CONNECTED V. AUTOMATED VEHICLES AS GENERATORS OF USEFUL DATA

September 30, 2014

MDOT
Michigan Department of Transportation

CAR
CENTER FOR AUTOMOTIVE RESEARCH

leidos

Sponsoring Organization:
Michigan Department of Transportation (MDOT)
425 Ottawa Street
P.O. Box 30050
Lansing, MI 48909

Performing Organization(s):

Center for Automotive Research (CAR)
3005 Boardwalk, Ste. 200
Ann Arbor, MI 48108

Leidos, Inc.
11951 Freedom Drive
Reston, VA 20190

Connected vs. Automated Vehicles as Generators of Useful Data
MDOT REQ. NO. 1259, Connected and Automated Industry Coordination
Task C.1. Connected vs. Automated Vehicles as Generators of Useful Data

September 30, 2014

Author(s):
Qiang Hong, PhD., CAR
Richard Wallace, M.S., CAR
Gregory Krueger, P.E., Leidos
Managing Editor(s):
William Tansil, Asset Management Administrator, MDOT
Matt Smith, P.E., PTOE, Statewide ITS Program Manager, MDOT
Collin Castle, P.E., Connected Vehicle Technical Manager, MDOT
Marty Lontz, MDOT
Additional Contributor(s):
Eric Paul Dennis, P.E., CAR

Abstract:

This report summarizes the types of data that may be available from the United States Department of Transportation (USDOT) connected vehicle effort, as well as data from advanced automated vehicle systems. The USDOT connected vehicle program emphasizes the creation of transportation-related data that will be available and usable by public transportation agencies for many applications. By contrast, advanced automated vehicle systems are generally being developed and deployed without consideration of data creation for public use. This report concludes that the emergence of automated vehicle systems represents an opportunity to obtain useful data from these vehicles. To some extent, these data may depend on connected vehicle

technology for transmission beyond the vehicle. Importantly, automated systems have the potential to capture both vehicle-centric and infrastructure-oriented data, while connected vehicle technology is more amenable to capturing vehicle-centric data.

ACKNOWLEDGMENTS

This document is a product of the Center for Automotive Research under a State Planning and Research Grant administered by the Michigan Department of Transportation.

EXECUTIVE SUMMARY

The Michigan Department of Transportation (MDOT) is a recognized leader in advanced transportation system technologies (including intelligent transportation systems, or ITS) and is committed to remaining a source of innovation. The management of cutting-edge projects requires bold leadership and acting without precedent. In keeping with its leadership, MDOT asked a team led by the Center for Automotive Research (CAR), assisted by Leidos, to examine the potential data that may be available from near-term deployments of connected and automated vehicle (CAV) technologies, considered separately and jointly.

Quite possibly, the strongest potential for future CAV technologies to provide data of value to transportation agencies is for the USDOT Connected Vehicle Program to result in broad deployment of the core connected vehicle system. A prototype core connected vehicle system is being tested in the Safety Pilot Model Deployment in Ann Arbor, MI. Variations of this core system are being, or will be, deployed in test beds around the country, possibly resulting in a step towards implementation of a national connected vehicle network. Historically, the USDOT Connected Vehicle Program was designed primarily to support National Highway Traffic Safety Administration (NHTSA) decision-making regarding potential rulemaking concerning vehicle-to-vehicle communication for safety applications. However, using connected vehicles to generate data for use by transportation agencies has been a strong secondary design principle. Much of the research underway in the USDOT connected vehicle program is directed toward such data-providing applications.

Automated vehicle technologies represent both an opportunity and threat for the possibility of gleaning data from CAVs. In recent years, automated vehicle systems have been deployed that can provide safety-related applications such as driver warnings, automated emergency braking, lane-keeping, and even hands/feet-free highway autopilot and traffic-jam assist. Higher levels of 2automation, and possibly even fully self-driving cars, might be deployed in future years. Automated vehicles have the potential to generate vast amounts of data that can support a variety of transportation agency needs and applications. While these automated technologies vary in their design, functional components, and the types of data they produce and consume, they generally are being deployed without reference to the USDOT connected vehicle efforts. Thus, the information produced and used by these systems is mostly proprietary and likely not available without any cost to public-sector

transportation agencies. Furthermore, NHTSA is still in the process of drafting regulations that would enable many connected vehicle applications, and eventual deployment of these applications is not guaranteed.

TABLE OF CONTENTS

1 Introduction 1
2 Transformations in Transportation Data Collection and Management Practices.......... 3
2.1 Big Data in Transportation... 4
2.2 Multi-Source and Open Data Environments... 5
2.3 Mobile Data Communications 6
3 Connected vs. Automated Vehicles 9
3.1 Definitions.. 9
3.2 Connected Vehicle Enabling Technologies 11
3.3 Automated Vehicle Enabling Technologies 12
3.4 Common and Distinguishing Technologies.. 14
4 Connected and Automated Vehicle Data Systems.. 16
4.1 Connected Vehicle Applications.. 16
4.2 Connected Vehicle Data Environment Concept 17
4.3 Connected Vehicle Physical Architecture and Data Exchange 18
4.4 Connected Vehicle Data Systems 22
4.5 Basic Safety Message Data.. 26
4.6 Automated Vehicle Generated Data 29
4.7 Availability of Data.. 35
5 Conclusions and Recommendations 38
References.. 40
Appendix A: List of Abbreviations 42

MDOT
Michigan Department of Transportation

CAR
CENTER FOR AUTOMOTIVE RESEARCH

leidos

1 INTRODUCTION

The advent of telecommunication and information technologies such as wireless communications, mobile apps, Web-based tools, sensors, Global Positioning System (GPS), and mapping have brought about major leaps in transportation data collection and management practices. Connected and automated vehicle (CAV) systems essentially rely on a combination of these technologies. They have the potential to enhance current operations, transform transportation system management, supplement or replace conventional data collection methods, and fundamentally alter the transportation data paradigm. Therefore, not surprisingly, in a recent request for information (RFI) about the Connected Vehicle Pilot Deployment Program issued by the USDOT, "data" was the most frequent word that appeared summed across all responses (Figure 1).

FIGURE 1: CONNECTED VEHICLE PILOT RFI RESPONSE WORD CLOUD[1]

Given the importance of data in transportation system management and operations, and the potential for data enhancements arising from CAV technology, the Michigan Department of Transportation (MDOT) asked the Center for Automotive Research (CAR) to investigate, and compare and contrast, connected and automated vehicles as generators of useful data. The objectives of this study are to identify connected and automated vehicle technologies and their potential impacts on transportation data sources and

---

[1] Source: USDOT ITS JPO. Connected Vehicle Pilot Deployment Program Request for Information (RFI) Update. 2014.

management. Chapter Two of the report describes the transformations in transportation data collection and management, followed by discussions of CAV technologies in Chapter Three. Chapter Four is devoted to the data environment concept, connected vehicle applications, data flow, and data elements. Finally, Chapter Five details remaining challenges and offers concluding remarks and recommendations.

2 TRANSFORMATIONS IN TRANSPORTATION DATA COLLECTION AND MANAGEMENT PRACTICES

Transportation data programs are important to support decision-making processes and help understand our transportation challenges. Like many other transportation agencies around the country, MDOT is involved with a wide spectrum of data programs ranging from the traditional traffic, travel, safety, system inventory and condition, and finance, to the rapidly developing intelligent transportation system's technologies and new data sources. Data management programs may define key data principles and include specific data elements within each of these categories depending on the organizational needs. The American Association of State Highway and Transportation Officials (AASHTO) developed a set of core data principles proposed for use by state DOTs, as follows:[2]

- VALUABLE - Data are a core business asset that has value and are managed accordingly.
- AVAILABLE - Access to data is critical to performing duties and functions, data must be open and usable for diverse applications and open to all.
- RELIABLE - Data quality is acceptable and meets the needs for which the data are intended.
- AUTHORIZED - Data are trustworthy and safeguarded from unauthorized access, whether malicious, fraudulent, or erroneous.
- CLEAR - Data dictionaries are developed and metadata established to maximize consistency and transparency of data across systems.
- EFFICIENT - Data are collected once and used many times for many purposes.
- ACCOUNTABLE - Timely, relevant, high quality data are essential to maximize the utility of data for decision making.

The advancement of connected and automated vehicle technologies, as well as the significant increase in the number of personal mobile devices and amount of "crowdsourced" information, will create new dynamics of transportation

---

[2] Slater 2013.

data environment. The new data environment will include various actors (individual, business, and public sector), and it can be characterized by the following key transformations: big data in transportation, multiple data sources and application platforms, and mobile data communications.

2.1 BIG DATA IN TRANSPORTATION

Big data is the term for a collection of large, quickly growing, and complex data sets that generally have some or all of following features:

- Digitally generated
- Passively produced
- Automatically collected
- Geographically or temporally trackable
- Continuously analyzed

Big data could be any structured or unstructured data such as text, sensor data, audio, video, click streams, log files, and more. The key attributes of big data are:

- Volume: the exponential jump in the amount of data we collect
- Velocity: bits and bytes have to be processed at high speed
- Variety: structured and unstructured data in many formats and from diverse sources
- Value: data needs to be converted into meaningful and trustworthy information [3]

Thanks to on-board computers, sensors, and wireless connectivity, big data have been expanding into the automotive and transportation sector. For example, big data are helping the automotive industry to improve operational efficiency in designing, building, and servicing vehicles. Automakers are also increasingly connecting data from vehicles and people’s behavior to the data about the environment in which the vehicle is operating (weather, traffic, hazardous situations, etc.).

The transport sector’s increasing ability to track the location of mobile devices has enabled both the monitoring of traffic to save time and reduce congestion

---

[3] United Nations 2012.

as well as the provision of new location-based services. For example, in 2012 TomTom, a leading provider of navigation hardware and software, had more than 5,000 trillion data points in its databases, gleaned from its navigation devices and other sources, describing time, location, direction and speed of travel of individual users.[4] The trucking industry is using telematics and electronic on-board recorders (EOBRs) to collect data and communicate in real time to improve safety and operational performance. The future of big data in trucking will be cross-referencing real-time driver data with data on weather, parking availability and traffic delays to deliver information to the driver as quickly as possible. Many of these programs are still in the initial phase of implementation, but the potential improvements are substantial.

Many transportation agencies see big data and its applications as an opportunity to improve the management and operation of transportation systems, increase the accuracy of prediction, enable informed decision making, and optimize transportation services. However, big data are only useful if policy questions are framed correctly and if available datasets are relevant to those questions.[5] Further, data collections are becoming too big, growing too fast, or too complex for existing information technology systems to handle them. The challenges include capture, curation, storage, security, search, sharing, transfer, analysis and visualization. The path forward will be a mix of technology, public policy, and consumer acceptance.

2.2 MULTI-SOURCE AND OPEN DATA ENVIRONMENTS

Variety is one of the key attributes of big data, which becomes possible because of multisource data environments. When freely available for people to obtain, use, and redistribute, big data becomes open data. One of the biggest enablers of intelligent transportation is an open-data approach.

MDOT creates and manages a vast amount of data, and they are increasingly sharing data with private sector users and the general public. Multi-source transportation data is reflected by:

- Multiple modes – highway, arterial, public transit, and freight system performance data

---

[4] OECD 2013.
[5] International Transportation Forum 2014.

- Multiple instruments – smartphones, sensors, social media, and specially equipped vehicles (Radar, Lidar, and other instruments) that enable continuous collection, communication, and processing of mobility data; anything from traffic and weather conditions to parking spots and rideshares.
- Multiple agencies – DOTs, transit operators, freight service providers
- Multiple data types – traffic speeds/volumes/occupancies, incidents, transit supply, asset condition, and performance

The Real-Time Data Capture and Management Program is one of the connected vehicle activities managed by the Joint Program Office (JPO) of the Research and Innovative Technology Administration (RITA) of the USDOT. The goal of the program is to establish multi-source, multi-modal environments, in which data from diverse sources can be used concurrently. For example, traveler information, private vehicle, transit vehicle, heavy vehicle, infrastructure, weather, and parking data in such an environment might be combined into one data warehousing effort in which users can obtain a coherent and concurrent view of the transportation system.[6] A required element of the USDOT Connected Vehicle Safety Pilot Model Deployment is the systematic collection of multisource data via Dedicated Short Range Communications (DSRC) and wireless access for vehicular environments (WAVE).

2.3 MOBILE DATA COMMUNICATIONS

Velocity is another key attribute of big data, which means data capture, processing, and transmission often take place nearly instantaneously. The concepts of *data in motion vs. data at rest, snapshot vs. continuous, and send vs. receive* are useful to explain velocity or mobile data communications.

DATA IN MOTION VS. DATA AT REST

"Data at rest" refers to static data that have been collected from various sources and are analyzed after the event occurs.[7] For example, traffic engineers generally analyze vehicle travel time and speed data, and develop congestion mitigation measures, using static data. The action takes place after

---

[6] USDOT RITA 2011.
[7] Nikki Nixon 2013.

the data collection was complete. Most current transportation data environments are archival in nature. Data captured and managed in these environments tend to be collected over time, assessed for quality and potentially aggregated at some intermediate point, and then at a later date (days, months or even years later) made available to researchers or other interested users.

The second state is called "data in motion", sometimes referred to as "data in transit", and it is the process of analyzing data on the fly without necessarily storing it. In many cases, data in motion will become basic features of connected and automated vehicle systems since they require hyper local and hyper current data flows and evolution to support targeted applications.

Data in motion often requires advanced computing platforms or real-time processing capabilities to glean insights from vast situation data. According to USDOT, the notion of what constitutes “real-time” data provision is based on two considerations. The first is the data capture interval or how frequently data is collected. The other is data latency or the time lag between when the data are collected and the time data can be shared with users. The Real-Time Data Capture and Management Program seeks to support development and testing of new or enhanced applications enabled by collecting new or existing types of data at more frequent intervals and enabling these data to be shared more rapidly.

Data latency among current transportation data management efforts varies widely, including its use in an asset management approach for larger versus smaller agencies. For some large data collection efforts, the latency could be months or years (e.g., the National Household Travel Survey, hardware asset management such as signs, guardrail, and culverts). Focused data collection activities that capture weather or travel time information from roadway sensors can be provided within minutes or even seconds of collection. Such latency is still appropriate for many current transportation operation and management applications. The proposed connected vehicle safety applications require less than 10 millisecond latency because a lag in data transmission and processing could miss the opportunity to support the targeted applications.

SNAPSHOT VS. CONTINUOUS

Another distinction of big data relates to the rate of change of a data set. Traditional data are generally snapshot data; data that describe a system state at a specific point in time. Handling big data often requires working with continuous data. Continuous data streams can provide near real-time metrics while they evolve over space and time, but they often require sophisticated algorithms to analyze and sufficient processing power to perform such analysis. In other words, it is difficult to turn large amount of continuous data into useful, actionable information.

SEND VS. RECEIVE

A DSRC/WAVE-connected vehicle broadcasts a basic safety message (BSM) ten times each second. The vehicle simultaneously receives BSMs from nearby connected vehicles. The BSM includes data elements (e.g., vehicle's ID, sequence number, time, location, velocity, and acceleration) at each 1/10th-second interval. The vehicle may also broadcast additional information as conceptualized in the BSM II, such as light status, wiper status, etc.

If automated vehicles are deployed *without* being DSRC/WAVE-enabled, they will not be continually broadcasting information in the way that a connected vehicle does. Automated vehicles deployed without DSRC connectivity will be more likely to be users of data and will likely rely on a cellular network link for map updates and other information. Depending on the eventual deployment concepts, automated vehicles may or may not broadcast information about the vehicle location and status.

With a significant increase in on-board computing power and the number of sensors collecting data, connected and automated vehicles will likely generate an even greater volume of data over time. They will also have the capabilities to either ingest and analyze data in real time or work with massive quantities of data. This data is likely to be utilized in a more connected way to lessen congestion, improve safety, reduce fuel consumption, and more effectively assist people's travel in general.[8]

The next chapter will discuss the definitions of connected and automated vehicles (CAV) and their enabling technologies. Chapter Four will discuss data environment and data elements associated with CAV applications.

---

[8] Morgan Stanley 2013.

3 CONNECTED VS. AUTOMATED VEHICLES

3.1 DEFINITIONS

As considered within the USDOT Connected Vehicle Research Program, connected vehicle systems focus on localized vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and vehicle-to-everything (V2X) systems communication to support safety, mobility and environmental applications. Their foundation is data communications that enable real-time driver advisories and warnings of imminent threats and hazards on the roadway. DSRC/WAVE standards are currently the leading medium for connected vehicle safety communication.[9] A large percentage of the USDOT list of non-safety critical V2I applications could be achieved via cellular (i.e. traveler info, dynamic routing, etc.). The development of the connected vehicle environment will serve the public good in a number of ways:

- Highway crashes could be dramatically reduced when vehicles can sense and communicate the events and hazards around them;
- Mobility will be improved when drivers, transit riders, and freight managers have access to substantially more up-to-date, accurate, and comprehensive information on travel conditions and options;
- Environmental impacts of vehicles and travel can be reduced when travelers can make informed decisions about modes and routes and when vehicles can communicate with the infrastructure to enhance fuel efficiency by avoiding unnecessary stops. [10]

The automobile industry, as well as some technology companies such as Google, also is working to develop automated vehicle (AV) technologies, which would control steering, acceleration, and braking without a driver's input.[11] At certain levels, AVs are designed to drive from point-A to point-B without direct manual control from the driver. The vehicle uses a combination of cameras, lidar, radar, sensors, positioning systems, and digital maps to determine its surroundings and uses artificial intelligence to navigate the vehicle to its destination.

---

[9] ITS America 2014.
[10] AASHTO 2013.
[11] GAO 2013.

Some studies suggested that there is a continuum between conventional, fully human-driven vehicles and AVs, which partially or fully drive themselves and may require no driver at all. NHTSA has created a hierarchy to help clarify this continuum. [12]

- Level 0 (no automation): The driver is in complete and sole control of the primary vehicle functions and is solely responsible for monitoring the roadway and for safe vehicle operation.
- Level 1 (function-specific automation): The driver has overall control but can choose to cede limited authority over a primary control (e.g., adaptive cruise control).
- Level 2 (combined-function automation): This level involves automation of at least two primary control functions designed to work in unison to relieve the driver of controlling those functions.
- Level 3 (limited self-driving automation): Vehicles at this level of automation enable the driver to cede full control of all safety-critical functions under certain traffic or environmental conditions, and in those conditions to rely heavily on the vehicle to monitor for changes in those conditions requiring transition back to driver control.
- Level 4 (full self-driving automation): The vehicle is designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip. By design, safe operation rests solely on the automated vehicle system (NHTSA, 2013).

CONVERGENCE OF CV AND AV: CONNECTED AUTOMATION

The level-4 automation systems currently being developed will require some means of connecting to the internet for map information, software updates, and access to live information such as traffic conditions. Prototype automated vehicles are not dependent on DSRC/WAVE connected vehicle systems. However, many stakeholders within the ITS industry believe that DSRC/WAVE connectivity will be required to achieve reliable fully automated (driverless) systems.

---

[12] NHTSA 2013.

While automated vehicles' dependence on dedicated infrastructure is much lower than it was in the early prototype stages several years ago, they may still need some basic level of infrastructure including road markings and signage, GPS mapping, and some level of vehicle-to-grid (V2X) communication. Some studies suggest that lack of infrastructure could be a challenge to accelerating penetration of automated vehicles.[13]

The selection and use of technologies will influence the data to be generated in connected and automated vehicle applications. Below we will discuss key enabling technologies for connected and automated vehicle systems.

3.2 CONNECTED VEHICLE ENABLING TECHNOLOGIES

Key technologies necessary to enable connected vehicle applications include onboard equipment, roadside equipment, communication systems, digital mapping and positioning technologies, computing and management systems.[14]

Onboard equipment (OBE), or mobile equipment, consists of devices (effectively similar to Wi-Fi radios) and systems present (installed during assembly or added later) in vehicles through which vehicle operators will interact with the connected vehicle environment. Other technologies, such as GPS and the vehicle's data bus, often are needed to provide basic information used in connected vehicle applications, including location, speed, and heading.

Also consisting of radio communication devices, roadside equipment (RSE) provides connectivity between vehicles and roadside systems and provides the foundation for V2I deployment, such as systems integrated with traffic signal controllers used for Signal Phase and Timing (SPaT) applications. Devices compliant with the 5.9 GHz DSRC RSE specifications are required for many connected vehicle applications, though some can operate through cellular networks.

Communications systems infrastructure is needed for V2I connectivity and might be necessary for certificate management in a V2V environment. DSRC technology was developed specifically for connected vehicle communications. It operates at 5.9 GHz frequency, using standards such as Society of

---

[13] Morgan Stanley 2013; MDOT and Car 2012.

[14] USDOT 2013.

Automotive Engineers (SAE) J2735 and the Institute of Electrical and Electronics Engineers (IEEE) 1609 suite. DSRC offers the greatest promise because it provides security, a fast network acquisition, low-latency, high-reliability communications link. DSRC also has the ability to work with vehicles operating at high speeds, prioritize safety messages, tolerate multipath transmissions, and is immune to extreme weather conditions (e.g., rain, fog, snow). For many non-safety applications, existing cellular technologies can be used.

Support systems facilitate interactions among vehicles, field infrastructure, and back office users. The system also includes the security credentials management that allows devices and users in the Connected Vehicle Environment to establish trust relationships. Most applications require back office systems for data processing, storage, retrieval, and end-user presentation using cloud computing and crowdsourcing technologies.

For certain safety applications, such as lane change warning and pedestrian crossing detection, positioning of the vehicle relative to the lane edges or pedestrian must be known. In some cases, this might require absolute position information matched to high-resolution maps stored in the vehicle, while in other cases relative position is sufficient to avoid imminent contact. This allows the vehicle to be located relative to its surroundings. Preliminary results from the connected vehicle model deployment Safety Pilot indicate that Wide Area Augmentation System-enabled (WAAS) GPS accuracy (about 10 meters) is insufficient for pedestrian crossing warning application.[15] Significant improvement could be achieved with more accurate technology, such as differential GPS systems for specific localized implementations.

3.3 AUTOMATED VEHICLE ENABLING TECHNOLOGIES

Like most robotic systems, automated vehicles use a “sense-assess-act” design. In-vehicle components of an automated driving system usually include active sensing technologies that gather information about its surroundings, sophisticated algorithms/artificial intelligence that process sensor data and

---

[15] USDOT Webinar: Transit Safety and Mobility Applications in a Connected Vehicle World. May 14, 2014

control the vehicle, and computational power that runs the vehicles.[16] Below are brief discussions of each of these technology components:

Cameras: Automotive monovision cameras are usually equipped with the software that can fast process the images and recognize common roadside infrastructure (lane markings, speed limit signs, etc.). More advanced stereovision cameras use two video sources, similar to human eyesight. This incorporates depth perception and helps the car better understand the relative position of moving traffic and potential obstacles.[17]

Radar: Automated vehicles often employ radar systems for range-finding applications. Two typical types of radar systems are in use, short and long-range. Short-range radar "feels" around the car's immediate surroundings, especially at low speeds, while long-range radar is used at high speeds and over relatively long distances. Normally, one device cannot do both.

Lidar: Lidar uses laser detection ranging to create a 3D profile of the surroundings around the car. One of the issues faced by this system in real life is that temporary changes (like snow or new traffic patterns) could disrupt the surrounding profile.

Ultrasonic: Sonar sensors are used to detect objects and calculate their distance, direction, and speed. They are also used extensively to understand what is happening with the car itself, including acceleration sensors, pressure sensors, light sensors, etc.

Human-machine interface (HMI): The HMI refers to the combination of systems inside the vehicle, including the infotainment/entertainment system, instrument panel, and controls that act as an interface between the car and the driver. When there is a human driver in control, the HMI must optimally manage driver distraction, and drivers must be thoroughly comfortable with the transition process between manual and automated driving (hand off control and take it back).

Domain controller: The domain controller functions as the hardware “brain” of the automated driving system. It decides what action is to be taken based on inputs received from the various sensors. It passes instructions to mechatronic

---

[16] RAND Corporation 2013.
[17] Morgan Stanley 2013.

units/actuators that physically control the drivetrain components, such as the steering wheel, throttle, brakes, suspension, etc.

Wireless Data Networks: Automated vehicles likely will need reliable, high-speed two-way data communication networks for navigation and content reception. This may include Long-term Evolution (LTE) networks, DSRC, WiFi, or some other wireless communications protocol.

GPS: GPS receivers pinpoint where the automated vehicle is and where it is heading on the map. For localization, the vehicles can combine GPS with correction methods such as inertial navigation systems (INS), as well as sensors for relative position in regards to other vehicles and objects. The current GPS deployment can provide an accuracy of 7.8 meters at a 95% confidence level. Real-world data from the Federal Aviation Administration (FAA) show that its high-quality GPS standard positioning service (SPS) receivers provide better than 3.5 meters horizontal accuracy.[18] Research suggests differential GPS (DGPS) may achieve lane-level accuracy in the near future with accuracies of 1.5 meters 95% of the time.[19] The requirement for vehicle positioning varies by application. For example, an accuracy of 5–10 meters is required to warn drivers of hazard at a fixed location (e.g., a crash site). By contrast, applications in small spatial and temporal areas, such as motion control, especially crash avoidance, require 1-2 meters accurate positioning in real time.[20]

DSRC: Many industry insiders believe that a mature DSRC connected vehicle network will be necessary to achieve fully automated vehicles. Others suggest that DSRC connectivity may not be necessary.[21]

3.4 COMMON AND DISTINGUISHING TECHNOLOGIES

The enabling technologies for connected vehicles and automated vehicles are generally quite different. The prototype core connected vehicle OBE includes only the DSRC transponder, GPS receiver, a small processor, and a simple HMI interface to provide warnings and other information. An automated vehicle may include the DSRC OBE, but most current research and

---

[18] http://www.gps.gov/systems/gps/performance/accuracy/
[19] Goodall 2014
[20] Wang 2014
[21] RAND Corporation 2013.

development projects are not assuming the existence of a nationwide DSRC network. Automated vehicles will likely include a GPS receiver, and may or may not continually broadcast the vehicle location and other information by cellular networks to back-office systems for support and services. As vehicles move toward more automated driving, significantly more computing power is necessary for data and threat analysis and vehicle control.

The prototype core connected vehicle is based on the DSRC OBE. Additional sensors may be included on the vehicle, but not necessarily so. Automated vehicles, with or without DSRC connectivity, can be expected to have a complex sensor suite in order to accurately sense the driving environment. Automated vehicles will likely have cameras with complex machine-vision software, as well as various range-finding sensors such as radar, sonar, and lidar. The key enabler of high-level automation will be accurate digital 3-D maps of the road network. Lidar and cameras can be fitted to automated vehicles that allow the vehicle to generate and automatically update the basemap. However, consumer-deployed automated vehicles may not include map-making capability, in order to reduce system cost and complexity.

The comparison of connected and automated vehicle enabling technologies is presented in Table 1.

TABLE 1: COMPARISON OF CONNECTED AND AUTOMATED VEHICLE TECHNOLOGIES

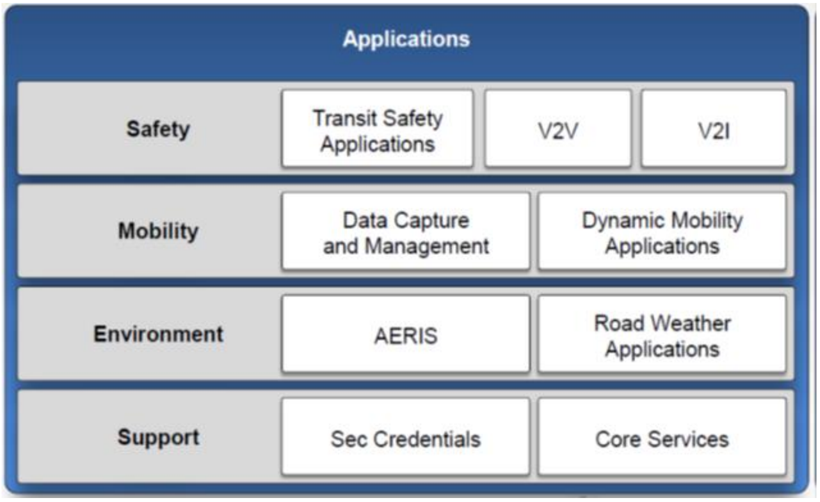
Technology	Connected Vehicle	Automated Vehicle	Connected & Automated Vehicle
Radar		✓	✓
Cameras		✓	✓
Lidar		✓	✓
GPS	✓	✓	✓
Cellular		✓	✓
DSRC	✓		✓
Onboard Equipment and Computing	✓	✓	✓
Roadside Equipment	✓		✓
Back-office Systems	✓	✓	✓

4 CONNECTED AND AUTOMATED VEHICLE DATA SYSTEMS

Connected vehicle data systems will be influenced by applications types, interface design and data exchange architecture. Below we will discuss connected vehicle applications, the data environment concept, physical architecture and interface design, and data systems related to the vehicle, infrastructure, traffic management, and travelers.

4.1 CONNECTED VEHICLE APPLICATIONS

The USDOT has made a significant investment in foundational research and initial development of connected vehicle applications. Current research efforts focus on safety, mobility and environmental applications through interoperable wireless communications among vehicles, infrastructure, and personal communications devices (Figure 2). Currently a few connected vehicle application efforts are complete. The proposed applications are not in the same state of maturity. However, according to USDOT ITS JPO, a large number of application development efforts across multiple programs will be substantively complete in late 2014.



Applications			
Safety	Transit Safety Applications	V2V	V2I
Mobility	Data Capture and Management	Dynamic Mobility Applications	
Environment	AERIS	Road Weather Applications	
Support	Sec Credentials	Core Services	

FIGURE 2: CONNECTED VEHICLE APPLICATIONS[22]

---

[22] Source: USDOT ITS JPO Connected Vehicle Pilot Deployment Workshop. 2014.

4.2 CONNECTED VEHICLE DATA ENVIRONMENT CONCEPT

As part of the connected vehicle research efforts, the USDOT ITS JPO is engaged in assessing the potential of the multi-source, active-acquisition data paradigm to enhance planning, operation, and management of future transportation systems. One core concept in the Real-Time Data Capture and Management Program is the *data environment*. A data environment is defined as:

- A well-organized collection of data of specific type and quality,
- Captured and stored at regular intervals from one or more sources, and
- Systematically shared in support of one or more applications.

A data environment is made up of one or more data sets systematically captured from connected vehicles (automobiles, buses, trucks, and fleets), mobile devices, travelers, and infrastructure. Data elements may include real-time data, or a combination of observed, derived, and/or simulated data from a broad spectrum of data sources. The data environment concept is shown in Figure 3.

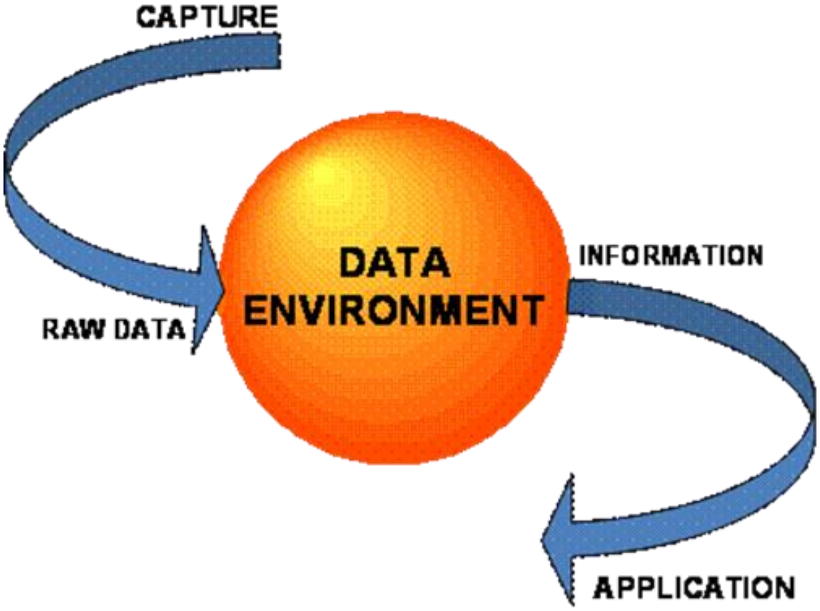


FIGURE 3: THE DATA ENVIRONMENT CONCEPT[23]

---

[23] Source: U.S. DOT ITS JPO. Real-Time Data Capture and Management Program Vision: Objectives, Core Concepts and Projected Outcomes. 2010.

According to SAE International, data are defined as representations of static or dynamic entities in a formalized manner suitable for communication, interpretation, or processing by humans or by machines. Connected vehicle datasets are defined in the DSRC standard. Data element, data frame, and message are three interrelated data concepts often seen in a connected vehicle environment.[24]

Data Element (DE) is a syntactically formal representation of some single unit of information of interest (fact, proposition, observation, etc.) with a singular instance value at any point in time, about some entity of interest (e.g., a person, place, process, property, object, concept, association, state, and event). A data element is considered to be indivisible. The types of data elements (integers, enumerations, short strings) are representative of the content found and anticipated in the future editions of the standard.

Data Frames (DF) are viewed as logical groupings of data elements to describe structures or parts of messages. A data frame is a collection of one or more other data concepts in a known ordering. These data concepts may be simple (data elements) or complex (data frames).

Message is a well-structured set of data elements and data frames that can be sent as a unit between devices to convey some semantic meaning in the context of the applications about which this standard deals, such as the basic safety message (BSM) proposed in connected vehicle safety applications.

4.3 CONNECTED VEHICLE PHYSICAL ARCHITECTURE AND DATA EXCHANGE

Connected vehicle technology needs to be interoperable between vehicles from different manufacturers and across different infrastructure implementations. To achieve this goal, a framework is needed that effectively integrates technologies and provides data exchange interfaces. The Connected Vehicle Reference Implementation Architecture (CVRIA) is based on a set of applications that have been defined by various connected vehicle programs. The goal is to identify key interfaces across the connected vehicle

---

[24] SAE International 2010.

environment and prioritize standards development activities.[25] CVRIA is developed in four viewpoints:

- Enterprise - Describes the relationships between organizations and the roles those organizations play within the connected vehicle environment
- Functional - Describes abstract functional elements (processes) and their logical interactions (data flows) that satisfy the system requirements
- Physical - Describes physical objects (systems and devices) and their application objects as well as the high-level interfaces between those physical objects
- Communications - Describes the layered sets of communications protocols required to support communications among the physical objects that participate in the connected vehicle environment

The physical architecture is the most relevant viewpoint to this study since it represents data flows and information exchanges in a connected vehicle environment. The broad types of physical classes that will interact and exchange information to support the connected vehicle applications are: centers, field, travelers, vehicles, and support. Each physical class includes several more specific application objects (Table 2):

---

[25] www.iteris.com/cvria/index.html

TABLE 2: COMPONENTS OF USDOT CONNECTED VEHICLE PHYSICAL ARCHITECTURE[26]

Class	Physical Object
Vehicles	Commercial Vehicle
	Emergency Vehicle
	Maintenance and Construction Vehicle
	Transit Vehicle
	Vehicle
Field	Commercial Vehicle Check
	Parking Management
	Roadway
	Roadway Payment
	Security Monitoring
Centers	Archived Data Management
	Commercial Vehicle Administration
	Emergency Management
	Emissions Management
	Fleet and Freight Management
	Information Service Provider
	Maintenance and Construction Management
	Payment Administration
	Traffic Management
	Transit Management
Travelers	Personal Information Access
	Remote Taveler Support
Support	Security and Credentials Management System

The USDOT Test Bed in Southeast Michigan recently implemented the first CVRIA compliant system. This implementation is intended to demonstrate the operations of a data management system for connected vehicle data. The system (diagrammed in Figure 4) defines the following data flows:

---

[26] Adopted from USDOT "National ITS Architecture" and "Connected Vehicle Reference Implementation Architecture (CVRIA)" http://www.iteris.com/itsarch/html/entity/paents.htm, http://www.iteris.com/cvria/html/physobjects/physobjects.html

- Data flows between vehicles and back office services using roadside dedicated short range communications (DSRC) radios and other Internet Protocol radio services;
- A back office data management system that validates messages and distributes them to both public and private sector applications (includes the data flows between the data management device and the applications);
- Input and storage data flows for a data warehouse that stores the most current relevant data for users; and
- Data flows to return current relevant data to users, either through the DSRC radio system or via other communications media.

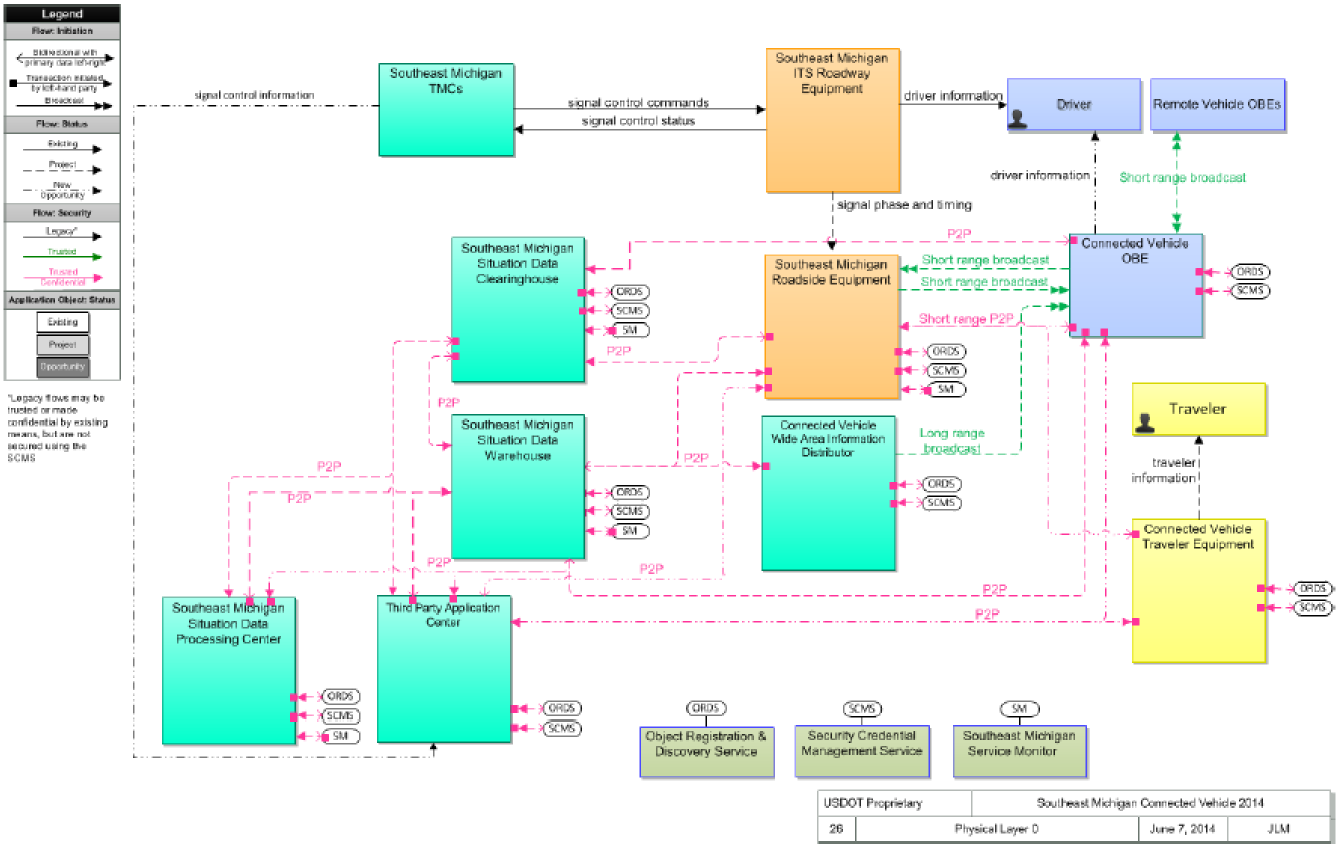


FIGURE 4: USDOT SOUTHEAST MICHIGAN CONNECTED VEHICLE ARCHITECTURE, PHYSICAL LAYER 0.

4.4 CONNECTED VEHICLE DATA SYSTEMS

In a connected vehicle environment, situational data about vehicles, infrastructure, and travelers are continuously generated and exchanged. The sections below discuss connected vehicle-, infrastructure-, traffic management-, and traveler-based data elements. Overlapping of data elements across these subsystems is likely.

VEHICLE DATA

Moving vehicles are increasingly used as data sources for numerous purposes. In a connected vehicle environment, the vehicle subsystem (VS) provides the sensory, processing, storage, and communications functions necessary to support efficient, safe, and convenient travel. These functions reside in various types of vehicles, including automobiles, and commercial, emergency, construction, maintenance, and transit vehicles.

Advanced sensors, processors, enhanced driver interfaces, and other on-board units (OBU) are able to record and deliver the data through wireless networks. The data include basic vehicle measures, vehicle safety data, environmental probe data, vehicle diagnostics data, and vehicle emissions data. Specific data elements from connected vehicles include, but are not limited to:

- Vehicle type and characteristics (length, width, bumper height)
- Time stamp
- Speed and heading
- Vehicle acceleration and yaw rate
- Turn signal status
- Brake status
- Stability control status
- Driving wheel angle
- Vehicle steering
- Tire Pressure
- Traction control state
- Wiper status and run rate
- Exterior lights
- GPS status and vehicle position (longitude, latitude, elevation)
- Obstacle direction
- Obstacle distance

- Road friction
- Current and average fuel consumption
- Vehicle emissions data - measured emissions of specific vehicles comprised of exhaust pollutants including hydrocarbons, carbon monoxide, and nitrogen oxides
- Air temperature and pressure
- Weather information such as rainfall rate and solar radiation data
- Electronic Stability Control

INFRASTRUCTURE DATA

The infrastructure subsystem is a critical component of the connected vehicle environment. The subsystem could include specific DSRC-capable roadside equipment (RSE) and more traditional ITS equipment distributed on and along the roadways, such as traffic detectors, environmental sensors, traffic signals, highway advisory radios, dynamic message signs, Closed Circuit Television (CCTV) cameras and video image processing systems, grade crossing warning systems, and freeway ramp metering systems.

The infrastructure subsystem is able to provide and exchange data elements related to roadway characteristics, road conditions, intersection status, and field equipment status. These characteristics are monitored or measured by RSE and ITS sensors to support advanced vehicle safety, traffic control, work zone management, and road construction and maintenance activities. Specific data elements include, but are not limited to:

Roadway characteristics

- Friction coefficient
- Road geometry and markings

Road conditions

- Surface temperature
- Subsurface temperature
- Moisture
- Icing
- Treatment status

Road surface weather conditions

- Air temperature

- Wind speed
- Precipitation
- Visibility

Intersection status

- Current operational status
- Signal phase and timing
- Intersection geometry
- Approaching vehicle information (position, velocity, acceleration, and turning status)

Field equipment status

- Dynamic message signs
- Variable speed limit signs
- Dynamic lane signs or control devices
- Ramp meters

Parking information

- Location of parking facilities
- Spaces available

TRANSPORTATION MANAGEMENT DATA

Connected and automated vehicles are capable of reporting current vehicle position, speed, heading, and snapshots of recent events including route information, starts and stops, speed changes, and other information that can be used to estimate traffic conditions and support transportation planning and asset management. The information can be delivered to the traffic management centers (TMC) or the traffic management subsystems as defined in ITS Architecture.

In a connected vehicle environment, a TMC or transportation operations center (TOC) manages a broad range of transportation facilities, monitors and manages traffic, and provides support to travelers to achieve certain traffic management strategies (e.g., mode selection, route choice, time of travel, etc.). A TMC provides the capability to exercise control over ITS devices that are used for automated highway system (AHS) traffic and vehicle control purposes.

The traffic condition measures or macroscopic traffic characteristics derived from connected vehicles include:

- Traffic speed
- Travel times
- Volumes
- Occupancy
- Density
- Origin and destination data for vehicles that opt to provide this information
- Incident status
- Video images

TRAVELERS' DATA

In a connected vehicle environment, travelers continuously interact with the vehicles, centers, and support subsystems to allow them to receive current transportation information and help them make smarter travel choices. For example, The Personal Information Access Subsystem (PIAS) provides the capability for travelers to receive formatted traffic advisories from their homes, place of work, major trip generation sites, personal portable devices, over multiple types of electronic media. These capabilities also provide basic routing information and allow users to select those transportation modes to avoid congestion, or more advanced capabilities to allow users to specify transportation parameters that are unique to their individual needs.

This subsystem provides travelers with the capability to receive route planning from the infrastructure at fixed locations such as in their homes, their place of work, and at mobile locations using personal portable devices and vehicle-based devices. In addition to end user devices, this subsystem may also represent a device that is used by a merchant or other service provider to receive traveler information and relay important information to their customers. This subsystem also provides the capability to initiate a distress signal and cancel a prior-issued manual request for help.

Travel information that can be generated through these interactions include:

- Trip origin, destination, and timing for vehicles that opt to provide this information
- Traveler's personal data such as address, trip records, and profile data for vehicles that opt to provide this information

- Service information (payment of tolls, parking reservations and fees, ridesharing options etc.) for vehicles that opt to provide this information
- Vehicle occupancy - The number of occupants detected by the vehicle
- Vehicle Miles Travelled (VMT) data by vehicle characteristics, time, and location

4.5 BASIC SAFETY MESSAGE DATA

The Society of Automotive Engineers (SAE) has developed standard J2735 to document the data formats within a connected vehicle environment.[27] Within the J2735 standard is the concept of the Basic Safety Message (BSM). The BSM is a core set of Over the Air (OTA) messages that supports safety applications developed under the Vehicle Safety Communications – Applications (VSC-A) Project by Crash Avoidance Metrics Partnership (CAMP) and NHTSA. This set of messages is further refined into two subsets, BSM Part I and BSM Part II. Data elements within the BSM Part I consist of vehicle status data, which are safety critical and must be included in every BSM broadcast from a vehicle, at a rate of 10Hz or ten times per second. Data elements within BSM Part II are either required by applications at regular intervals, but at a reduced frequency, required to notify applications of a given event for optional applications outside of the VSC-A project, or optional elements.[28]

Figure 5 shows the messages format of the BSM.

---

[27] SAE International, 2009

[28] USDOT, 2011

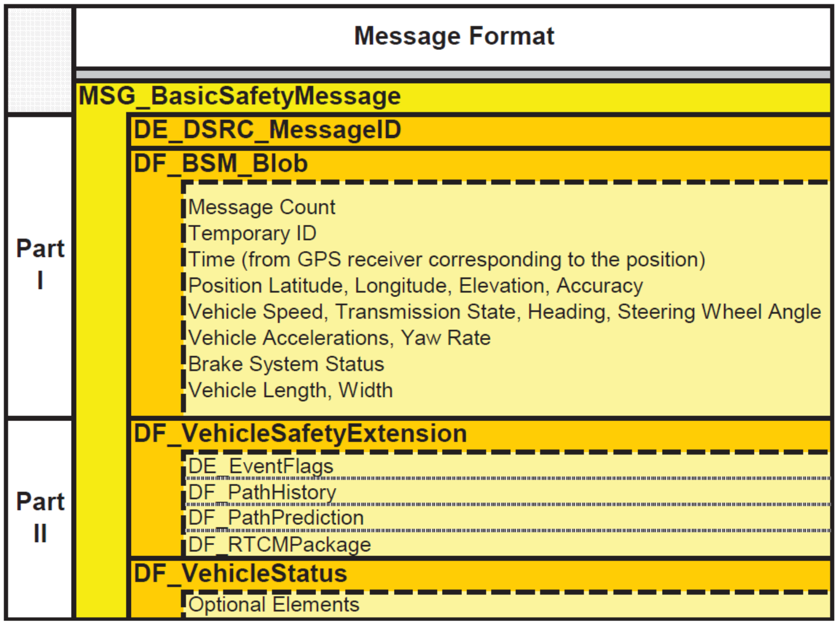


Figure 5: BSM Message Format.[29]

Modern vehicles are equipped with a specialized internal communications network called a vehicle bus, which interconnects various components of a vehicle and shares data between them. Indeed, most have more than one, though the controller area network (CAN) is common to most light vehicles in the U.S. The vehicle bus contains a variety of data that can be used to populate the BSM. However, the ability to access this data varies by manufacturer and model. Additionally, vehicle networks continue to become more complex as vehicles are equipped with more technology and often contain multiple buses to support advanced functionality.

For the USDOT Southeast Michigan Test Bed, two vehicles were purchased in 2007, a Ford Edge and a Jeep Grand Cherokee. These vehicles gave the project team greater access to data elements available on the CAN bus; a standard message-based protocol for vehicle buses. Table 3 shows sample data elements available from a single CAN bus on each vehicle that were used to populate BSMs within the Southeast Michigan Test bed.

---

[29] USDOT FHWA, 2012

TABLE 3: USDOT TEST BED VEHICLE BSM AVAILABILITY.

Jeep	Ford
BSM Part I	**BSM Part I**
BrakeSystemsStatus\StabilityControlStatus	AccelerationSet4Way
SpeedAndTransmission	SteeringWheelAngle
SteeringWheelAngle	BrakeSystemsStatus\BrakeAppliedStatus
AccelerationSet4Way	SpeedAndTransmission
BrakeSystemsStatus\BrakeAppliedStatus	BrakeSystemsStatus\TractionControlState
AccelerationSet4Way	BrakeSystemsStatus\TractionControlState
	BrakeSystemsStatus\StabilityControlStatus
	BrakeSystemsStatus\AntiLockBrakeStatus
BSM Part II	**BSM Part II**
ExteriorLights	ExteriorLights
ThrottlePosition	ThrottlePosition
WiperStatusFront	AmbientAirTemperature
AmbientAirTemperature	AmbientAirPressure
AmbientAirPressure	

In working to obtain this data, the project team found only a limited amount of information was available from the CAN Bus. A significant effort was undertaken to reverse engineer the CAN bus data on the On-Board Diagnostics (OBD-II) port to enable the Test Bed team to collect and use that information. The potential NHTSA Rule for V2V connectivity might well address the need for data on the CAN bus by requiring, at a minimum, BSM Part 1 data to be available and accessible on the vehicle network. Because vehicle buses vary across manufactures and models, aftermarket solutions to obtain CAN bus data may not provide access to all of the information required in BSM Part 1. This would affect the potential of those vehicles to be fully integrated into a connected vehicle system and support application development.

The USDOT performed an analysis of the utility of the BSM for use in mobility applications as part of their Dynamic Mobility Applications (DMA) program. They had two primary findings[30]:

1. The BSM is useful for a limited subset of mobility applications, but it is not solely sufficient as it is not intended that there will be complete roadway coverage using DSRC, so an alternate means of communicating with the vehicle is required if data is required in locations where there is no DSRC coverage.
2. The data in BSM Parts I and II can adequately provide the vehicle-based information needed for most mobility applications if the data is cached, bundled, or sent over an alternate ubiquitous communications media.

4.6 AUTOMATED VEHICLE GENERATED DATA

The USDOT connected vehicle program has long focused on designing the system to provide information to a variety of traffic and transportation related applications. In contrast, prototype automated vehicles currently in development are not designed for the purpose of providing any data or information useful for administration of the road network and broad transportation system. Automated vehicles are not currently being designed for broader interoperability with public ITS systems. If automated vehicles generate any data, it will likely be proprietary data for the use of the vehicle owner, operator, manufacturer, or service provider.

The primary data component that automated vehicles will produce is three dimensional digital image data, which concurrently enables the operation of the automated driving system. The base map will be useful mainly to the vehicle itself. However, with some additional processing, the digital basemap could likely be used to track traffic and infrastructure conditions in near real-time. There are currently not any proposed business models that would include a publically available automated vehicle base map. Future business models may include this, or may provide an opportunity to purchase data from third-party automated driving service providers.

---

[30] USDOT FHWA, 2012

AUTOMATED VEHICLE DATA SYSTEMS

The type of data generated by an automated passenger vehicle depends on the level of automation supported. As mentioned, automated vehicles use a "sense-assess-act" design. The degree to which a vehicle must perform "sense-assess-act" functions is correlated to the level of automation and determines the type and amount of data collected, processed, and utilized by the vehicle. An automated (but not connected) vehicle is "selfish" in the sense that the automated vehicle only collects information related to the host vehicle and makes decisions accordingly. Therefore, data collected or generated by an automated system are primarily related to the host vehicle's surrounding environment, current state of the host vehicle, and decisions made by the automated system.

CURRENT MARKET

Driver control versus system control is the primary factor in differentiating between lower and higher levels of automation. Partial automation systems (i.e., Level 1 and 2) depend primarily on the driver for vehicle control and monitoring of the environment with assistance from the system. Conversely, the system is primarily responsible at higher levels of automation. Partially-automated vehicles equipped with driver assistance systems such as adaptive cruise control (ACC) and crash-imminent braking (CIB) are currently available on the market and are supported by Level 1 and Level 2 automation technologies. These technologies may perform tasks for the driver in controlled environments (e.g., manage the gap to the vehicle ahead while using cruise control on the highway) or take action for the driver in safety critical situations (e.g., applying the brakes automatically to prevent a potential rear-end collision when a stopped vehicle is detected ahead).

Vehicles with Level 1 systems like ACC and CIB are equipped with sensors such as vision (i.e., camera), radar, and ultrasonic. Driver assistance systems vary by manufacturer but most systems combine data collected by sensors (i.e., sensor fusion) and process it to determine when automated functions should be performed. Vision sensors may provide object detection and classification information (i.e., passenger car, motorcycle, pedestrian, fixed object), while radar sensors provide object position information such as the range, range rate, and angle of a forward object. The system may use radar for larger objects at higher speeds and ultrasonic sensors to detect objects at lower speeds. An automated system may be comprised of multiple components:

- a “sense” component which collects information from surrounding environment,
- an “assess” component which combines and processes the sensor and other data available on the vehicle, and
- an “act” component (i.e., domain controller) which monitors all the data available, houses the logic which determines if automated control is necessary, and interfaces with the vehicle to control the vehicle if needed.

For partial automation systems like ACC and CIB, the type and usefulness of the data may vary with each component of the system. The “sense” component can provide raw data from the radar like the range to the vehicles and/or objects ahead, which likely has limited use outside the context of the host vehicle system. Data fused together by the “assess” component can provide information such as the distance to the most forward object or the number and types of vehicles ahead. This data could be shared with surrounding vehicles to support future connected vehicle concepts such as Cooperative Adaptive Cruise Control where the distance to a vehicle ahead is essential. The number and types of vehicles ahead could be used to provide a localized snapshot of the mix of traffic or volume, which when assessed from multiple vehicles could provide system insight into traffic conditions. The “act” component can generate data related to decisions made by the automated system such as applying the brakes to avoid a collision in a safety critical situation. This could support queue warning or other traveler information applications.

FUTURE MARKET

Private companies and universities around the country are developing and testing prototype vehicles with Level 3 automation capabilities that are not currently available to consumers. Systems supporting higher levels of automation also require collecting more detailed information about the surrounding environment of the host vehicle.

Level 3 and above automated systems may also rely on multiple sensors like partial automation but incorporate more advanced sensors such as Lidar to gather more detailed information about the host vehicle’s surrounding environment. These automated systems may also rely on positioning systems and communication with backend technologies for navigation and up-to-date location-based content.

Lidar is one of the primary methods being used today for autonomous vehicle navigation and mapping of the area surrounding the vehicle. Systems like the Velodyne HDL-64E S2[31] can generate over 1.3 million data points per second, identifying objects on the road ahead and threats to the side or rear of the vehicle. From this data, control algorithms can be developed for the autonomous vehicle to permit autonomous driving by mapping out a safe path and identifying threats. An example of the processed output from a Lidar system is shown in Figure 6 and Figure 7. As Lidar systems for transportation applications continue to mature and as the autonomous driving industry continues to evaluate their needs for resolution, the potential output of Lidar systems will continue to evolve as well.

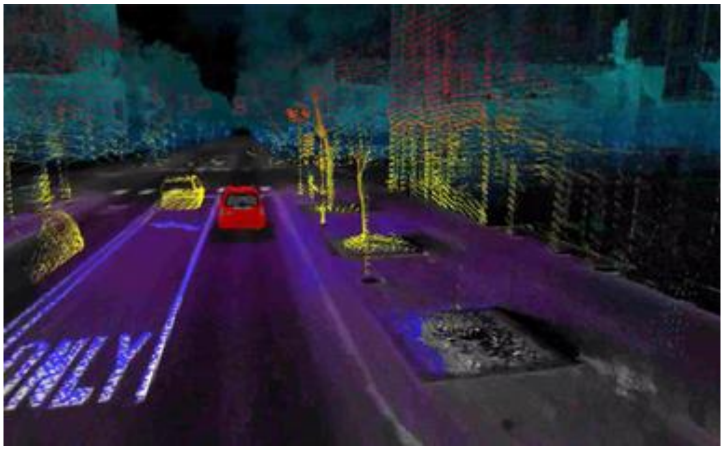


FIGURE 6: RENDERING OF DATA FROM A LIDAR SYSTEM.[32]

---

[31] http://velodynelidar.com/lidar/products/brochure/HDL-64E%20S2%20datasheet_2010_lowres.pdf
[32] http://www.autoguide.com/auto-news/2014/07/need-know-autonomous-vehicles.html

FIGURE 7: VELODYNE LIDAR READOUT AND DASHBOARD VIEW.[33]

The data from Lidar systems can be assembled into a three dimensional map of the surrounding environment. Currently this data and resultant 3-D map is used (or planned to be used) by algorithms within an autonomous vehicle to detect threats and navigate the vehicle through those threats (such as other vehicles, pedestrians or debris on the road). While the majority of the data from Lidar systems is very time dependent and used to identify immediate threats, some of that data, if extracted from the vehicle, could be used to help maintain State and local transportation agency base maps, roadway inventory, and imagery. Collecting and updating this data is an extremely time and resource-intensive responsibility for public agencies but essential to maintaining safe and reliable roadways and bridges. When funding is constrained, public agencies may rely on alternative sources such as Google Maps and Google Street View to understand field conditions.

However, there are multiple challenges regarding data from in-vehicle Lidar systems:

- The data needs to be extracted from the vehicle. Depending on the vehicle architecture, the raw Lidar data is likely sent directly to the Autonomous

---

[33] http://velodyne.com/blog/tag/autonomous-vehicle/

Driving module which controls the steering, throttle and break systems of the vehicle. There is no other output from the Lidar that would make the raw data accessible to outside systems.

- Lidar data is significant in size and content and could present a challenge for vehicle communication networks. One Lidar system is capable of recording over 1.3 million data points per second, and there could be multiple Lidar systems on a vehicle. This would impact not only the direct connection between the vehicle and a local site (DSRC radio location, Cell tower, etc.), but also the overall backhaul to a central processing center. For high traffic volume areas, this bandwidth would be multiplied over the number of vehicles in the area talking to that local site. At the same time, this data offload would compete with other uses of the available bandwidth, potentially including safety applications in a V2V environment and other telematics uses such as navigation or traffic and mode data. Potentially, this could be overcome by adding additional processing power so the vehicle can filter and send only pertinent information. That pertinent information will need to be defined and the algorithms to process/send that information would need to be standardized.
- Once the raw (or pre-processed) Lidar data is received by a DOT data management or processing center, it will need to be processed and converted into usable information. Typical vehicle bus data is simply a numeric value, such as speed, GPS location or brake pressure. Raw Lidar data is simply a value that depicts the distance to a point, relative to the vehicle. Back end data systems will need to correlate Lidar data to the location of the Lidar unit on the vehicle. Some of this data could be processed on the vehicle, which would require additional processing power on the vehicle as well as collaboration between the vehicle manufacturer and the data consumer.

The other major technology in supporting automated driving is the use of Machine Vision. Machine vision is the use of video cameras mounted on the vehicle to identify threats for input to the automated driving module of the vehicle that controls the steering, throttle and braking. Currently, much of the automated vehicle industry is focused more on radar systems, because they are less limited by inclement weather, dirt, or other blockages. Many current lane-keeping systems, however, use machine vision to detect lane markings. One example, developed by VisLab in

Italy, is shown in Figure 8. This is similar to the lane marking information from a Lidar solution shown in Figure 7.

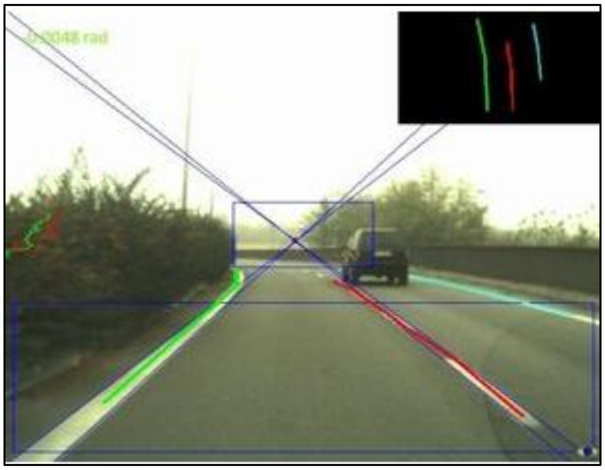


FIGURE 8: VISLAB LANE MARKING DETECTION.[34]

The data from machine vision systems has the same challenges as those from Lidar-based systems. The raw data is likely not readily available to the vehicle bus, the raw data (or video image) will require significant bandwidth for distribution outside of the vehicle to a processing center and the raw data will require significant post-processing by the user to acquire pertinent information.

4.7 AVAILABILITY OF DATA

Data availability is a challenging concept in the vehicle as different levels of availability need to be addressed. What data is available in the vehicle? Are there technical limitations that make the data unavailable? And what data will carmakers make available?

Different vehicles have different levels of data availability, especially beyond the core data of the widely used CAN BUS data. The data available in vehicles depends on a variety of factors, including vehicle manufacturer, model, year and options selected. While some data is standard, such as speed, heading and GPS location, other data desired by applications simply may not

---

[34] http://vislab.it/products/veld/

exist. Additionally, outside of the basic CAN data, data definitions are not standard for all data elements. For example, vehicle manufacturers have implemented different settings for intermittent windshield wipers.

Simply having the data on the vehicle does not mean it is readily available or accessible through the primary CAN bus. Not all data goes through the CAN bus, and each vehicle make and model has a different architecture for internal data systems. As a result, accessing data may be difficult or impossible without redesigning the internal systems. As an example, Miller and Valasek recently evaluated multiple vehicle architectures to determine which were most vulnerable to cyber-attacks. Figure 9 and Figure 10 show the architecture of a 2014 Honda Accord LX and a 2014 Chrysler 300 and show the different access points for different data elements as well as how the complexity of the architecture changes between the different vehicles.

Finally, even if data are available and readily accessible through the OBD-II port or other access point, if the automaker has not permitted access to the data, through a license or by documenting the data so systems integrated into the CAN bus can identify/interpret the data in a standard format/definition, then the data are still unavailable to other parties. Currently, many companies in the automotive aftermarket specialize in reverse engineering data on the CAN bus to enable different aftermarket systems to be integrated into the vehicle. If the automakers presently made all data readily available and accessible, these companies would not need to perform the step of reverse engineering these data. In other words, until the automakers are more open to publishing this information, significant efforts are needed to find, document, and pull the data from the vehicles.

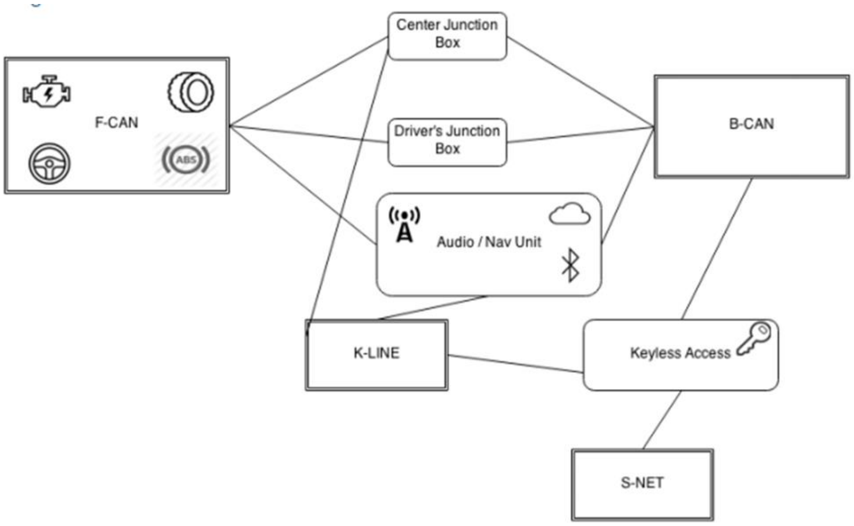


FIGURE 9: CAN ARCHITECTURE OF A 2014 HONDA ACCORD LX[35]

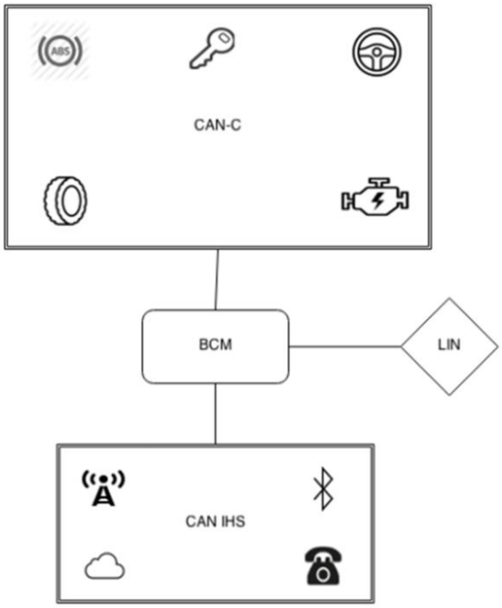


FIGURE 10: CAN ARCHITECTURE OF A 2014 CHRYSLER 300.

---

[35] Source: A Survey of Remote Automotive Attack Surfaces, Miller, Charlie and Valasek, Chris. http://illmatics.com/remote%20attack%20surfaces.pdf

5 CONCLUSIONS AND RECOMMENDATIONS

Connected and automated vehicles will fundamentally change our transportation systems. The societal benefits of these technologies include the reduction of crashes, energy consumption, pollution, and ultimately overall costs associated with the movement or people and goods. Connected and automated vehicle systems will also provide tremendous opportunities to improve the density, accuracy, completeness, and timeliness of data collection from a multisource data environment. The increased quantity and diversity of connected and automated vehicle-generated data may also drive new transportation systems analyses and provide new directions for system planning and management.

While there is still work to be done to develop applications that make use of connected vehicle data for operations, areas of known interest would include:

- Traffic operations and management
- Safety applications
- Performance measures and system analysis
- Asset management
- Travel behavior analysis
- Modeling and predictive traffic studies
- Traveler information systems
- New transportation services

Although the potential of connected and automated vehicle data to enable new applications is significant, it may also be constrained by the very nature of the data. For example, many connected and automated vehicle applications will require close collaboration between the public sector, industry, and individuals in the form of data sharing and information exchange agreements. The use of connected and automated data by the public sector will largely depend on the design of these partnerships. For example, data privacy concerns have the potential to limit direct applications of connected and automated vehicle data to origin-destination studies, even though such studies are highly valuable for transportation system planning.

The USDOT Connected Vehicle Program is designed largely to provide standardized and useful transportation data to public agencies. Many prototype connected vehicle applications are dependent on broad availability of vehicle data provided through DSRC/WAVE-connected vehicles. If a national DSRC network is not implemented, it is unclear what types of data will be available from advanced vehicles. Advanced automated vehicle systems are currently being deployed and designed without reference to DSRC, and are already seeing commercial success. While these advanced vehicle systems may generate information-rich datasets, such data may be proprietary and remain unavailable for use by transportation agencies. Even with full implementation of a national DSRC connected vehicle network, it is not certain that the information available from automated vehicle systems will be available for public agencies, given that the USDOT likely will be unable to require manufacturers to share data generated by proprietary systems.

There are many other data related challenges too. CAR has or will address these challenges and opportunities through separate studies. For example, ethical issues and decision-support relating to Intelligent Transportation Systems (ITS) data were discussed in recently completed *ITS Data Ethics in Public Sector*. Data management related issues will be addressed in *Management Procedures for Data Collected via Intelligent Transportation Systems*. Institutional and multi-modal related issues will be addressed in the *Use of ITS Technology for Management of Freight and Transit Assets*. Opportunities to incorporate crowd sourced data into planning and operation processes will be addressed in an upcoming research report.

Finally, although technological change is certain, it is difficult to predict the exact timeline of connected and automated vehicle applications in the real world. The assessment of their impacts on future transportation systems should be an evolving task. Similarly, the use of connected and automated vehicle generated data needs to be an ongoing effort in order to reflect new advancements in technologies, regulations, and policy developments.

REFERENCES

AASHTO. Connected Vehicle Field Infrastructure Footprint Analysis: Preparing to Implement a Connected Vehicle Future. 2013.

Bessler, S., and Thomas Paulin. *Literature Study on the State of the Art of Probe Data Systems in Europe*. FTW Telecommunications Research Center Vienna. 2013.

GAO (U.S. Government Accountability Office). Intelligent Transportation Systems: Vehicle-to-Vehicle Technologies Expected to Offer Safety Benefits, but a Variety of Deployment Challenges Exist. Nov 2013.

Goodall, N. J., Brian L. Smith, and Byungkyu "Brian" Park. Microscopic Estimation of Freeway Vehicle Positions from the Behavior of Connected Vehicles. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations. March 12, 2014.

International Transportation Forum. *Big Data in Transport: Applications, Implications, Limitations*. http://2014.internationaltransportforum.org/Big-Data. 2014.

ITS America. *Connected Vehicles – Next Generation of ITS.* http://www.itsa.org/industryforums/connectedvehicle. 2014.

MDOT and CAR (Michigan Department of Transportation and Center for Automotive Research. Connected Vehicle Industry Delphi Study. September 2012. Accessed May 2014 at: http://www.michigan.gov/documents/mdot/09-27-2012_Connected_Vehicle_Technology_-_Industry_Delphi_Study_401329_7.pdf

Morgan Stanley. Autonomous Cars: Self-Driving the New Auto Industry Paradigm. November 2013.

NHTSA. Preliminary Statement of Policy Concerning Automated Vehicles. 2013.

Nixon, Nikki. *Data in Motion vs. Data at Rest*. http://www.internap.com/2013/06/20/data-in-motion-vs-data-at-rest/. 2013.

OECD. Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data". 2013.

RAND Corporation (2013). Autonomous Vehicle Technology: A guide for Policymakers.

SAE International (2009). Surface Vehicle Standard: Dedicated Short Range Communications (DSRC) Message Set Dictionary.

SAE International (2010). DSRC Implementation Guide - A Guide to Users of SAE J2735 Message Sets over DSRC.

Slater, G., *AASHTO Draft Core Data Principals*. Presented at the Joint Midyear Meeting of the Transportation Research board and the American Association of State Highway Officials, Washington, DC, June, 2013. http://onlinepubs.trb.org/onlinepubs/conferences/2013/MPO/Slater.pdf

United Nations (May 2012). Big Data for Development: Challenges and Opportunities.

USDOT RITA. Real-Time Capture and Management Evaluation and Performance Measures – Evaluation Framework. Report No. FHWA-JPO-11-136. 2011.

USDOT. Vehicle Safety Communications – Applications (VSC-A) Final Report. Report Number DOT HS 811 492A. September 2011

USDOT FHWA. Vehicle Information Exchange Needs for Mobility Applications: Version 2.0, Revised Report. Report Number FHWA-JPO-12-021. August 1, 2012.

USDOT RITA. Connected Vehicle Data Capture and Management (DCM) and Dynamic Mobility Applications (DMA): Assessment of Relevant Standards and Gaps for Candidate Applications. Report No. FHWA-JPO-13-019. October 2012.

Wang, Jianqiang, Daiheng Ni, and Keqiang Li. RFID-Based Vehicle Positioning and Its Applications in Connected Vehicles. Sensors 2014, 14, 4225-4238.

APPENDIX A: LIST OF ABBREVIATIONS

3D, 3-D	Three Dimensional
AASHTO	American Association of State Highway Officials
AHS	Automated Highway System
AV	Automated Vehicle
BSM	Basic Safety Message
BSM II	Basic Safety Message Type-2
CAR	Center for Automotive Research
CAV	Connected and Automated Vehicle
CCTV	Closed Circuit Television
CVRIA	Connected Vehicle Reference Implementation Architecture
DE	Data Element
DF	Data Frame
DOT	Department of Transportation
DSRC	Dedicated Short-range Communication
EOBR	Electronic On-board Recorder
GPS	Global Positioning System
HMI	Human Machine Interface
IEEE	Institute of Electrical and Electronics Engineers
INS	Inertial Navigation System
ITS	Intelligent Transportation Systems
ITS JPO	ITS Joint Program Office
LTE	Long-term Evolution
MDOT	Michigan Department of Transportation
NHTSA	National Highway Traffic Safety Administration
OBE	On-board Equipment
PIAS	Personal Information Access Subsystem
RFI	Request for Information
RITA	Research and Innovative Technology Administration

RSE	Roadside Equipment
SAE	Society of Automotive Engineers
SPaT	Signal Phase and Timing
TMC	Traffic Management Centers
USDOT	United States Department of Transportation
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-grid, or vehicle-to- “everything” (V2V + V2I + vehicle to various devices and systems)
VMT	Vehicle Miles Travelled
VS	Vehicle Subsystem
WAAS	Wide Area Augmentation System
WAVE	Wireless Access in a Vehicular Environment (IEEE 802.11p + IEEE 1609)