgEventRefName>
            <ServiceRequestUnitOrgEventRefDescription></
ServiceRequestUnitOrgEventRefDescription>
            <ServiceRequestUnitOrgEventRefPriorityRefName>1</
ServiceRequestUnitOrgEventRefPriorityRefName>
            <ServiceRequestEffectiveTimestamp>2018-11-02T23:01:47</
ServiceRequestEffectiveTimestamp>
            <ServiceRequestLastModifiedTimestamp>2018-11-02T23:01:47</
ServiceRequestLastModifiedTimestamp>
            <Unit>
                  <UnitServiceRequestUnitName>0001</
UnitServiceRequestUnitName>
                  <UnitServiceRequestStatusRefName>Available</
UnitServiceRequestStatusRefName>
                  <UnitServiceRequestPreviousStatusRefType>AVAILABLE</
UnitServiceRequestPreviousStatusRefType>
                  <UnitFirstDispatchUnitStatusRefName>Dispatched</
UnitFirstDispatchUnitStatusRefName>
                  <UnitFirstDispatchTimestamp>2018-11-02T23:01:36</
UnitFirstDispatchTimestamp>
                  <UnitFirstEnrouteUnitStatusRefName>EnRoute</
UnitFirstEnrouteUnitStatusRefName>
                  <UnitFirstEnrouteTimestamp>2018-11-02T23:01:37</
UnitFirstEnrouteTimestamp>
                  <UnitFirstOnSceneUnitStatusRefName>OnScene</
UnitFirstOnSceneUnitStatusRefName>
                  <UnitFirstOnSceneTimestamp>2018-11-02T23:01:38</
UnitFirstOnSceneTimestamp>
                  <UnitFirstSecondaryEnrouteUnitStatusRefName></
UnitFirstSecondaryEnrouteUnitStatusRefName>
                  <UnitFirstSecondaryOnSceneUnitStatusRefName></
UnitFirstSecondaryOnSceneUnitStatusRefName>
                  <UnitLastClearUnitStatusRefName>Available</
UnitLastClearUnitStatusRefName>
```

```
                <UnitLastClearTimestamp>2018-11-02T23:01:47</
UnitLastClearTimestamp>
                <UnitServiceRequestEffectiveTimestamp>2018-11-
02T23:01:47</UnitServiceRequestEffectiveTimestamp>
                <UnitServiceRequestLastModifiedTimestamp>2018-11-
02T23:01:47</UnitServiceRequestLastModifiedTimestamp>
                <PersonUnit>
                     <PersonUnitName>John Doe</PersonUnitName>
                     <Employee>
                          <EmployeeGivenName>John</EmployeeGivenName>
                          <EmployeeMiddleNames></EmployeeMiddleNames>
                          <EmployeeSurName>Doe</EmployeeSurName>
                          <EmployeeBadgeNumber>0001</
EmployeeBadgeNumber>
                          <EmployeeIdNumber>0001</EmployeeIdNumber>
                     </Employee>
                </PersonUnit>
                <UnitHomeUnitOrgPath>CAD##96 TOWN NAME</
UnitHomeUnitOrgPath>
           </Unit>
     </ServiceRequest>
</Incident>
```

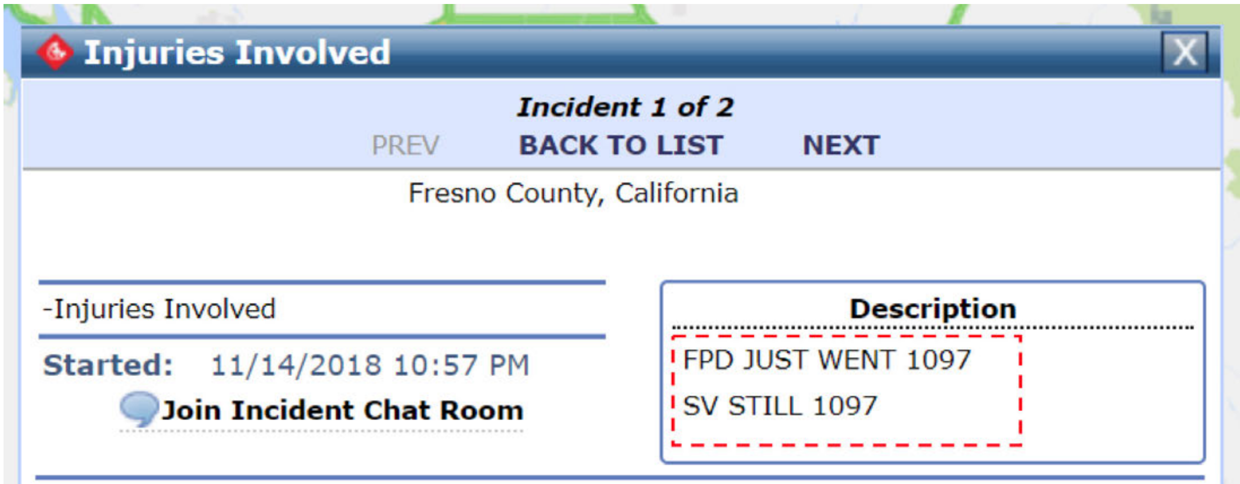

## Data Availability

Most public safety agencies at the State level have CAD systems that were developed and maintained by different vendors.

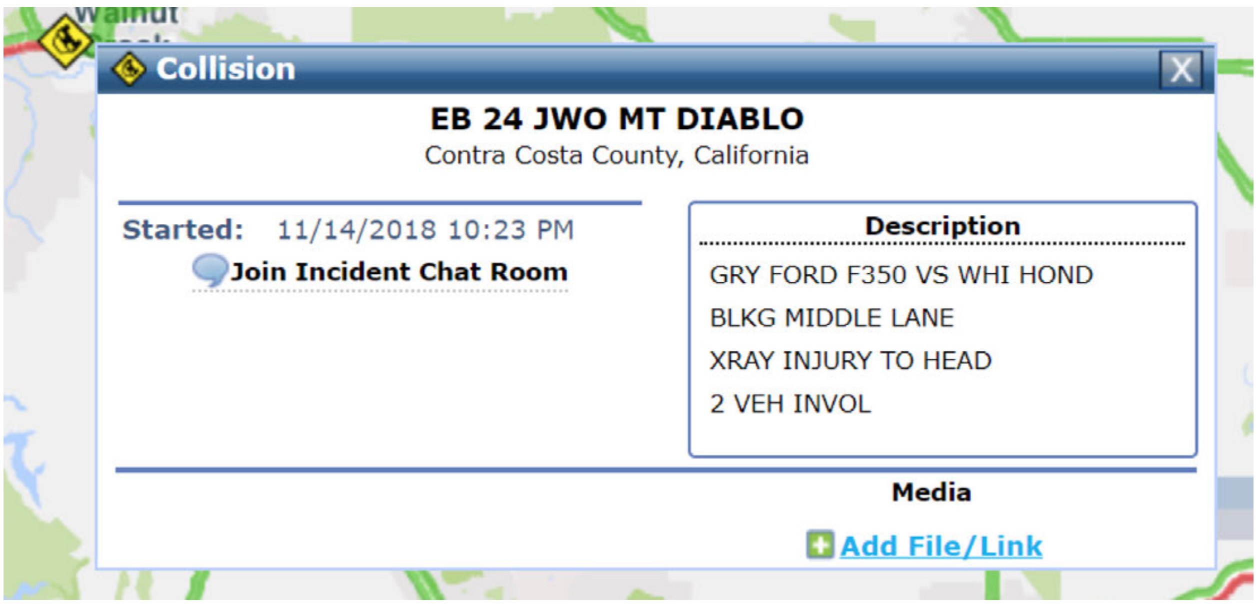
In a case where public safety and TMCs share information, each public safety agency works directly with a TMC to develop sharing mechanisms through their individual vendor applications. CAD data is not readily available to all TMCs across the country. Agencies are still working with their public safety partners to garner support and develop methods for exchanging this valuable data.

CAD data coverage is directly related to a specific public safety agency jurisdiction. This means that municipal police CAD data may cover an urban area, while county police CAD data may cover arterial and county roads, and State police CAD data may be focused on interstates.

The quality of data varies depending on the agency and system capabilities. Some systems provide a rich set of data elements, while others rely on mostly free-text entry that, while readable to a human, may not be very useful when it comes to automated integration into other systems. Figure 51 and figure 52 show examples of different CAD data that has been integrated into different ATMS platforms. The first two are from the State police in California, and the third is from the county police in Maryland. While one uses short-hand, abbreviations, and “10-codes” to describe the event, the other agency uses machine-readable text and “signal codes” to describe the event.




a) Difficult to read, free-text typed into the California Highway Patrol's computer-aided dispatch system (dashed box) and transmitted to Caltrans.






b) Example text free-typed to describe an incident, including use of verbal shorthand.

**Figure 51. Screenshot. A California Highway Patrol computer aided dispatch message. Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

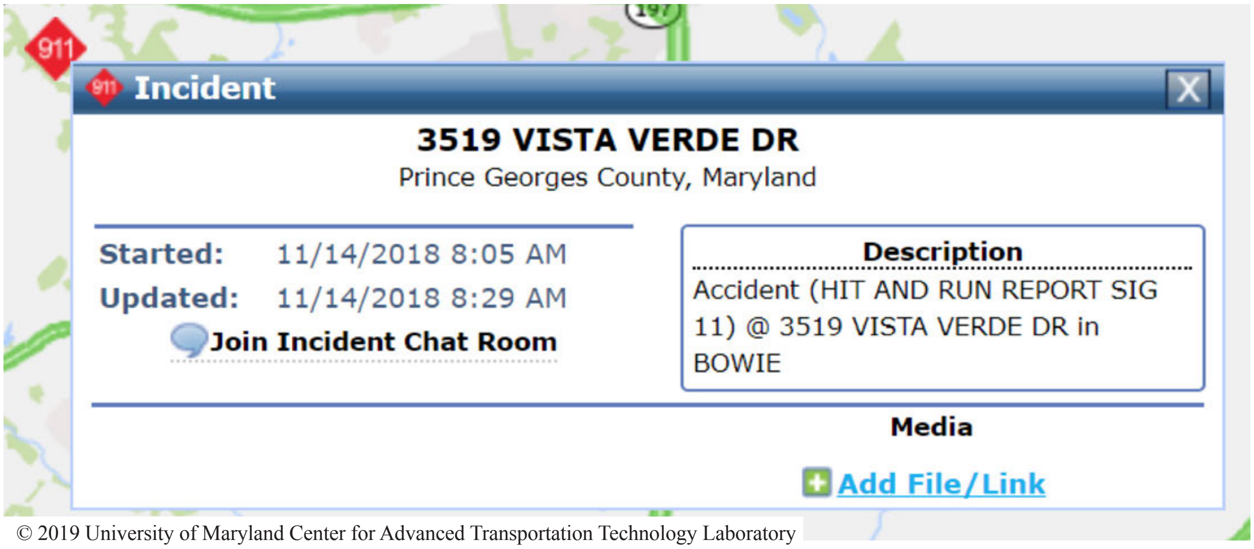



**Figure 52. Screenshot. Prince George's County Maryland computer-aided dispatch system messages.**

**Source: University of Maryland Center for Advanced Transportation Technology Laboratory.**

## Pros and Cons

The most valuable contribution of CAD data to a TMC is in the form of faster incident detection and more effective response. TMCs can dispatch appropriate resources more quickly and clear the scene more effectively if they have a better understanding of the type of incident and its severity.

In addition to improvements in incident clearance time, automated information exchange reduces the burden of communication workload and increases situational awareness.

However, CAD data comes with many challenges. Most of the time, these systems were not developed to support transportation operations. This means that when adapted to share information with TMCs, there are several challenges:

- Data sensitivity/security:
    - CAD systems collect data for all calls, not just transportation incident calls. This means that feeds must be filtered to exclude non-transportation calls, such as criminal activity types of calls that may include personally identifiable information.
- Data structure:
    - CAD systems are developed for call takers who must interpret incoming information and quickly input it into the system. This means that many CAD systems rely on abbreviations and shorthand as well as free text entry.
    - Free text entry, while readable to a human, can be very challenging to process in an automated manner. This means that systems may not be able to automatically determine incident type and extract relevant information accurately. Simple misspellings may invalidate data and be either ignored or thrown away before they even reach a TMC.

- Information overload:
  - TMC operators already deal with a large amount of information, both incoming and outgoing. They monitor CCTV, radio, and ITS outputs to determine the status of the transportation system. Sometimes they also deal with phone calls and customer inquiries.
  - The addition of CAD data may present a challenge, especially in cases where TMC is already aware of an incident and the system is unable to intelligently handle duplicates or related incidents, or in cases where CAD data volume is so high that the operators cannot keep up with inflow and triage it appropriately.
  - Some TMC systems use CAD data to generate alerts for operators when relevant incidents arrive via CAD. However, CAD design, free text abbreviations, and shorthand can often trigger false-positives, which eventually results in operators ignoring the incoming information.

## Use Cases for CAD Data

**Use Case: California Highway Patrol Computer Aided Dispatch Data for Public Consumption.** CHP is generally the first to become aware of incidents on major California highways through 911 calls. In an effort to keep the general public informed about congestion in major metropolitan areas, media outlets always look for the fastest and most reliable source of incident information. This resulted in CHP receiving a high volume of calls from media, and even their DOT partners, requesting information about current conditions and incidents. CHP found that the amount of time and effort expended on responding to these calls and requests was significant and was impacting the agency's ability to accomplish its primary mission of responding to and clearing incidents.

CHP took a bold step to solve this problem by creating an XML data feed of all their transportation-related CAD data, except confidential information, and publishing it for public consumption. These data included detailed text about each and every publicly viewable incident. This means that anyone interested could obtain access to data and do whatever they wanted with it. Initially, some feared that this would result in major liability issues. Instead, the incoming call volume to CHP reduced significantly and resulted in lower cost and better focus on incident response.

While the information in the XML feed was detailed, much of it was in the form of free text fields, making automated integration somewhat difficult. Nevertheless, availability of this information is invaluable for media, the public, and DOT partners looking for better situational awareness.

**Use Case: Virginia Department of Transportation Realtime Traffic Incident Management Information System.** VDOT was interested in improving safety and mobility by sharing existing information currently available to VDOT with its partner public safety agencies. To break down information and communication silos, VDOT worked with their partners to develop a realtime data sharing system, illustrated in figure 54, using automated data extraction and filtering that limited the impact to operators' workload.

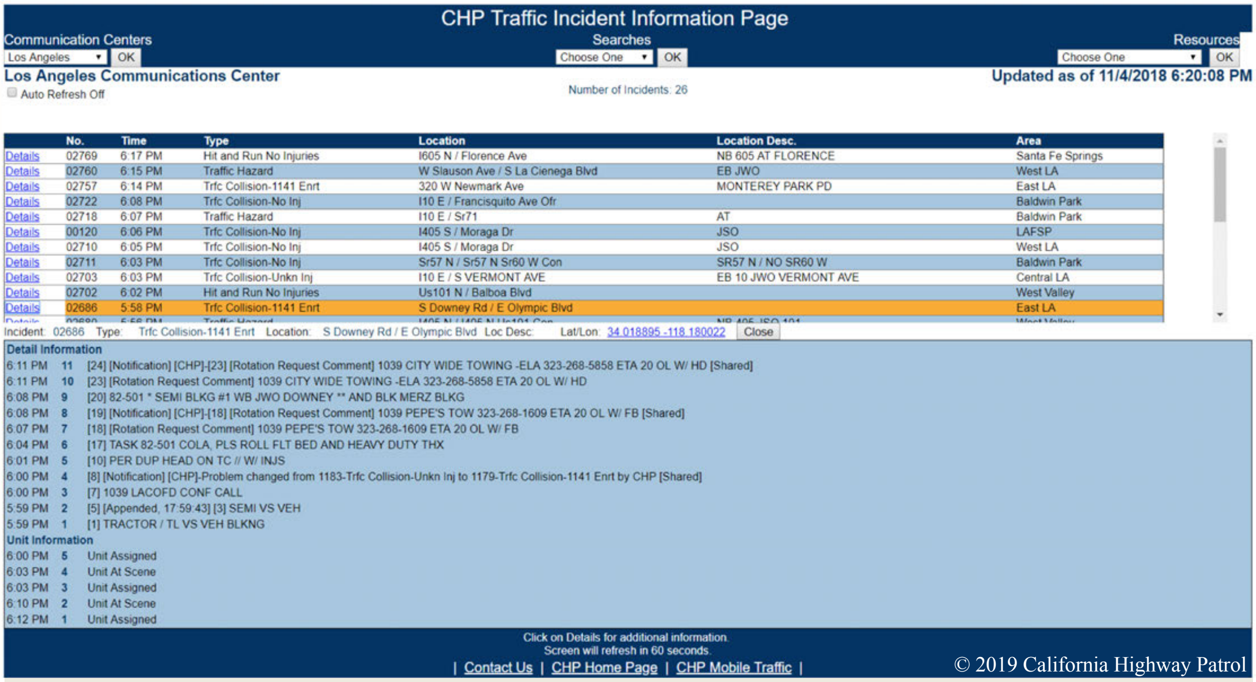



**Figure 53. Screenshot. California Highway Patrol computer-aided dispatch log example.**
**Source: California Highway Patrol public data feed (http://cad.chp.ca.gov/Traffic.aspx).**

The Realtime Traffic Incident Management Information System (RTIMIS) collects data from the CAD system and transmits it via secure connection to the integration partners, including VDOT TMC operators. This integration provided a 34-percent reduction in clearance time across 67 miles of I-95 by making TMC operators aware of incidents more quickly than before.[15] In addition to clearance time reduction, RTIMIS CAD integration reduced communication workload as well as improved general situational awareness.

VDOT found that CAD and TMC operators often have different goals, which impacts the quality of data being exchanged. For example, information that TMC operators may consider critical may not be important to CAD operators, so they may not enter it as frequently. Similarly, CAD systems use free text fields and shorthand that makes it very difficult to automate integration of valuable CAD data.

**Use Case: Florida Highway Patrol and SunGuide Integration.** The FDOT ATMS platform, SunGuide, receives a filtered realtime CAD data feed. Through this integration, FDOT has seen a reduction in incident verification, response, management, and clearance time. FDOT operators become aware of incidents more quickly using CAD data and can evaluate incident severity and impact and dispatch appropriate field units and towing units more quickly. In addition to quantifiable improvements in incident response, the Florida Highway Patrol dispatchers have experienced a reduction in their workload as they do not have to take extra steps to notify their FDOT counterparts when requiring traffic incident management support.

15 Virginia Department of Transportation, Realtime Traffic Incident Management Information System Presentation, April 23 – 24, 2018. Available at: http://i95coalition.org/wp-content/uploads/2018/06/3-I95CC-CAD-Workshop-Virginia-Presentation-Apr2018.pdf?x70560.

FDOT CAD integration is read-only with operators having the ability to dismiss alerts, create new events based on the alert, or associate an existing event to an incoming CAD event.

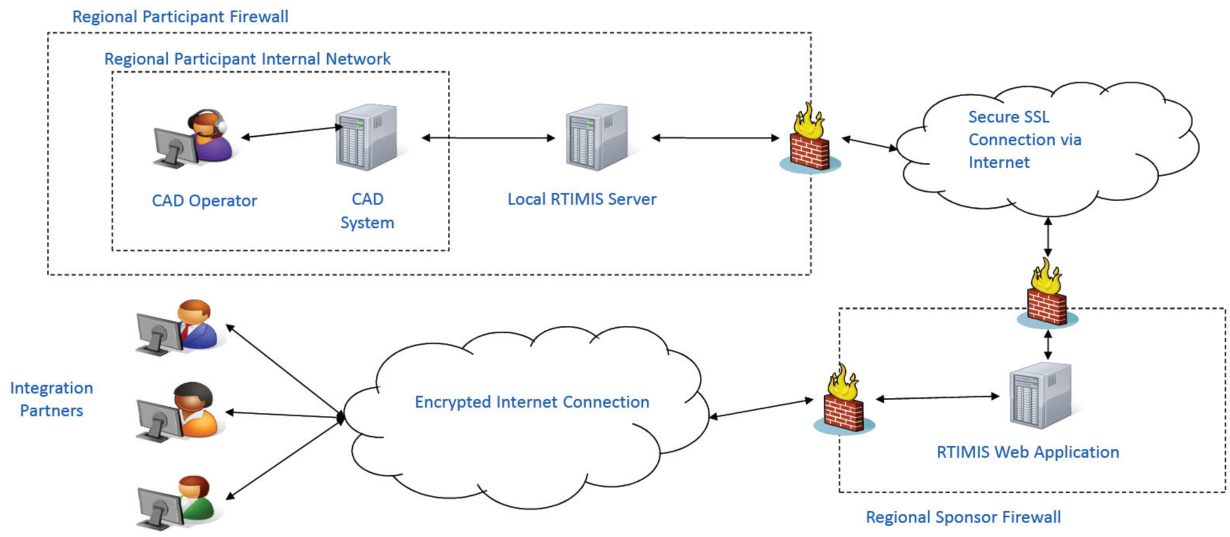





**Figure 54. Virginia Department of Transportation Realtime Traffic Incident Management Information System high-level architecture diagram.**

**Source: Virginia Department of Transportation, Realtime Traffic Incident Management Information System Presentation, April 23 – 24, 2018. Available at: http://i95coalition.org/wp-content/uploads/2018/06/3-I95CC-CAD-Workshop-Virginia-Presentation-Apr2018.pdf?x70560.**

# CHAPTER 3. CURRENT AND EMERGING PRIVATE-SECTOR BUSINESS MODELS

Private sector providers currently use the business models and delivery mechanisms displayed in table 13. Each one of these business models can have sweeping impacts on pricing, contracting, and agency use. These business models and delivery mechanisms are not mutually exclusive. For example, blending "perpetual ownership" with a business model that includes "archive only" means that the agency would have perpetual ownership of archived data, but not necessarily access to realtime data.

Business models determine how revenue is generated or how the private sector generates value from their product. Data providers may sell their products directly to agencies or to a third-party vendor that then sells the data to an agency. Private sector vendors may instead receive payment through advertisers.

**Table 13. Industry data procurement and sharing business models and delivery mechanisms.**

| Model/Delivery Mechanism | Explanation |
|---|---|
| Direct Sales to Agency | Private sector vendor sells data directly to agency. Vendor can usually offer the lowest possible pricing for a given data set. |
| Third-Party Resellers | Data provider sells data to an intermediary company who then sells it to end customers. When a third party sells data to an agency, it is difficult to know the extent of additional markup or overhead. Both the originating data provider and the third-party reseller must remain profitable. |
| Advertising | In rare cases, agencies will receive data or services for free (or at a reduced rate) because the private sector gains revenue on the products through advertising services. Agency-branded websites (like 511 systems) usually provide these services. Many of these systems are falling out of favor. It is growing more complicated to sell advertising on agency websites that are falling in popularity—primarily due to the rise in popularity of newer navigation apps on smartphones. |
| Integration with Ancillary Products (e.g., Visualization Tools) | Adding additional services and products to data can make it more valuable, and can make the sale easier to digest by agencies. For example, a department of transportation may procure a new advanced traffic management system or 511 system from the private sector, and as a requirement of the request for proposal, may bundle third-party data into the contract. Similarly, a data provider may bundle data, analytics on the data, and other ancillary products together. Some data providers will offer realtime analytics, visualization tools, tiling services, and more. Each one may not have enough value by itself, but when bundled, they have greater earning potential. |

**Table 13. Industry data procurement and sharing business models and delivery mechanisms. (continued)**

| Model/Delivery Mechanism | Explanation |
|---|---|
| Trade and Promote | Some providers are less interested in revenue directly from the agency, and are more interested in integrating received agency data back into private sector systems resold to others. In addition to providing private sector data to the agency in exchange for access to agency data, the private sector may also require that the agency prominently promote the private sector's products and provide significant, visible attribution. |
| Shared Revenue | The agency agrees to one or more revenue generation models (like advertising on agency-owned websites), but shares or outright relinquishes the title to any revenue generated from such advertisements. |
| Realtime Use Only | In this model, agencies use the data for realtime purposes only. The agency cannot archive the data for later use. |
| Single-Purpose Use | Single-purpose licensing may allow an agency to purchase the data for use in only one, usually very limited, application area. For example, licensing a probe speed data set for use in identifying potential incident locations would disallow the agency from using the data for congestion monitoring or performance management. |
| Archive Only | While it may not seem applicable, archived data can be incredibly useful for realtime operations. Newer decision support systems that leverage modeling and simulation or machine learning can benefit from an archived dataset. Archived data can train systems, help in prediction, and more. The private sector may sell access to archives of historic data at a lower cost than realtime data. |
| Limited Geography and/or Time Ranges | Even though data may be available for the entirety of the State's network, vendors will nearly always charge less for access to data that is limited in geographic scope—perhaps only to select corridors and/or only during certain times of day. |
| Leased Equipment, Dual Use | Roadside equipment manufactures install and maintain the equipment at some cost to the agency and allow the agency to use the data for one or more purposes. However, the owner of the equipment may be entitled to resell the data to other third parties. |
| Perpetual Ownership | The agency buys all rights and the title to the data in perpetuity. Even after the contract with the data provider is complete, the agency still owns the data and can use it for other purposes, but cannot resell the data to others. |
| Single-Agency Use Only | The agency can only use the data for itself. Partner agencies, like metropolitan planning organizations, research institutions, and neighboring departments of transportation, cannot use the data. |
| Hosted and Aggregated | Only the vendor can store the data in the cloud. Aggregated data is available via application program interfaces, websites/visualizations, or other methods. The agency or application does not get raw data. The agency is otherwise unable to integrate the data into native systems. |

To illustrate how these different business models affect cost and acceptable use, consider the following agency use case in table 14. This compares the procurements of probe-based speed data from HERE Technologies by two separate departments of transportation (DOT). Both agencies are purchasing identical data from the exact same data provider. However, each agency took a separate approach to procurement and business models.

**Table 14. Comparison of Florida and North Carolina procurement strategies for probe-based speed and travel time data.**

| | Florida Department of Transportation (DOT) | North Carolina DOT |
|---|---|---|
| **Purchased Data** | Realtime probe data (speeds and travel times). | Realtime probe data (speeds and travel times). |
| **Business Model and Contracting Approach** | Procurement directly by the State of Florida. No provisions in the contract for sharing with partner agencies outside of the State. | Procurement through the I-95 Corridor Coalition's Multi-state Vehicle Probe Data Project. Data sharing agreement requires data to be available to any agency in the I-95 Corridor Coalition (not just the department of transportation). No limits on non-commercial use. Perpetual ownership. |
| **Impacts to Business** | Limits ability to share information across borders (or see data from neighboring States). Disallows multistate performance monitoring, cross-border congestion analytics, or cross-border traveler information services. | No information found. |
| **Cost Implications** | Lower cost. | No information found. |

Another illustration, provided in table 15, is for two agencies that procured O-D data and services from two separate vendors. Each vendor has unique business models that drastically affected cost and capabilities.

**Table 15. Comparison of origin-destination data procurement strategies by two agencies.**

| | New York City Department of Transportation (DOT) | Virginia DOT |
|---|---|---|
| **Purchased Data** | Origin-destination (O-D) data (including breadcrumb trails) from INRIX. | O-D data through Streetlight. |

**Table 15. Comparison of origin-destination data procurement strategies by two agencies. (continued)**

| | New York City Department of Transportation (DOT) | Virginia DOT |
|---|---|---|
| **Business Model and Contracting Approach** | Provides raw O-D data (including breadcrumb trails) to the agency for unrestricted use. Perpetual ownership. | Single-use data license. Data provided for a specific study/study area. Data cannot be stored and re-used for other studies. Data provided through an on-line application in aggregated forms along with reports and visualizations. |
| **Impacts to Business** | Because of perpetual licensing and unrestricted use, the City owns the data for use in any study and can share the data with partners. Because it is raw data, it is up to the agency to decide *how* to use it. Analytical tools are not provided. | Use was restricted to a specific study for a specific date range. Data cannot be used for other studies. Because this was project-specific, procurement was easier and cost was significantly lower. Because an analytics platform offered the data, the data was initially easier to use and interpret than raw data would have been. |
| **Cost Implications** | Higher initial cost, but potential higher value. | Lower initial cost. Conducting a significant number of studies over many date ranges could increase the cost beyond that of the raw data purchase. |

Data providers offer *many* different business models when selling their data and services. Most will also lower the cost of their data if it is restricted to single-purpose/single-study use. The examples above are more indicative of agency procurement strategies and how the business model chosen impacts long-term capabilities and costs.

## PRICING STRUCTURES

The pricing structure of data varies significantly between businesses and can depend on the procurement method and type of data or service provided. Examples include:

- Pricing data based on centerline miles.
- Pricing data based on population of the covered region.
- Pricing based on the number of calls (or requests) to application programming interfaces (APIs).
- Pricing based on number of users.
- Pricing based on number of studies conducted with the data.
- Discounts for bundling of data or services.

These models vary. Some data providers may charge based on mileage of the National Highway System only and then offer arterials at lower prices. It is difficult to find published, stable price sheets on data products from the private sector due to several factors:

- **Costs change over time:** Commoditization can result in some data products losing value. Pricing sheets today would become obsolete in 1 year or less. Other data sets may grow in value over time. Brand new data products may be difficult to sell as convincing an agency of the value of the data can be difficult. The private sector may initially offer the new data products at substantially reduced rates to get the public sector interested in them. The price can change after determining the actual value.
- **Persistence:** Vendors do not want old pricing sheets circulating through agencies as they may be misinterpreted or considered current when pricing has changed.
- **Competition:** In the low-bid marketplace, pennies can mean the difference between winning and losing a contract. The private sector is wary of the competition learning about their pricing for fear that their competition will use that against them.
- **Third-party markup:** When selling data or services through a third-party or bundled in procurements with other contracts, the prime consultant may charge additional overhead, markup, or profit on the data and/or services.
- **Requests for proposal and low-bid contracts:** Low-bid contracts accepted by agencies can disrupt marketplaces and result in constantly changing price sheets. This is because low-bid contracts make it difficult for data providers to win contracts based solely on the quality of their data or services and force providers to compete on price alone. This has many other negative side effects as well, including the reduction in innovation and reinvestments that providers may make. If a provider knows it will largely be judged on price alone, there is less incentive to propose an innovative and leading-edge solution.

## REVENUE SHARING WITH AGENCIES

The private sector can view shared revenues from data sharing as a marketing strategy rather than a viable business model. Many agencies have experimented with exclusively sharing speed, volume, and incident data in exchange for traveler information services to the public through 511 and other means. Revenue can be acquired through sponsorship of 511 systems, and while revenues that go beyond operating costs are supposed to be sent back to the State, all too often this income either falls short of predictions or any small amounts of revenue gained are reinvested in the system. Many revenue sharing schemes surrounding data sharing still end up costing the agency as much as (if not more) than traditional contracts.

Parking, tolling, and automated enforcement (speeding or red light running) are the notable exceptions to revenue sharing success. However, the business models that make these relationships a success have less to do with data sharing and more to do with contracts that enable the private sector to operate (or install) parking, tolling, and enforcement infrastructure for the agency. Sometimes the agency pays for these services directly, and other times the concessionaire pays for the entire infrastructure and shares a percentage of the revenue collected with the State.

## EXCLUSIVE ACCESS

A few agencies have entered into exclusivity arrangements with the private sector. For example, the private sector may offer to redistribute an agency's closed-circuit television (CCTV) video to partner agencies and the media at no cost to the agency. However, in return, the private sector will request exclusive access to the video for sharing with the media. The private sector then has exclusive rights to essentially re-sell the video to television stations, third-party traveler information providers, or other non-public agencies.

## THE VALUE OF AGENCY DATA

Most private sector data providers still have a desire for agency-produced datasets; however, the type of data deemed most valuable has changed. Table 16 shows examples of data types and their perceived values as considered by the private sector.

**Table 16. Agency data value as described by private sector companies.**

| Agency Data Type | Value Narrative |
|---|---|
| Volume and Speed Sensor Data | Used to be valued by traveler information providers, the media, and third-party probe data. However, the growing penetration rate of probes (cell phones and Global Positioning System (GPS) data) has made this agency data unattractive. |
| Planned Event or Lane Closures Data | Still valued by media, third-party data providers/aggregators, and others as the agency is essentially the only source of this data. |
| Agency Advanced Traffic Management System/511 Incident Data | With the improvements in crowdsourced data collection, this dataset is slowly losing value, but many companies want it since agency-provided incident data often has information that goes beyond the basics of location and incident types. Agencies that provide more detailed information about which lanes are closed and which responders are on the scene are viewed has having greater value. |
| Computer Aided Dispatch Data | Considered the "gold standard" by both agencies and the private sector alike, not for the completeness of the data, but rather for the speed and coverage that are provided. The private sector still considers this data valuable even though it can be difficult to interpret. |
| Driving Records/Division of Motor Vehicles Data | Insurance companies find this valuable. |
| Police Crash Reports | Insurance companies find this valuable. |
| Live Closed-Circuit Television Feeds | Valued by the media and third-party traveler information providers. Studies have shown that the public chooses its local television and radio news programs based off the quality of its weather, traffic, and sports coverage (in that order). Broadcasters can raise advertising rates when higher quality traffic video is available. Some broadcasters and third parties even install and sell their own camera feeds. |
| Parking Availability | Traveler information providers, vehicle manufacturers, and parking services value this data. While many private companies are the main providers of this data, there are cases where agencies own or manage this data and can provide it to the private sector for use. |

## USE CASES OF VALUE EXCHANGE FOR AGENCY DATA

As part of their connected citizens programs, public agencies are swapping data on a large scale with Waze; that is, no money is directly passing from the agency to Waze. Waze is one of the few private sector companies offering to swap data with public sector agencies. Waze offers agencies access to their crowdsourced incident data in exchange for access to agency incident data, attribution, and co-marketing of the Waze platform. While the Waze acceptable-use terms are evolving, some agencies are seeing restrictive data sharing agreements from Waze, which is limiting the value of the data for anything beyond realtime operations.

Several State DOTs have tried to generate revenue through marketing and selling CCTV video feeds as a subscription-based service. Several States have also arranged for in-kind services, which seem to have had more success than the subscription-based services. Some years ago, the Florida Department of Transportation (FDOT) had evaluated efforts to sell CCTV camera feeds to media partners based on a subscription service model dependent on the size of the market in question. For example, a larger market like Orlando would generate more revenue and be more desirable than the smaller Jacksonville or Tallahassee markets. FDOT ultimately decided against pursuing this effort, although it continues to provide CCTV images to media partners through written use agreements.

FDOT recently entered into a co-sharing and license agreement with Harris Corporation. Harris has developed the Helios platform, which enables traffic and other cameras to function as weather sensors providing better situational awareness. In an exchange of data, FDOT will grant Harris access to CCTV video images. In turn, Harris will allow FDOT to access the Helios web service. Harris also provides a web-based API in order to retrieve the Helios Analytic Observations or the weather condition observations. The agreement stipulates that API requests cannot exceed 100,000 per month. The weather condition observations gathered will be integrated into the FDOT FL511 site and provided to the regional transportation management centers (TMCs). As an additional contractual term, FDOT must provide credit to Harris for the use of Helios analytics. Neither the FDOT nor Harris Corporation is charged any fees under this agreement.

The State of California makes its CCTV cameras available to the public through the use of an application that is downloadable from their website. Similarly, the State of Oregon provides its CCTV feeds free of charge, making the data as accessible as possible. Media outlets and interested parties pay for any connection necessary to access the information.

The Utah Department of Transportation (UDOT) has a non-exclusive agreement with Total Traffic & Weather Network (Total Traffic) for traveler information dissemination in exchange for value in-kind services. UDOT grants Total Traffic access to the traffic operations center (TOC) for routine radio traffic reports from 5 a.m. to 11 p.m. Total Traffic provides and maintains, at its own expense, equipment and communications infrastructure necessary for the operation of its travel information dissemination activities. The equipment and communications infrastructure used by Total Traffic must not interfere technically or operationally with the UDOT system at any time.

During the radio traffic report broadcast, Total Traffic must mention the UDOT TOC or use any other agreed-upon language providing attribution to UDOT. As part of this agreement, UDOT also receives a minimum of 20 30-second commercial announcements per Clear Channel station using the traffic reporting services per month to advertise CommuterLink™ activities. In addition, all Clear Channel stations receiving the traffic reporting service and that have websites must provide a legible website link to Utah's CommuterLink™ website at no charge to UDOT. According to the agreement, UDOT does not guarantee continuity of access, nor the accuracy, timeliness, or quality of information. Information is provided on an "as is" and "with all faults" basis. Any reliance on said information is at the risk of the company. UDOT also accepts no liability for system down time.

In an agreement where the New Jersey Turnpike Authority administers the contract for both the Turnpike and the New Jersey DOT, the Authority receives $30,000 in revenue per year from TrafficLand based on the current number of cameras configured. On a per-unit basis, additional revenues may be provided should the New Jersey Turnpike Authority and the New Jersey DOT increase their number of configured cameras. The Agreement for Redistribution of Streaming Traffic Video indicates compensation in the amount of $2,500.00 per month for all cameras available to TrafficLand with additional compensation of $50 per month for each group of 50 cameras added to the original quantity.

The Michigan DOT has a video sharing contract with Skyline Technology Solutions (Skyline) where the company processes and hosts all of the Michigan DOT's traffic video and provides traffic video feeds for the Michigan DOT website at no charge. The contract does not restrict video redistribution by Skyline at a cost to external customers. Skyline provides a hosted site where MDOT's trusted partners (dispatch, county, police, and other TMCs) can access the video for free. This arrangement helps Michigan DOT avoid costs associated with purchasing and maintaining servers to host and distribute video to partners and the media.

## AGENCY DATA VALUATION

Historically, agencies have found that the cost of managing even the most basic cost-recovery business models were overwhelming and not worth the effort. Most agencies are not equipped to receive funds—even in small amounts. Both legal and ethical issues have proved too great for the relatively modest returns.

However, if an agency does want to determine the value of a particular data asset, the agency should assess the following:

- **Identify the opportunity:** What is unique about the data or asset that makes it valuable? Who desires the data? Have companies requested the agency's data already, or will the agency seek out potential buyers? How many potential customers are there?
- **Assess your uniqueness:** Does the agency have exclusive rights to its data? Can anyone else offer an equivalent product?
- **Assess the market:** Data for dense, urban areas with multiple, competing media platforms are likely more desirable and would provide greater competitive advantage. Do not expect data assets in rural parts of a State to be considered as valuable as those in congested urban regions.
- **Tiers of service:** Will the agency offer an expected level of service and/or reliability for its data, or will it simply provide the data as-is with no warranty or expectations of availability? Will the agency offer different product types or combinations, e.g., silver, gold, and platinum data products?

- **Evaluate your costs:** Will providing data to the private sector cost the agency anything? Potential costs could include:
  - IT/networking costs.
  - 24 hours a day/7 days a week operations and maintenance support.
  - 24 hours a day/7 days a week technical/user support.
  - Contracting/management/legal.
  - Hardening/productizing of your data product.
  - Periodic system upgrades.
  - Insurance/liability.
- **Profit or cost-recovery:** Does the agency simply want to recover the cost of making the data available, or is it looking to bring in funds to support other efforts?
- **Test the market:** Before officially releasing a cost model, it is important to test the model with one or two private sector partners to gauge their reaction to the pricing and terms. Face-to-face conversations are preferred to help in determining the private sector's appetite for a proposal. Be ready with alternative business models as well as both floor and ceiling numbers.

It is important to think about management costs, contracting costs, and legal fees. Even if an agency does not officially charge its time to specific projects, there is a real cost to taking on additional efforts and contracting. That cost could be purely financial, or it could be realized when other work or contracts get dropped or bumped due to a lack of available resources.

It is also critical to evaluate the liability cost associated with charging for data: Does the agency stand to lose anything by providing the data (either free or at some cost)? Are there philosophical concerns with limiting access to the data to those who can afford to buy it? Would the agency be stifling innovation in restricting access?

Alternatively, does the agency have anything to gain by sharing/selling the data? Positive public perception and allies in the private sector can sometimes be more valuable than the cash or cost of the transaction.

# CHAPTER 4. PROCUREMENT STRATEGIES

When procuring data or services, agencies need to exert considerable effort to ensure proper communication of the agency's needs to bidders. When developing requirements for data, services, or both, two approaches are typically used:

- Issuing a detailed set of system requirements to identify the functions and top-level design of the system. This approach requires a major effort on the part of the agency and can become expensive and frustrating—especially if new requirements are identified midway through the initiative or the requirements were not thought through entirely.
- Developing the requirements as part of the contract prior to beginning. This method can allow for greater dialogue between the bidders and the agency, which ultimately can lead to a better product.

Both of the options have their advantages; however, the second approach typically results in an end-state that is more favorable to the agency.

Well-written and well-defined requests for proposal (RFPs) state the agency's preferred business model, data use rights, and delivery mechanisms. Agencies that leave these out of their procurements open themselves up to potential misunderstandings as to what is delivered and what will be allowable. The private sector may try to propose alternative business models in procurements, but that opens the agency and private sector to risk.

The requirements for the use of specific technologies or techniques to collect and deliver data, however, should be left out of RFPs. Development of new technologies, methodologies, and tools happens both quickly and often, and requiring outdated technologies can result in artificial limits being placed on the agency and the consultant as they work to perform analytical tasks. It is generally better practice to allow data or service providers to drive these decisions based on what they perceive to be the most efficient and effective tools and methods at the time of delivery.

## SOLE SOURCE

Agencies that procure new data or services via sole-source procurements usually do so because they have thoroughly defined needs, have evaluated what is available from the private sector, and have found only one potential provider that can meet all of those needs. Sole-source justification is a good procurement strategy when and if an agency is fully aware of all of the alternative solutions and data providers and there is a clear advantage of one provider's data over another. Sole-source procurements are often appropriate for truly emerging data sets because it is unlikely that multiple sources exist for the data.

Less competition for sole-source procurements does not necessarily result in higher prices for an agency. Sole-source procurements are significantly less effort for the private sector—thus lowering overall costs and overhead associated with developing proposals. The private sector is often more open to negotiating favorable terms for an agency with sole-source procurements.

Sole-source procurements present problems when an agency does not truly understand its own needs or what may or may not be available from other data providers.

## TRADITIONAL REQUESTS FOR PROPOSALS

Traditional RFPs are usually challenging for both the agency trying to procure data and the private sector. For the public sector, RFPs can take a very long time to let and award. Technology and data offerings can change significantly in a short period of time, and what is documented and specified in the early stages of RFP development could be significantly out of date and even irrelevant by the time the RFP is awarded.

It is also very difficult to write a good RFP. Agencies face the following challenges:

- **Being overly specific:** Some agencies are far too specific and demanding in their RFP language—specifically surrounding technical requirements. Agencies that list a large number of technical requirements in an RFP risk inadvertently raising costs, confusing the private sector, or receiving complicated bids from the private sector. In addition, not every data provider can provide the exact same services. Often an agency will merge requirements from multiple data vendor systems. This can lead to a complicated RFP that no single vendor can truly meet. The agency may then select a vendor that claims they can meet all of the requirements, but it really cannot, leaving the agency feeling duped.
- **Being too vague:** Being too ambiguous with requirements and needs can be another risk. Agencies may poorly define their requirements, and then receive many bids from unqualified candidates that do not truly understand the agency's needs and think they have a solution that might fit the poorly crafted scope of work. This leaves the agency reviewing more responses than needed, and it makes it very difficult for reviewers to judge appropriately which company should be given an award.
- **Being influenced by the private sector:** While not always a bad thing, RFPs can be influenced by the private sector. This typically happens during the RFP conceptualization and drafting period—well before the RFP is let. Agencies can use contractor suggestions and inputs to their advantage, crafting an RFP that will receive a reasonable response from the private sector; however, when RFP development occurs without consideration of multiple providers' abilities, this can lead to unfair bias or even favoritism, which in turn results in an RFP that some competitors may perceive as being "written for" a specific vendor based on their known strengths. It is far better to have the agency meet with all potential data providers multiple times before drafting the RFP. However, the agency cannot simply combine each private sector data offerings into the requirements because this can create a list of requirements that no one can truly satisfy, as mentioned above.
- **Low-bid States:** Many agencies exist in low-bid environments—meaning that they must always select the lowest bid response. While there are provisions in most States to reject bids that cannot meet all of the requirements of the RFP, it is rare that reviewers will outright reject a response partly because companies are good at writing responses that seem to meet all of the requirements of the RFP—when in reality, they might not.

## INTERGOVERNMENTAL AGREEMENTS

Intergovernmental agreements (IGAs) are useful when an agency needs to obtain the data or the services of a public-sector university, wants to partner with another department of transportation (DOT) in another State, or desires to work with a quasi-governmental entity. The risks associated with IGAs are similar to those of sole-source agreements; namely, the agency must be sure that it is working with a provider that will meet all of its needs before going down this path. IGAs can also produce results at a lower cost because they are typically leveraging a resource that has already been vetted and procured or is being purchased in bulk.

## POOLED FUND STUDY

Pooled Fund Studies (PFSs) are excellent ways to leverage the resources of multiple agencies to fund data and services from one or more providers. Agencies "pool" their resources (funds) to develop a scope of work and procure support. A single agency can administer PFSs on behalf of everyone else. While there is an increased level of burden on this lead State, the benefits are typically significant and far exceed the administrative costs. The work or data procured by a PFS can be sole-sourced or put out for bid.

Several benefits of PFSs include:

- **Scope Synergy:** Because multiple agencies work together to define the scope of work/services, there is a greater likelihood that the scope will ultimately be more complete and thoughtful, thus leading to a better product.
- **Lower Cost:** Almost every data and/or service provider will lower prices when products are purchased in bulk. Therefore, each participating agency will likely pay less for the data/service that it wants to purchase.
- **Scalability:** PFSs can grow in size over time. For example, if three agencies initially form a PFS that becomes successful, other agencies can later join the PFS—gaining from the wisdom and investment of the first three agencies and potentially further lowering costs. While there are many examples of this, the current American Association of State Highway and Transportation Officials (AASHTO) Transportation Management Center (TMC) Pooled Fund Study is a recent example of TMCs getting access to additional data and analytical services surrounding the National Performance Measures Research Data Set for reporting under the Moving Ahead for Progress in the 21st Century (MAP-21) third performance measure rule, PM3.

## ON-CALL CONSULTANT AS THE PROCURER

Several agencies have leveraged their on-call consultant contracts to procure data or services for the agency. These types of purchases sometimes appear to circumvent the procurement process entirely; however, some agencies have effectively used these contracts in a reasonable and ethical way. For example, the City of Austin, Texas, recently had one of their on-call consultants procure realtime probe data and data analytics services on the city's behalf. The consultant put together their own RFP from which multiple private sector entities responded. The consultant then reviewed the responses in consultation with the DOT and chose a winner. The city essentially outsourced the entire procurement process to their consultant.

## PROCURED AS PART OF A LARGER SYSTEM

There is a growing trend among some agencies to bundle data, analytics, and other services with the procurement of their advanced traffic management system (ATMS) or 511 platforms. While this makes sense if the DOT already knows exactly what data it needs and requests very specific data from a specific vendor in their RFP, the risks include:

- **Overhead:** Most consultants need to markup the cost of data—to cover management, overhead, contract administration, legal, etc. An agency will need to evaluate if the cost implications of procuring within a bundled contract outweigh the cost of an independent procurement or purchasing strategy.
- **Mixed solutions:** If the agency does not specify exactly which vendor, data, or service they want within their ATMS/511 procurement, then the agency risks receiving bids from ATMS vendors that choose data providers on lowest cost, strategic and/or existing business relationships. The agency must then choose to make an award to a great ATMS platform with sub-optimal data or an inadequate ATMS platform with potentially great third-party data.
- **Locked-in:** Bundling data purchases with an ATMS purchase means that the agency can get locked into the data provider for the life of the ATMS contract—limiting the agency's ability to change providers without broader contract implications.

# CHAPTER 5. **POLICY CONSIDERATIONS**

This section covers some of the many policy considerations that will impact departments of transportation (DOT) when considering a data source, business model, or procurement strategy.

## ACCEPTABLE USE TERMS AND CONDITIONS

The data-use agreement is a cornerstone of any public-private data partnership. This agreement governs several key components of the partnership including, but not limited to:

- Who can use the data within the agency?
- Who can data be shared with and in what format – raw, aggregated, anonymized, etc.?
- What, if any, attribution is required when publishing data and data products?
- What type of processing and publications are allowed – research, reports, public traveler information, operational use only, etc.?
- What is the expected level of accuracy and reliability of data?
- How is the data validated?

Acceptable use terms and conditions are new to many agencies. Traditionally, agencies procured and deployed their own infrastructure and received data as the byproduct of that deployment. Because of this, agencies had full control over data: they could store it or share it with other agencies, universities, or the public; use it for realtime operations, planning, or studies; or dispose of it if they deemed it not useful or worth the cost of storage and management. In fact, many agencies were often unaware of all the data they were collecting. As technology advanced, the value of data became clearer as it supported information-driven decisionmaking. The private sector seized the opportunity to generate and monetize transportation data to support agencies in their operations, planning, and performance management efforts.

As agencies began to work with the private sector data providers, they faced questions they never had to think about or answer. Those questions became key components of data partnerships that could either make or break the agency's ability to achieve its transportation systems management and operations (TSMO) mission. Early partnerships often resulted in agencies paying a high price for data that they could have generated with their own infrastructure, effectively paying a vendor to access the agencies' own data. Other times agencies obtained valuable data, but found that overly restrictive data use agreements prevented them from doing anything useful with that data.

Over time, agencies learned that a successful partnership required freedom and flexibility in data use on the agency side, while protecting the proprietary nature of a private partner's data sets. Model agreements ensure the following key components:

- Ability to utilize data for any purpose within the agency.
- Pay once, use many times.
- Share data with trusted partners or the public while protecting the proprietary nature of the data set.
- Tie contract payment to vendor performance.

Successful acceptable use terms and conditions are a result of a meaningful effort to understand agency internal needs across different departments, potential uses and interactions with partner agencies, private sector partners, academia, and the public. However, it is not a good idea to develop these types of data use agreements in a vacuum. Instead, agencies must consider other successful model agreements and engage their peers in discovering best practices, lessons learned, and common pitfalls. Finally, agencies must understand that data use agreements are a two-way street, and a path to a true partnership with private sector companies requires mutual benefit and respect.

## DATA VERIFICATION AND VALIDATION

No data is perfect. However, it is critical for agencies to understand what data is and is not capable of, and use it in an appropriate manner. To accomplish this, agencies must understand the problem they are trying to solve, and work with the private sector to understand how a data set may help solve that problem. Similarly, private-sector companies must manage expectations and provide data of satisfactory quality. This section discusses important aspects of verification and validation and how it can or should impact procurements.

A true data partnership includes a collaborative approach to problem solving. Agencies define problems they are attempting to solve and provide a set of goals they are attempting to achieve. Private sector partners evaluate the fitness of their data set in context of agency goals and develop data packages and data use agreements that provide the capability to solve those problems and reach the end goals.

Private sector data providers continue to innovate and improve their data packages. As part of that innovation agencies must consider potential uses of new data packages and work closely with their private-sector partners.

A benchmark for data performance is necessary for a two-way partnership to work. This benchmark may consist of many different components, including:

- Data delivery interval and definition of acceptable data gaps.
  - If data delivery happens once a minute to support realtime operations, how many missed intervals negatively affect realtime operations? How are those data gaps treated? Are they replaced by modeled or archived data?
- Geographic data coverage.
  - Does data consistently provide coverage of the agreed-upon network at the required spatial granularity?
- Data accuracy and responsiveness.
  - How much data is allowed to deviate from ground truth? Is there a sliding scale depending on the type of the road or environment? How quickly does data need to reflect changes occurring in the field?

For these components to be verified and validated correctly, it is important that a neutral third-party performs the verification and validation. While it is okay for the vendor to provide internal benchmarks, and for the agency to define expectations in a contract, neither party participating in the contract is in a position to fairly evaluate the performance of the data.

Impartial third parties that can perform verification and validation can be universities that have the capabilities and expertise to analyze data and provide objective evaluation. Alternatively, other private parties could perform this evaluation, but it is imperative that those private parties avoid conflicts of interest.

Validation efforts must be comprehensive, consistent, and timely. These efforts require data analysis expertise and the ability to not only process data and compare the results to benchmarks, but also understand the trends in data to ensure that data is consistent and reasonable. Given the critical nature of this data, it is important that validation processes are fast enough to identify any potential issues. This will allow the providers time to either correct them or for agencies to exercise their contractual options.

## PAYMENT TERMS

This section discusses how companies (and agencies) prefer to structure payment terms and conditions, and how those terms and conditions can affect the price of the data. It covers more innovative terms and conditions that incorporate payments based on the quality of the data.

Most public agencies are comfortable with the traditional infrastructure procurement process. For the most part, agencies procure technology or infrastructure through a bid process, install it in the field (either by in-house staff, vendor, or contractor), and maintain it over the life of the equipment or infrastructure.

Data procurement is different because a third party generates the data and agencies are exclusively a consumer of that data. In this arrangement, there are several approaches to procurement and payment terms.

### Raw Data Procurement

Agencies can work with third-party providers to purchase raw data. In this arrangement, agencies can purchase data in several different ways:

- **Realtime only:** Agencies obtain access to a realtime data feed that they can consume and store as they please.
- **Realtime and archived data:** In addition to obtaining a realtime data feed, agencies can purchase some amount of archived data (i.e., one or more years) delivered in a one-time data transfer.
- **Aggregated data:** In some cases, agencies may prefer to purchase data aggregated up to a certain level to support their existing operational capabilities. This is generally not an ideal approach as it reduces the agency's capabilities in the long term if they identify other uses for the data in the future.

The number of data points purchased often determines the payment terms for raw data procurement. For example, segment basis determines probe vehicle data marketing. Agencies can choose to purchase a certain number of segments, often covering certain geographic extent (State, county, and municipality) and/or network type (freeways, arterials, etc.).

In some instances, raw data procurement cost may be dependent on the number of users. For example, a State DOT may purchase data for use by a certain number of users at the DOT and for an additional cost, add partner metropolitan planning organization (MPO) users. This approach is less common as it is difficult to predict who would be using the data and when, and may be difficult for the third-party providers to enforce.

Purchasing raw data is a cost-effective approach when an agency has internal resources to process that data and make it useful. The cost of data in this arrangement may be low due to the competitive nature of the market.

### Data and Service Procurement

In addition to purchasing raw data, agencies may acquire additional data services. For example, they may purchase software-based dashboards, analytics, or decision support systems. This approach reduces the effort (and potential cost) for agencies to make data useful and actionable. The tradeoff is that the service may package data in a specific way that reduces flexibility when it comes to use of that data. For example, a dashboard may only be capable of displaying several realtime operational metrics; even if the agency has additional metrics they want to monitor that could be produced using third-party data. The good news is that, in a competitive market, third-party providers are often willing to work with agencies to develop additional capabilities and tools.

Another similar approach is through partnerships with other private companies or universities. For example, an agency may purchase raw data from the provider and a data analytics platform from a university partner who works with the provider to develop needed capabilities.

Purchasing raw data and associated services is generally more expensive than purchasing raw data alone, but may be more cost-effective as agencies often have major resource constraints or lack of specific expertise to process data. This has been one of the most common cost models over the last several years.

### Single-Use Data and/or Service Package

Raw data purchase often allows agencies perpetual access to purchased data, but in some instances, providers may offer data and service for a specific single use. This may be a particular study or analysis, or it could be to support a specific operational strategy implementation. In these cases, the provider only supplies the necessary data for this specific use, often aggregated, calculated, and packaged as a final result rather than as a raw data set.

These single-use data/service packages are not common, but have been a way for agencies to "test-drive" a new data set or service or to satisfy an immediate need at a relatively lower cost than if they procured the entire data set and analytics platform. However, the major drawback of this approach is that procurement satisfies a very narrow need and the agency does not have ability to leverage this investment elsewhere (i.e., for another department) or at a later time (i.e., for a similar study in the future). This approach is not sustainable or cost-effective in the long term.

### Bartering

In some instances, agencies can work with third-party data providers to obtain data in exchange for advertisement or attribution, or in exchange for agency data that may be of value to the private sector partner. This has been a popular model with crowdsourced data providers, such as Waze. Waze is capable of generating data on current conditions, but has no way of informing their customer of planned closures. In this arrangement, Waze partners with an agency to provide crowdsourced realtime data to support agency operations in exchange for receiving and displaying planned event and planned closure information. This approach does not involve any monetary exchange, but is mutually beneficial to both partners.

While this approach has some clear benefits, it does severely limit the agency's ability to control the quality, timeliness, or any other aspect of data delivery. There are no specific service-level agreements in place and partners are free to change their data interfaces as they see fit without agency approvals since there are no strict contracts or agreements in place. So far, this model has been exclusive to crowdsourced data and has not been prevalent among many third-party data providers.

### Data Quality

In most procurement arrangements (except with bartering), agencies have the ability to structure payments based on data quality and performance. As part of third-party provider contracts, agencies can tie their payments to the results of the validation process to ensure that the provider performs as expected. This allows protection for agencies throughout the contract while encouraging providers to continue to innovate and improve their data products to remain competitive.

## DATA MANAGEMENT AND MAINTENANCE

When it comes to data, a core asset for agencies today, agencies must consider policies surrounding short-term and long-term data management and data operations and maintenance (O&M). These policies include cost considerations, technical considerations, and agency information technology (IT) policies that might impact management.

One of the key decisions agencies must make in today's data management ecosystem is whether to store/manage data in-house or utilize one of the various hosted options. The implications of this choice can have long-term impacts on costs, control, and utility of data.

### In-House Versus Hosted Options

To store and process their data, agencies previously relied on in-house systems. With the emergence of third-party provided data, agencies are facing a choice of hosting data in-house or using a hosting service (with either the data provider, trusted third party partners host, or a commercial hosting solution). There are benefits and drawbacks for each approach.

Storing and processing data, at a minimum, requires computing resources, IT management resources, and data management experts (e.g., database administrators, analysts, and developers). Traditionally, most agencies have IT departments that provide IT infrastructure management and access to computing resources. Similarly, agencies have in-house transportation experts with varying levels of expertise when it comes to data analysis. While this type of approach worked well on traditional data sets that were smaller and easier to manage, currently produced data sets are significantly larger and more complex. Traditional tools, such as Microsoft Excel and small databases, are proving to be inadequate when it comes to management and processing of today's datasets.

Table 17 outlines general pros and cons of an in-house approach versus using hosted options for today's data set types and sizes.

**Table 17. Pros and cons of hosting options.**

| Hosting Option | Pros | Cons |
|---|---|---|
| **In-House** | ▪ Full control over data storage and processing infrastructure, strategies, algorithms, and data retention.<br>▪ Good understanding and tight control of budgets to support in-house storage and processing. | ▪ High cost of procuring the infrastructure, space, and expertise to manage data and processing.<br>▪ High investment needed in cybersecurity to ensure protection of infrastructure and data.<br>▪ Competitive workforce market makes it difficult to attract and retain talent.<br>▪ Data becomes obsolete due to inability to continue to innovate and stay current. |
| **Commercial Cloud Hosting** | ▪ Infrastructure management and cybersecurity become the responsibility of the cloud provider.<br>▪ Smaller initial investment since cost of storage in cloud is generally less than the cost of infrastructure and staff needed to store data in-house. | ▪ Lack of control over budget and expenses. While the cost of storage per byte may be low, the cost of data transmission may be difficult to estimate and control. Systems with a large number of transactions may end up driving cost significantly.<br>▪ Still requires in-house expertise. |
| **Trusted Partner Hosting** | ▪ Lower up-front investment.<br>▪ Lower risk due to stronger and more focused relationship between agency and the partner.<br>▪ Dedicated resources and expertise tailored to agency needs. | ▪ A university focused on pure research or a company with a single product offering may not be an ideal agency partner. |

## In-House Data Management

In-house data management solutions provide an ultimate level of control over data and storage/processing of that data. While this approach provides more control, it also requires a larger investment in workforce and physical infrastructure. For example, agencies may need to hire software developers and data experts who have strong expertise in data management and computer programming to effectively ingest and transform data into actionable TMC information. This can be a major challenge in a highly competitive market for these skills. Sometimes there may be many “data silos” even within a single agency, therefore multiplying the in-house data storage and management cost, while still maintaining barriers when it comes to data control.

Large in-house investments in data management systems can lock those agencies into a rigid system that does not adapt well to changing data sources and the emergence of new data elements and types. This can result in the agency falling behind until another large-scale data management system is procured (often doomed to become quickly obsolete).

Finally, in-house data management systems require not only an initial investment and operations and maintenance considerations, but also associated cybersecurity needs, especially in cases where personally identifiable information (PII) is collected and stored.

## Commercial Cloud Hosting

As data became a core asset in many industries, there has been an emergence of commercial cloud providers. These commercial cloud providers offer a variety of services ranging from barebones Infrastructure as a Service (IaaS), to more complex execution environment acting as a Platform as a Service (PaaS), and finally a full blown Software as a Service (SaaS) model.

These commercial cloud providers have identified a market and present a great option for agencies that do not have strong IT capabilities, dedicated technical staff, or sufficient up-front capital to establish their own IT infrastructure. Even agencies that do have these capabilities can benefit from scalability and elasticity of commercial cloud providers. Elasticity is the ability for infrastructure to grow or shrink in response to demand in order to manage load and cost of the system.

The biggest challenge with the use of commercial cloud hosting is that, while the costs may appear affordable when pricing data storage by number of bytes of data stored, the pricing model can get very complex and uncertain when it comes to transactions and processing of that data. For example, it may be cheap to load raw data into the cloud, but very expensive to extract it back when needed to perform after-action reviews or share that data and results with partners. If an agency has made a plunge and invested heavily in a cloud-based system, it may find itself held hostage to that cloud provider, even when costs rise, since the cost of extracting data and capabilities and establishing them elsewhere can be prohibitively expensive or time consuming.

Finally, even if an agency elects to use a commercial cloud platform, the agency still must have internal expertise and capacity to manage the system and develop necessary capabilities in the cloud.

## Trusted Partner Hosting

Trusted partner hosting is often a more agency-friendly solution that carries lower costs and lower risks. This approach utilizes a trusted third-party partner, such as a university, sister agency, or a trusted consultant to establish a cloud-based solution. For example, an agency may work with a university to build a system and store data at a lower cost than in-house or commercial cloud, while still having access to expertise and resources otherwise unavailable to the agency. Similarly, universities and sister agencies may be sharing internal State or local networks that provide an additional layer of security and improved performance.

Trusted partners can often provide similar levels of service as commercial cloud providers, but with a more focused mission. For example, a local university that provides a cloud storage and processing system for a TMC often is not also in the business of providing hosting solutions for major retailers and general consumers. This means that the staff is focused on building, operating, and maintaining the TMC system and addressing the needs of the agency rather than a broad spectrum of customers with potentially competing requirements and priorities.

The key to the successful partnership is a mutual understanding and appropriate partner selection. University, consultant, or sister-agency partners must be equipped and prepared to operate a production system and understand TMC needs. A university entity with a pure research focus or a company with a single product offering may not be the best option for an agency that is looking to remain flexible, innovative, and responsive to ever-changing data trends.

## Disaster Recovery and Security Considerations

When deciding between in-house and cloud storage, it is important to consider data sensitivity and risk. Cloud providers usually have a different threat profile, but also different threat management mechanisms. This means that cloud providers may be attractive targets for an attacker, especially a sophisticated one such as a foreign government, while a typical agency may not have the same risk. However, cloud providers acknowledge this threat and usually have a robust security policy to both prevent and recover from cybersecurity compromises. On the other hand, agencies may have limited funds and expertise to implement robust security mechanisms.

The nature of cybersecurity dictates that no system is ever 100 percent secure. The key is to balance the risk, vulnerability, and costs. Commercial cloud providers often exercise higher security standards than some individual agencies can afford to do, but even cloud providers are not immune to cybersecurity attacks. In fact, cloud providers have become a point of failure because a sophisticated attacker can focus on taking down a single cloud provider, which in turn could impact hundreds or thousands of cloud customers. Therefore, a TMC utilizing a commercial cloud to store and process data may become collateral damage in an attack on the cloud, even if the TMC would have never been the specific target. Even if TMC data is not especially sensitive, and the risk of a data breach is not as critical, a denial-of-service attack could render a TMC inoperable, with agency staff having no power or influence over resolution.

If a TMC decides to host its data and systems internally, it becomes a slightly smaller target, but is still vulnerable to cybersecurity attacks. Agencies must focus on developing a layered security architecture that contains at least the following components:

- Physical security.
    - Data centers with secured door access and server rack locks.
    - Limited authorized staff physical access and access audits.
    - Proper disposal of old hardware.
- Encryption.
    - Encryption of transmitted and stored information.
- Principle of least privilege.
    - Any entity or component must be able to access only the information and resources necessary for its legitimate purpose.
- Network and computer security.
    - Regularly maintained servers, workstations, and software – frequent firmware and software updates, patches, and virus scanners.
    - Secure boundaries – firewalls, routers, and intrusion prevention/detection systems, etc.
    - Network segmentation – separating different components on the network to avoid compromise of one segment affecting others.
- Social engineering.
    - Staff education and awareness of security threats that capitalize on human fallacies and cultural and social norms, such as never disclosing passwords or other sensitive information to potential impersonators.
    - Strong password policies and multi-factor authentication.

In addition to the security investment, agencies must ensure they have strong disaster recovery plans as well. With recent growth of ransomware attacks, it is critical for agencies to back up their data to several locations and be prepared to automatically switch processing to a backup location, or replace compromised data (corrupted by a hacker or encrypted by ransomware). Backups should be frequent and maintained in several geographically distributed locations that are not connected.

## OPEN SOURCE SOFTWARE VERSUS OPEN DATA

The transportation industry is becoming more acquainted with the terms "open source software" and "open data" now that they deal with software and data as core components of their business. However, there is a lot of confusion and misunderstanding of the difference between open source software and open data and in understanding what each term means.

Open source software uses a source code released under a license that allows others to change and distribute the code to anyone for any purpose. Open source software allows collaborative development of software, or building of specialized functionality on top of common core software. Open source software offers numerous benefits, such as lower cost, quicker innovation, sustainability, flexibility, and improved reliability and security. However, it also has certain drawbacks, some of which have major impacts in the transportation industry. For example, advanced traffic management system (ATMS) vendors are in the business of providing proprietary service and software to agencies to allow them to manage traffic more effectively and more efficiently. If agencies require their vendors to open source their software, vendors see that as a loss of competitive advantage over their competitors, but because they want to lock in a contract with an agency, they may agree to open source their software. This appears to be a win-win, when in fact it is a detrimental arrangement. Vendors often provide bare minimum functionality that satisfies client requirements, but have a legitimate concern that exposing advanced capabilities followed by adoption by other vendors will reduce the original vendor's ability to market the product.

Open data makes data freely available to everyone to use and republish without restrictions. Open data can be generated by open source software or proprietary software. Open data can provide transparency in how an agency operates its network and systems and allows the public, which may be indirectly funding the generation of that data, to understand challenges an agency faces and how that agency prioritizes its goals. In addition, open data allows a broader community to contribute and develop innovative applications, concepts, and capabilities that individual agencies either would not have otherwise, or would not have sufficient resources to implement.

The primary challenges with open data include the following:

- Privacy.
  - Someone must make a decision on what data is available. Care must be taken not to expose private information or reduce the value of data by removing potentially useful elements.
- Interpretation.
  - "If you torture the data long enough, it will confess to anything" – Darrell Huff.
  - Poorly documented or poor quality open data can lack clear context or be complex. Lack of understanding or malicious manipulation of underlying data to advance a specific agenda can lead to inconsistencies, incorrect, or various interpretations that can result in impactful consequences.
- Cost.
  - Collecting, processing, storing, managing, and disseminating data can be expensive, and if agencies cannot recover the cost, agencies and private sector entities may not be willing to make the necessary investment.
  - If a public agency provides open data using public funds, then the question becomes whether it is acceptable for private sector entities to monetize that data by repackaging it or including it in their applications.

While both open source software and open data have significant benefits, agencies must be careful to evaluate those benefits against potential drawbacks and be flexible in their requirements during procurement and contracting.

# CHAPTER 6. **CONTRACT CONSIDERATIONS**

Agency staff are not required to use any particular business model offered by the private sector. A well-informed agency has significant negotiating and bargaining power, and executing that power requires only a modest amount of effort. If an agency is offering funding for data, the agency should have the ability to request specific business models and approaches. If the requests are unreasonable, then no one will bid on the work. However, most bids receive proposals, and having an open dialogue with all potential bidders about the pros and cons of the agency's projected outcomes is critical.

Once an agency has selected a data provider (or winning bidder) and contract negotiations start, the company may request revisions to contract documents—including payment terms, acceptable use terms, or the business models used. Agencies should be wary of late-game change-ups to the scope of work and/or acceptable use terms and conditions. Less scrupulous companies sometimes attempt this bait-and-switch approach. That said, most ethical companies will point out in the proposals what it is that they are offering to the agency—even if it differs from what was requested in the request for proposal.

As discussed in the Federal Highway Administration report "Applying Archived Operations Data in Transportation Planning—A Primer," it is wise to consider a phased approach to data procurement. A best practice is to write contracts in a phased approach—start small and ensure the meeting of early goals. If the projects are successful, then work can progress on bigger tasks—adding layers of complexity and building off prior work and available data sets. Initiating extremely large tasks that are not easily broken down into smaller deliverables can be a recipe for confusion, cost overruns, disappointment, and waste.

Lastly, check references! Almost every data provider has worked with or is already working with another State or agency. Make sure you independently verify the data provider's data and business practices with friends and colleagues in the agency. It is also advisable to call multiple people within each agency to ensure one person does not skew the results.

# APPENDIX. EXAMPLE DATA USE AND DATA SHARING AGREEMENTS

This section includes example data use agreements and data sharing agreements that are already in use today.

## ORIGINAL DATA OWNERSHIP AND DATA LICENSING LANGUAGE

The University of Maryland inserted the following text into the original request for proposal for the procurement of probe data from the private sector in 2008. At the time, this was a revolutionary approach. Any agency procuring data from any third-party today should consider similar language—especially when multiple vendors are available.

**Attachment A**

Section 6.0 DATA OWNERSHIP AND DATA LICENSING

of University of Maryland Contract N136906

as reflected in Modification M002 of April 23, 2008

**6.0 DATA OWNERSHIP AND DATA LICENSING**

It is the intent of this contract to secure for the Coalition, its member organizations, and their officially designated representatives full rights to the traffic data to use in support of internal organization operations, and sufficient rights to the traffic data to disseminate traveler information to the public consistent with the organizations' traffic management and operations responsibilities. Paragraphs 6.1 through 6.6 further define the rights and uses. The Contractor's proposal should affirm the Contractor's ability to support the data rights presented herein. The Contractor may define additional restrictions to safeguard the commercial value of the Contractor's traffic data, but any such restrictions should not impede the use of the data for the envisioned purposes. Any restrictions imposed by the Contractor will be assessed in the technical evaluation of proposals.

6.1 The Contractor shall retain ownership of all traffic data provided to the I-95 Corridor Coalition as a result of this contract. The Coalition, its member organizations, and their officially designated representatives shall have the right to use the traffic data provided under this contract for transportation planning and operational analyses, service and data quality validation analyses, and all other internal organization applications. This includes the right to archive all the traffic data and use it for internal organization purposes for an unlimited period of time in the future, but excludes the right to sell or otherwise transfer the traffic data either (a) to other public entities which are not part of the Coalition, or (b) to any private entities for purposes not directly related to Coalition activities hereunder.

The Coalition and its member organizations will cooperate with the contractor to protect the commercial value of its data. This includes placing appropriate copyright notification on data disseminated to the public using methods and communication mediums that provide a reasonable opportunity for unlawful copying and use of the traffic data provided by the contractor to the Coalition. The copyright notice affixed to such data dissemination will read, "Copyright © 20__ INRIX, Inc. All rights reserved."

6.2 Realtime traffic data delivered by the Contractor may be provided by the Coalition, its member organizations, and their officially designated representatives to external users, subject to the following restrictions:

- Information may only be disseminated to the public using dynamic message signs (also known as variable message signs), portable message signs, highway advisory radio, 511 information systems, and Coalition and member supported websites and web services.
- Information disseminated to the public is restricted solely to travel times and speeds only for such media as dynamic message signs (also known as variable message signs), portable message signs, highway advisory radio, telephone-based 511 information systems.
- No restrictions are placed on the dissemination of data to the public utilizing Coalition and member supported websites and web services, including web-based 511 information systems.

6.3 The Coalition, its member organizations, and their officially designated representatives shall have the right to create visualizations and summary statistics of the archived traffic data (i.e., maps, graphs, charts, tables, etc.) for presentation and distribution to the general public. The University and Coalition will cooperate with the Contractor and make reasonable efforts to protect against the unlicensed distribution of data. However, neither the University nor the Coalition will assume any liability for unlicensed use of the data by third parties or unlicensed access to the data by third parties.

6.4 Contracting organizations, including universities, providing services on behalf of the Coalition or its member organizations, shall be subject to the same rights and restrictions given herein, but limited to the context of the contracted service. This includes organizations engaged by or acting on behalf of the Coalition to evaluate the accuracy, latency, and other parameters of the traffic data. Any contracting organizations, including any universities, desiring access to the traffic data for purposes not funded or sanctioned by the Coalition or its member organizations, must negotiate with the Contractor for access and rights to the traffic data.

6.5 Nothing in this contract shall preclude the Coalition, its member organizations, or their officially designated representatives from displaying or otherwise presenting any information to external users that has been obtained from other sources or other organizations that are not a party to this contract. Nothing in this contract shall preclude the Coalition and member organizations from displaying or otherwise presenting any information that is deemed essential to the safety of the traveling public.

6.6 Data provided by the Contractor may be incorporated into the Coalition's Integrated Corridor Analysis Tool (ICAT). ICAT is a geographic information system (GIS)-based transportation network for the 16 state Coalition region and linked databases of information about the region's roads, traffic volumes, and travel patterns. The dissemination of ICAT data may include summaries of historic traffic data with minimum summary periods of fifteen minutes and spatial resolutions that include road sections between interchanges, major intersections, major landmarks, and major destinations. The minimum length of such sections will generally be greater than four miles, with exceptions provided for bridges, tunnels and other unusual road network topography.

# I-95 CORRIDOR COALITION'S MULTI-STATE, MULTI-PROVIDER DATA USE AGREEMENT V2 (7/2014 – PRESENT)

**I-95 Corridor Coalition Traffic Flow Data Program**
**R009**
**Agreement for Use of Data**

1.0 This Agreement for Use of Data ("Agreement" or "Data Use Agreement") is entered into, by and among INRIX, Inc., a corporation organized under the laws of the State of Delaware and having a primary business address set forth in Section 8 below, HERE North America, a limited liability company organized under the laws of the State of Delaware and having a primary business address set forth in Section 8 below, and TomTom North America, Inc., a corporation organized under the laws of the State of California and having a primary business address set forth in Section 8 below (collectively referred to hereinafter as the "Data Vendors") and

________________________________________________

________________________________________________

(hereinafter "Data Licensee"), having an address as set forth on the signature page below. Data Vendors and Data Licensee together are hereinafter referred to as the "Parties".

2.0 Background: I-95 Corridor Coalition (hereinafter "Coalition"), an unincorporated coalition of members and affiliate members (collectively referred to herein after as "Coalition Members" or "Coalition Member" in singular), has authorized the University of Maryland (hereinafter "UMD"), an agency and instrumentality of the State of Maryland, to act on behalf of the Coalition to enable the Coalition to continue to serve as a valuable knowledgeable collaborative resource throughout the I-95 Corridor. Acting on behalf of the Coalition, UMD has executed agreements with the Data Vendors under which they have agreed to license real-time traffic data, including travel time, speed, complimentary metrics and associated products provided pursuant to task orders (hereinafter referred to as "Licensed Data") to Coalition Members and their subcontractors under the following agreements: UMD and INRIX have executed contract # 83794N-1-INRIX ; UMD and HERE North America have executed contract # 83794N-2-HERE ; and UMD and TomTom North America have executed contract # 83794N-3-TomTom (collectively hereinafter referred to as the "Contracts").

This Agreement, to be executed by all Data Vendors and Data Licensees, sets forth the terms and conditions under which Coalition Members and their subcontractors may access and use any Licensed Data purchased through the Contracts, irrespective of geography or time; for appropriate use of and liability for misuse of Licensed Data; and warranties regarding Licensed Data.

For the avoidance of doubt, Data Licensee acknowledges that Data Vendors' obligation to deliver the Licensed Data to Data Licensee is limited to the duration and the terms of active task orders under the applicable Contracts. The foregoing shall be without prejudice to Data Licensee's right to use the Licensed Data that it has received as set forth in clause 4.0 below.

3.0 Certification: Data Licensee certifies that it is a member or affiliate member of the Coalition in good standing or an entity under contract to a Data Licensee (that may include universities) that directly supports a Coalition Member in good standing via a written agreement (hereinafter

Version R009 10/22/2014 Page 1 of 10

referred to as 'Subcontractor'), and requires and is authorized to access/use the Licensed Data procured under the Contracts.

If Data Licensee is a Subcontractor, it shall complete Attachment A to this Agreement, which shall be incorporated as part of this Agreement. Coalition Members shall notify all Data Vendors and UMD upon the termination of its written agreement with any Subcontractor.

Data Vendors agree that Data Licensee is entitled to access and use Licensed Data under the terms of this Data Use Agreement. Notwithstanding whether Data Licensee elects to purchase or not purchase Licensed Data for its particular jurisdiction, Data Vendors agree that Data Licensee is entitled to access and use, at no cost, Licensed Data purchased by any Coalition Member, subject to the terms of this Data Use Agreement.

4.0 Grant of License: Data Vendors hereby grant Data Licensee a nonexclusive, fully paid up right and license to reproduce, use, distribute, make derivative works based on, and archive Licensed Data consistent with Data Licensee's traffic management, operations and planning responsibilities. Data Licensee is entitled to receive all Licensed Data purchased by any Coalition Member regardless of geographical or political boundaries of Data Licensee's respective jurisdiction.

5.0 Rights and Limitations of License: The license granted under this Agreement is subject to the following restrictions:

(a) Data Licensee shall not have the right to sell or otherwise transfer or disclose Licensed Data either to public or private entities that are not licensed to receive such data without prior written authorization from Data Vendors unless Data Licensee is required by applicable laws or regulations or pursuant to an order of a court of competent jurisdiction or a valid administrative or congressional subpoena to disclose Licensed Data. In that event, Data Licensee shall provide the affected Data Vendors prompt notice of the demand, unless prohibited by law, so they may take appropriate action to prevent disclosure, if they wish. Data Licensee shall provide a copy of any such notice to UMD. Nothing herein shall be deemed to authorize Data Licensee not to comply with any lawful order pending action by Data Vendors.

(b) Data Licensee shall limit access to Licensed Data to those of its employees and subcontractors who have a need to access and use Licensed Data in order to fulfill their contractual duties and shall require all such persons authorized to access and use Licensed Data to agree to abide by the terms of this Data Use Agreement. Any Data Licensee that desires access to Licensed Data for purposes not authorized by this Agreement must negotiate directly with Data Vendors to acquire such additional access and rights.

(c) Data Licensee may disseminate real-time traffic data delivered by Data Vendors to the public, subject to the following restrictions:

i. Licensed Data may only be disseminated to the public using dynamic message signs (also known as variable message signs), portable message signs, highway advisory radio, 511 information systems, and Coalition Members' supported websites, web services, social media, and smart phone applications; and

ii. Licensed Data disseminated to the public shall be restricted solely to travel times and speeds only that is disseminated via dynamic message signs (also known as variable message signs), portable message signs, highway advisory radio, telephone-based 511 information systems; and

iii. This Agreement does not place any restrictions on dissemination of data to the public through Coalition Members' supported websites, web services, social media, and smart phone applications, including web-based 511 information systems.

(d) The license granted by Data Vendors to Data Licensee authorizes Data Licensee to create visualizations and summary statistics of the archived traffic data (i.e., maps, graphs, charts, tables, etc.) for presentation and distribution to the general public ("Derivative Works"). Data Licensee shall own all copyrights in all such Derivative Works to the extent those works are protected by copyright.

(e) Nothing in this Agreement shall preclude Data Licensee from distributing, displaying or otherwise presenting any traffic data or derivative works deemed essential to the safety of the traveling public.

6.0 Right to Acquire Non-Licensed Data: Nothing in this Agreement shall prohibit Data Licensee from acquiring, displaying or otherwise presenting or sharing information that Data Licensee has obtained from sources other than Data Vendors.

7.0 Prevention of Unauthorized Use: Data Licensee will cooperate with Data Vendors to protect the commercial value of Licensed Data by taking such measures as:

(a) retaining all proprietary or restricted use notices included on Licensed Data as received; and

(b) not obstructing or modifying proprietary or restricted use notices included on Licensed Data as received; and

(c) ensuring that all copies of Licensed Data include all proprietary or restricted use notices included on Licensed Data as received . To the extent Data Vendors do not include any proprietary or restricted use notices on Licensed Data as delivered to Data Licensee, Data Licensee shall insert, at a minimum, the following notice on any copies of Licensed Data that Data Licensee makes: "PROPRIETARY INFORMATION OF DATA VENDOR. USE BY ENTITIES OTHER THAN AUTHORIZED, LICENSED USERS PROHIBITED"; and

(d) storing and disseminating Licensed Data using methods, communication mediums and technologies that provide reasonable protections against their unlawful copying and unauthorized access and use.

8.0 Notice of Unauthorized Use: In the event Data Licensee becomes aware of an inappropriate use or unauthorized disclosure of Licensed Data, Data Licensee shall provide immediate verbal notice as soon as practicable and subsequent written notice within 24 hours of its verbal notice to UMD and to the Data Vendor(s) whose data are the subject of inappropriate use or unauthorized disclosure as follows:

**University of Maryland Program Managers**

**INRIX Inc.**

**HERE North America LLC**

**TomTom North America, Inc.**

Subject to the above paragraph, all notices and approvals required to be made under this Agreement shall be made in writing and delivered (i) in person; (ii) by facsimile, with confirmation of transmission, (iii) by electronic mail (email) with return confirmation of delivery, or (iv) by first class mail, postage prepaid and addressed to the contact for each party specified above or such other person and address as each party may hereafter designate in writing. Notice shall be deemed effective upon receipt.

9.0 Indemnification: Data Vendors hereby indemnify and agree to hold harmless UMD, Data Licensees and their respective officers, employees and agents from and against any and all claims, actions, costs, judgments or damages of any type arising out of an allegation that Licensed Data infringes the intellectual property or proprietary rights of any third party or a breach of the representations and warranties of Data Vendors.

(a) Upon becoming aware of an allegation of infringement or a breach of a Data Vendor's representations and warranties, Data Licensee shall promptly notify the affected Data Vendor(s) and UMD.

(b) Data Vendors' duty to indemnify is conditioned upon (i) Data Vendors having sole control of the defense and settlement of the claim (provided that Data Vendors may not settle or compromise or defend any claim unless they unconditionally release all other parties from all liability, and further provided that Data Vendors must obtain prior approval of any such settlement or compromise from counsel for UMD and the Data Licensee which shall not be unreasonably withheld or delayed); (ii) Data Licensee provides, at Data Vendors' expense, information and reasonable assistance upon Data Vendors' request; and (iii) Data Licensee has not already compromised or settled the claim.

10.0 Liability: Under no circumstance will Data Licensee be responsible for another Data Licensee's breach of its duties under this Data Use Agreement. Each Data Licensee shall be liable for its own violations of this Agreement.. IN NO EVENT WILL ANY PARTY OR ITS OFFICERS, AGENTS, OR EMPLOYEES BE LIABLE TO THE OTHER PARTIES FOR ANY INCIDENTAL, SPECIAL, INDIRECT, EXEMPLARY OR CONSEQUENTIAL DAMAGES OF ANY KIND, INCLUDING BUSINESS EXPENSE, MACHINE DOWN TIME, LOSS OF PROFITS, OR DAMAGE OR INJURY TO PROPERTY FOR ANY CLAIMS, DEMANDS OR DAMAGES ARISING OUT OF OR RELATED TO THE PERFORMANCE OR ITS OBLIGATIONS UNDER THIS AGREEMENT OR THE USE OF LICENSED DATA BY ANYONE EVEN IF THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

11.0 Term and Termination

(a) This Agreement will commence with respect to an individual Data Licensee as of the last date of execution by each Data Vendor, the Data Licensee, and the UMD representative affirming the good standing of the Data Licensee.

(b) This Agreement will terminate upon the occurrence of any of the following:
(i) One or more Data Vendors or UMD gives written notice to Data Licensee of its breach of one or more obligations under this Agreement and Data Licensee's failure to cure its breach within thirty (30) days of receipt of notice of breach; or
(ii) UMD notifies Data Vendors and Data Licensee that Data Licensee is not a member or affiliate member in good standing of the I-95 Corridor Coalition and Data Licensee fails to restore its good standing within thirty (30) days of receipt of notice; or
(iii) A Coalition Member or its Subcontractor gives written notice to the Data Vendors and UMD that Subcontractor no longer has a need to access/use Licensed Data in which case this Agreement will terminate only with respect to the Subcontractor; or
(iv) Data Licensee gives written notice to Data Vendors and UMD that it wishes to terminate this Agreement.

Notices shall be provided to the addresses listed in Section 8.0 above.

Except for the reasons stated above, this Agreement will remain in effect and will not terminate.

(c) Termination under section 11.0(b) (i) and (ii) will become effective upon expiration of the 30-day period if the breach has not been cured. Termination under Section 11.0(b) (iii) and (iv) will become effective immediately upon receipt of notice.

(d) In the event of termination of this Agreement:
(i) Data Vendor will cease to provide Licensed Data to Data Licensee; and
(ii) Data Licensee will no longer be able to access Licensed Data maintained in archives and analysis tools at UMD; and
(iii)Data Licensee must destroy any and all Licensed Data in its possession and certify their destruction to UMD within thirty (30) days of the effective date of termination

12.0 Representations and Warranties

(a) Data Vendors represent and warrant that all Licensed Data shall be original and unencumbered.

(b) Data Vendors represent and warrant that they either own the Licensed Data or are authorized by the owner(s) of Licensed Data to grant licenses to Data Licensees under this Agreement or that Licensed Data are in the public domain.

13.0 General

(a) The validity, interpretation and effect of this Agreement shall be governed by the laws of the state where Data Licensee is located without regard to its conflicts of laws rules when Data Licensee is an agency or instrumentality of state government.

(b) No Party may assign its rights or obligations under this Agreement, except with the prior written approval of the other Parties. Such approval will not be unreasonably withheld.

(c) This Agreement may be modified only by written agreement of authorized representatives of all Parties.

(d) This Agreement supersedes any previously executed agreement between Data Licensee and Data Vendor/s with respect to Licensed Data.

(e) Nothing herein shall be construed to create a partnership, joint venture, or teaming agreement between or among the Parties and nothing herein shall be construed to imply that any Party's employees are employees of another Party.

(f) The Parties shall use their best efforts to resolve any disagreement that arises out of this Agreement amicably.

(g) No provision of this Agreement shall be waived unless in writing and signed by all Parties to this Agreement. The waiver of any provision of this Agreement shall not be deemed to be a continuing waiver or the waiver of any other provision of this Agreement.

(h) If any one or more of the provisions contained in this Agreement is held to be invalid, illegal, or unenforceable in any respect for any reason, then such invalidity, illegality, or unenforceability shall not affect any other provision hereof or any other application of the affected provision.

(i) This Agreement, together with Attachment A (if applicable), embodies the entire understanding between and among the Parties. There are no contracts, understandings, conditions, warranties or representations, oral or written, express or implied, with reference to the subject matter hereof which are not merged herein.

(j) This Agreement may be executed in counterparts, all of which when taken together will be deemed one original. The Parties agree to accept digital delivery of this executed Agreement.

Signature page follows

By signing below, the Parties certify that they agree to the above terms and are duly authorized to bind their respective entities.

______________________________________________

______________________________________________

**Data Licensee** (agency, affiliate, subcontractor)

______________________________________________

By: ______________________ Date______________
Name:

______________________________________________
Title

______________________________________________
Address

______________________________________________
Name and Title of Primary Contact:

______________________________________________
Telephone/Facsimile/E-mail

**INRIX, Inc.**

By: ______________________ Date______________
Name:

______________________________________________
Title

**HERE North America, LLC**

By: ______________________ Date______________
Name:

______________________________________________

Title

Version R009 10/22/2014 Page 8 of 10

**TomTom North America, Inc.**

By: ______________________________ Date______________
Name:

____________________________________________________________
Title

**UNIVERSITY OF MARYLAND CERTIFICATION**

University of Maryland hereby certifies that as of the date below, Data Licensee is a Coalition Member or affiliate in good standing, or a Subcontractor of a Coalition Member or affiliate in good standing.

By: ______________________________ Date______________

**ATTACHMENT A**
**Any Data Licensee that is**
**a Subcontractor must provide the following information**
**About its Contract with a Member/Affiliate to "Establish Need to Know"**

1. Name and Address of Subcontractor/Data Licensee:

   ______________________________________________________________

2. Contractor from whom data is requested (check all that apply)
   ____ INRIX ______ Here North America ______ TomTom

3. Explain why You Need Access to Licensed Data:

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

4. Prime Contract/Subcontract Source ______________________________

5. Contract/Agreement Number ______________________________

6. Contract/Agreement Period-of-Performance:

   From: ________________ to: __________________

7. Contact Information from Contracting Agency (Coalition Member/Affiliate):

   Name: ________________________________________ Phone: _______________________

   E-mail Address: ______________________________________________________________

Subcontractor's authorized official certifies that the information provided above is current and accurate.

By: _____________________________ Date_______________

Name and Title: _______________________________________________________________

Version R009 10/22/2014 Page 10 of 10

# I-95 CORRIDOR COALITION'S MULTI-STATE, SINGLE-PROVIDER DATA USE AGREEMENT V1 (7/2009 – 7/2014)

## Federal Highway Administration/INRIX National Performance Management Research Data Set Data Sharing License

### NATIONAL PERFORMANCE MANAGEMENT RESEARCH DATA SHARING LICENSE

As provided for in Contract No. DTFH61-17-C-00003 ("FHWA Contract") between the DOT/Federal Highway Administration ("FHWA") and the University of Maryland ("UMD"), in which INRIX, Inc. ("INRIX") is a designated sub-contractor, INRIX grants AGENCY a non-exclusive, non-transferable, non- sublicensable (except as set forth in this agreement) license to use Data (identified below) for the Purpose (specified below) on the terms and conditions herein as evidenced by delivery of a copy of Data to Agency Agency's use of the Data is evidence of the acceptance of the restrictions provided for in this license.

**1. IDENTIFICATION OF DATA and PURPOSE**

1.1 Data. Data refers to the National Performance Management Research Data Set provided by INRIX under the FHWA Contract. Data includes average travel time from vehicle probe data in three classifications: 1.) all vehicles, 2.) passenger vehicles, and 3.) freight trucks at a statistically significant sample size for both freight and passenger traffic for the following geographic coverage:

- United States Interstate System;
- NHS including NHS intermodal connectors;
- Strategic Defense Network Roadways (STRHANET); and
- Border crossings on principal arterials.
- Arterials within a five-mile radius on either side of the border at the top 20 United States/Canada Border Crossings as measured by average daily truck trips.
- Arterials within a five-mile radius on either side of the border at the top six United States/Mexico Border Crossings as measured by average daily truck trips.

1.2 Purpose. AGENCY may use Data:

- to support performance management activities such as creating performance indicators, measures and evaluations;
- to disseminate summaries of the Data to the public consistent with the organizations' transportation planning, programming, management and operations responsibilities as they pertain to performance management activities;
- in transportation planning and operational analyses, service and data quality validation analyses; and
- in applications for Agency's internal business.
- to provide a copy of a spreadsheet of the data used in developing a plan or capital program based in part or on performance measurement if requested for validation of decisional materials.

AGENCY may not use Data to make data sets or aggregated average travel time databases publicly available. For avoidance of doubt, the intent of this license is to enable AGENCY to provide summaries and statistics based on the Data but not to provide the Data in a form that would enable unlicensed parties to build databases of the Data.

**2. LICENSE**

2.1 Agency. Agency warrants that it is a State Department of Transportation or Metropolitan Planning Organization receiving federal transportation funds and is authorized by the US Federal Highway Administration to receive Data.

2.2 Contractors. Agency may grant contractors the right to use Data for work performed for Agency under the Purpose defined in this Agreement. Agency shall be responsible for the performance of its contractors and contractors must be bound to preserve the confidentiality and security of Data on terms at least as protective as those provided for in this Agreement. Agency may archive Data for an unlimited period of time for use only for the Purpose. Contractors must also execute this agreement, including completing Attachment A.

2.3 Restrictions. Agency shall not (a) sell or share Data with other public entities except as required to fulfill the Purpose; (b) sell, disclose or otherwise transfer Data to private parties except to contractors to the extent expressly permitted in Section 2.2; (c) disassemble, decompile, alter or otherwise reverse engineer Data; (d) combine, incorporate, utilize, or distribute Copies of Data with or in connection with any product or system which, alone or in combination with such Copies, infringes any other person›s or entity›s intellectual property rights or any other rights; (e) export from anywhere any part of Data or any direct product thereof except in compliance with, and with all licenses and approvals required under, applicable export laws, rules and regulations; or (f) use Data in any manner not expressly authorized herein.

2.4 Data Rights. INRIX represents and warrants that it has the right to grant all licenses granted by it hereunder. Agency acknowledges that INRIX, its licensors and suppliers: (a) owns or has the right to use all intellectual property rights in and to Data and (b) retain all such rights under this Agreement.

**3. FEES**

3.1 Fees. There is no fee payable by Agency to INRIX for Data described on the first page. Agency may agree with INRIX by written addendum for additional data to be provided for an agreed fee.

**4. TERM and TERMINATION**

4.1 Term. The term of this Agreement shall commence on the date of signature and continue through April 9, 2022, unless terminated as provided below (“Term”).

4.2 Termination for Breach. Either party may terminate this Agreement by written notice (i) immediately upon material breach by the other party if such breach cannot be remedied, or (ii) if the other party fails to cure any material remediable breach within 30 days of receipt of written notice.

4.3 Termination for Bankruptcy. INRIX may terminate this Agreement immediately if any of the following events occur: (a) Agency's voluntary bankruptcy or application for bankruptcy; (b) involuntary bankruptcy or application for bankruptcy for Agency not discharged within 60 days; (c) appointment of receiver or trustee in bankruptcy for all or a portion of Agency's assets; (d) an assignment for the benefit of Agency's creditors.

4.4 Termination of FHWA or UMD Contract. This Agreement shall terminate upon the termination or expiration of either the FHWA or UMD Contract. While content received during the term of this license can continue to be used as provided for in this Agreement after termination or expiration of this Agreement, no additional data shall be licensed under this Agreement after termination or expiration of the FHWA or UMD Contract.

4.5 Survival. Sections 4.4, 7 & 8 and the second sentence of 2.4 shall survive termination or expiration of this Agreement.

**5. DELIVERY & QUALITY**

5.1 Delivery. Subject to acceptance of the terms in this Agreement, INRIX will make Data available as agreed with the FHWA and UMD.

5.2 Error Reporting. Agency shall promptly provide any information it learns concerning errors, problems, complaints, and related matters concerning Data to INRIX, without charge, for INRIX's unlimited use, including incorporation into Data. Agency shall not retain, acquire or assert any right, title or interest in or to Data or the intellectual property rights thereto based on the transfer of such information to INRIX or INRIX's use or incorporation of such information in Data or otherwise.

**6. COPYRIGHT & MARKETING**

6.1 Copyright Notices. Agency shall conspicuously display INRIX's and its suppliers' copyright notices, marks and other proprietary rights legends ("INRIX Legends") as specified by INRIX, on Copies of Data, on on-screen displays, on splash and start-up screens, in the instructions (printed and electronic), and other written materials incorporating Data and distributed by or on behalf of Agency hereunder (collectively, "Collateral").

6.2 Attribution. In all instances where Data is used or where Collateral references Data, Agency shall attribute INRIX as the creator and source of origin of Data, and shall not in any way be misleading in that regard or represent or imply that Agency or any third-party is the creator or source of origin of Data. Agency shall further provide an acknowledgment of FHWA support and a disclaimer in any publication of any material, based on or developed using Data contract in the following terms:

"This material is based upon work supported by the Federal Highway Administration under contract number DTFH61-17-C-00003." "Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Federal Highway Administration."

6.3 License of Marks. During the term of this Agreement, INRIX grants Agency a non-exclusive, non-transferable, non-sublicensable right to use the INRIX Marks as required under this Agreement. Agency must conspicuously indicate in any materials displaying the INRIX Marks that the INRIX Marks are registered trademarks or service marks of INRIX, as applicable. Nothing stated herein shall constitute a grant or other transfer to Agency of any right, title or interest in the INRIX Marks or any other intellectual property rights of INRIX. Agency's use of the INRIX Marks shall inure to the benefit of INRIX. Upon termination or expiration of this Agreement for any reason, Agency shall immediately cease all use of INRIX Marks.

**7. CONFIDENTIALITY.** Each party agrees that all business, technical and other information it obtains from the other is the confidential property of the disclosing party ("Confidential Information"). The receiving party will hold in confidence and not disclose any Confidential Information of the disclosing party, other than to its parent company and affiliates. Upon termination of this Agreement or request of the disclosing party, the receiving party will return or destroy (and certify such destruction) all Confidential Information of such disclosing party. The receiving party shall not be obligated with respect to information the receiving party can document: (a) is or has become readily publicly available without restriction through no fault of the receiving party; or (b) is received without restriction from a third-party lawfully in possession and empowered to disclose such information; or (c) was rightfully in the possession of the receiving party without restriction prior to its disclosure by the other party; or (d) was independently developed by the receiving party without access to such Confidential Information; or (e) is required to be disclosed by law or order of court of competent jurisdiction.

**8. DISCLAIMER/Limitation of Liability.**

**8.1 EXCEPT AS EXPRESSLY SET FORTH IN THIS AGREEMENT, DATA IS PROVIDED "AS IS" AND INRIX MAKES NO REPRESENTATIONS OR WARRANTIES. INRIX EXPRESSLY DISCLAIMS ANY IMPLIED WARRANTIES OR CONDITIONS OF ANY KIND, INCLUDING, WITHOUT LIMITATION, ANY WARRANTY OR CONDITION OF QUALITY, PERFORMANCE, MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. EXCEPT AS SPECIFICALLY SET FORTH IN THIS AGREEMENT, INRIX DOES NOT WARRANT, GUARANTEE, OR MAKE ANY REPRESENTATIONS REGARDING THE USE, OR THE RESULTS OF THE USE, OF DATA OR ANY OTHER MATERIALS IN TERMS OF CORRECTNESS, ACCURACY, RELIABILITY, OR OTHERWISE.**

8.2 Except as otherwise provided and to the maximum extent permitted by law, the liability of both parties shall be limited to direct damages only, thus excluding liability for any other damages such as indirect, special, incidental, or consequential or punitive damages (including but not limited to lost profits, lost data, lost revenue, lost savings, lost business and loss of goodwill).

8.3 In no event shall INRIX's aggregate liability with respect to any matters whatsoever arising under or in connection with the Agreement exceed $50. Agency acknowledges and agrees that the fees and allocation of the risks (as expressed in the indemnities and the limits on warranties, liabilities, damages and remedies) contained herein reflect the economic basis of this Agreement, in the absence of which this Agreement would not have been made.

## 9. OTHER PROVISIONS

9.1 General. This Agreement is the entire agreement between the parties regarding the subject matter hereof. It supersedes all prior oral or written communications, representations, and agreements of the parties relating thereto. This Agreement may be modified or amended only by a written instrument duly executed by the parties.

9.2 Waiver. No waiver will be deemed effective unless set forth in writing and signed by the party charged with such waiver, and no waiver of any right arising from any breach will be deemed to be a waiver or authorization of any other breach or of any other right arising under this Agreement.

9.3 Severability. If any provision of this Agreement is held to be invalid, illegal, or unenforceable, the remaining provisions hereof shall be unaffected thereby and remain valid and enforceable as if such provision had not been set forth herein. The parties agree to substitute for such provision a valid provision that most closely approximates the intent of such severed provision.

9.4 Governing Law. This Agreement shall be construed and governed by the substantive laws of the State of Illinois without giving effect to the conflict of laws provisions unless the state law applicable to AGENCY or the rules applicable to AGENCY requires the application of such state law in which case such state law shall apply. The United Nations Convention of Contracts for the International Sale of Goods shall not apply to this Agreement.

9.5 Assignment. The rights and obligations of each party under this Agreement may not be transferred or assigned directly or indirectly without the prior written consent of the other party, which consent will not be unreasonably withheld, except that INRIX may assign this Agreement to a parent, subsidiary, or any entity that acquires substantially all of its stock, assets or business.

9.6 Notices. All notices under this Agreement must be in writing and delivered by hand, fax or nationally recognized overnight courier at the addresses set forth on the cover page, or at such other address as either party shall have furnished to the other in writing. Such notices shall be effective (a) if sent by overnight courier, two business days after mailing, and (b) if sent otherwise, upon receipt.

9.7 Force Majeure. INRIX shall not be liable to Agency for a failure to perform any of its obligations under this Agreement, due to circumstances beyond its reasonable control, provided it notifies Agency of the delay.

9.8 Relationship. This Agreement is between INRIX and Agency. No third-party beneficiaries are intended. In connection with this Agreement each party is an independent contractor and as such does not have any authority to bind or commit the other. Nothing herein shall be deemed or construed to create a joint venture, partnership or agency relationship between the parties for any purpose.

9.9 Agency agrees Data is a “commercial item” as that term is defined at 48 C.F.R. (“FAR”) 2.101, licensed in accordance with this License, and each copy of Data delivered or otherwise furnished shall be treated and marked (or embedded as appropriate) with the following “Notice of Use”:

Notice of Use Contractor (Manufacturer/Supplier) Name: INRIX, Inc. Contractor (Manufacturer/Supplier) Address: 10210 NE Points Drive, Suite 400, Kirkland, WA 98033

The Licensed Data is a commercial item as defined in FAR 2.101 and is subject to the Data Sharing License under which the Data was provided.

## ATTACHMENT A

**Any Data Licensee that is a Contractor must provide the following information about its Contract with a Member/Affiliate to "Establish Need to Know"**

Name and Address of Subcontractor/Data Licensee:

_______________________________________________________________________________

2. Explain why You Need Access to Licensed Data:

_______________________________________________________________________________

_______________________________________________________________________________

_______________________________________________________________________________

_______________________________________________________________________________

3. Prime Contract/Subcontract Source ________________________________

4. Contract/Agreement Number ________________________________

5. Contract/Agreement Period-of-Performance:

From: __________________ to: ____________________

6. Contact Information from Agency:

Name: ___________________________________________ Phone: _________________________

E-mail Address: ___________________________________________________________________

Contractor's authorized official certifies that the information provided above is current and accurate.

By: _______________________________ Date_________________

Name and Title: ____________________________________________________________________

## WAZE CONNECTED CITIZENS PROGRAM AGENCY AGREEMENT

Waze frequently updates the terms and conditions of their Connected Citizens Program. They have multiple license agreements, including one for researchers, one for a one-way agreement, and another for a two-way agreement. An example of their two-way agreement from 2015 can be viewed here: https://arlington.granicus.com/MetaViewer.php?view_id=&clip_id=3099&meta_id=142727.

U.S. Department of Transportation
**Federal Highway Administration**

Federal Highway Administration
U.S. Department of Transportation
1200 New Jersey Avenue, SE
Washington, DC 20590

FHWA-HOP-18-084
July 2019