Emerging Technologies in Transportation

PATH Team

November 8, 2021

Introduction

The purpose of this whitepaper is to give an overview of the technologies that are currently in the focus of attention from industry experts, investors, researchers and government. It is structured as follows. The rest of this section reviews transportation priorities on the National and California levels. Then, the whitepaper describes technologies of interest: internet of things; high-precision positioning; electrification; artificial intelligence; information systems - data management, privacy and security; connected and automated vehicles and drones. After that, the whitepaper discusses transportation application areas: traffic monitoring and performance measurement; congestion management; incident management; safety management; climate change and environment protection; public transportation; equity; parking and curb management; logistics; emergency response and evacuation.

This whitepaper **does not** cover:

- Pavement and construction;
- Maritime and aerospace technologies (except for some drone applications);
- Hyperloop and hypersonic travel.

National Priorities

As evidence of climate change continues to grow, the United States government has increased its focus on reducing environmental impact, especially with respect to the surface transportation system. The current administration has also emphasized the need to act urgently to ensure the country's infrastructure is resilient to severe-weather events caused by climate change. This has led to greater policy focus and increased federal spending in areas such as electrification, multi-modal transportation, active transportation (walking and biking), emerging technologies, and rebuilding infrastructure. Expanding the nation's workforce, maintaining safety, and social equity are other major priorities for the current administration.

As shown on their website,[1] the United States Department of Transportation (USDOT) is focusing on three major priorities:

- Safety: keeping traveling Americans safe.
- Innovation: engaging with emerging technologies.
- Infrastructure: maintaining the nation's infrastructure.

The first of these priorities has long been a goal of USDOT and its importance is becoming even clearer with new emerging threats such as COVID-19. The second

[1] USDOT. https://www.transportation.gov/priorities

priority, innovation, is also the focus of this white paper. USDOT views innovation as a catalyst to creating high-paying jobs and building a modern, sustainable transportation system. The final priority is seen as a prerequisite to ensuring our country's infrastructure is resilient to climate change. Infrastructure spending is also recognized as a stimulant to the nation's economy.

California State Priorities

California has long been a leader when it comes to protecting the environment, and the environment has become an even greater priority for the state as evidence of climate change continues to mount. On September 20, 2019, Governor Newsom issued Executive Order N-19-19, which proposes a bold climate agenda directing every agency of state government to redouble their efforts to reduce greenhouse gas (GHG) emissions and mitigate the impacts of climate change. Specifically, he directed the California State Transportation Agency (CalSTA) to:

- Align the state's climate goals with transportation spending where feasible;
- Reduce VMT by strategically directing discretionary transportation investment in support of housing production near available jobs and in accordance with the state's smart growth principles;
- Reduce congestion through innovative strategies designed to encourage people to shift from cars to other modes of transportation;
- Fund transportation options that contribute to the overall health of Californians and reduce GHG emissions, such as transit, walking, and biking;
- Mitigate increases in transportation costs for lower-income Californians.

On September 23, 2020, Governor Newsom went a step further by issuing Executive Order N-79-20. This order states that by 2035, all new cars and passenger trucks sold in California must be zero-emission vehicles. Because transportation currently accounts for more than 50 percent of California's GHG emissions, this order aggressively moves the state further away from its reliance on climate change–causing fossil fuels while creating jobs and spurring economic growth.

Caltrans' priorities as an organization closely align with the state's priorities of protecting the environment and ensuring transportation equity. These priorities are demonstrated in two major plans that were recently released: Caltrans Strategic Plan and California Transportation Plan (CTP) 2050.

Caltrans Strategic Plan

In early 2021, Caltrans released the Caltrans 2020-2024 Strategic Plan. The mission, vision, and goals that are in the plan closely align with California's transportation priorities of protecting the environment and ensuring equity for its citizens.

The Caltrans mission is to "provide a safe and reliable transportation network that serves all people and respects the environment."

Caltrans' vision is "a brighter future for all through a world-class transportation network."

Caltrans has established the following six goals as an organization:

- Safety first
- Cultivate excellence
- Enhance and connect the multimodal transportation network
- Strengthen stewardship and drive efficiency
- Lead climate action
- Advance equity and livability in all communities

Caltrans Transportation Plan 2050

CTP 2050, which was released in February 2021, is the state's long-range transportation plan that articulates strategic goals, policies, and recommendations to improve multimodal mobility and accessibility while reducing GHG emissions. The purpose of the plan is to present innovative, sustainable, and integrated multimodal mobility solutions to help guide the planning and implementation of a low-carbon transportation system. The plan aims to address many objectives, such as:

- Improving travel times and easing traffic congestion;
- Increasing safety and security on bridges, highways, and roads;
- Fostering healthy lifestyles through active transportation;
- Expanding economic opportunities through the movement of people, freight, services, and information;
- Creating a low-carbon transportation system that protects human and environmental health.

Technologies of Interest

Internet of Things

The Internet of Things (IoT) describes physical objects equipped with sensors, processing software, communication ability, and other technologies that connect and exchange data with other devices and systems over the Internet or other communications networks.

The IoT can assist in the integration of communications, control, and information processing and dissemination across various transportation systems. Application of the IoT extends to all components of transportation systems (i.e. vehicle, infrastructure, driver, other road user - bicyclist or pedestrian). Dynamic interaction between these components of a transport system enables inter- and intra-vehicular communication, vehicle-to-everything, intelligent traffic control, intelligent parking, electronic toll collection systems, logistics and fleet management, vehicle control, safety, and road assistance.

Communication Systems for IoT

The IoT makes use of a wide range of communication technologies - WiFi, cellular (LTE, 5G), satellite, radio frequency (Bluetooth, NFC)[2], mesh networks,[3] and low-power wide area networks (LPWANs). Most of these technologies have been around for a long time and are well known. **LPWAN** is a recent technology (the term appeared in 2013) that attracts a lot of attention - it will be the focus of our discussion.

LPWAN is a type of wireless telecommunication wide area network designed to allow long-range communications at a low bit rate among things (connected objects), such as sensors operated on a battery. The low power, low bit rate, and intended use distinguish this type of network from a wireless WAN that is designed to connect users or businesses, and carry more data, using more power. The LPWAN can accommodate packet sizes from 10 to 1,000 bytes, and the data rate ranges from 0.3 kbit/s to 200 kbit/s per channel.

It has the following characteristics:

---

[2] Radio frequency (RF) technologies, such as Bluetooth and near-field communications (NFC), enable only short range communication.

[3] A mesh network is a network in which devices are linked together, branching off other devices or nodes. These networks are set up to efficiently route data between devices and clients. They provide a consistent connection throughout a physical space, and are better suited for medium-distance IoT applications. An example of a mesh network is Zigbee - an IEEE 802.15.4-based specification for a suite of high-level communication protocols used to create personal area networks.

1. Long range: The operating range of LPWAN technology varies from a few kilometers (km) in urban areas to over 10 km in rural settings. It can also enable effective data communication in previously infeasible indoor and underground locations.
2. Low power: Optimized for power consumption, LPWAN transceivers can run on small, inexpensive batteries for up to 20 years.
3. Low cost: LPWAN's simplified, lightweight protocols reduce complexity in hardware design and lower device costs. Its long range combined with a star topology reduce expensive infrastructure requirements, and the use of license-free or licensed bands reduce network costs.

Unlike prior wireless technologies, LPWAN provides battery-efficient, ubiquitous wide-area connectivity, enabling more IoT applications that were previously prohibitive due to cost. However, a major tradeoff is the amount of data that can be transmitted. Yet, according to the 3rd Generation Partnership Project (3GPP), above 85% of all IoT devices use less than 3 MB of data per month.

LPWAN is not a single technology, but a group of various low-power, wide area network technologies that take many shapes and forms. LPWANs can use licensed or unlicensed frequencies and include proprietary or open standard options.

The proprietary, unlicensed **Sigfox** is one of the most widely deployed LPWANs today. Running over a public network in the 868 MHz or 902 MHz bands, the ultra-narrowband technology only allows a single operator per country. While it can deliver messages over distances of 30-50 km in rural areas, 3-10 km in urban settings and up to 1,000 km in line-of-site applications, its packet size is limited to 150 messages of 12 bytes per day. Downlink packets are smaller, limited to four messages of 8 bytes per day. Sending data back to endpoints can also be prone to interference.

Random phase multiple access, or RPMA, is a proprietary LPWAN from Ingenu Inc. Though it has a shorter range (up to 50 km line of sight and with 5-10 km no line of sight), it offers better bidirectional communication than Sigfox. However, because it runs in the 2.4 GHz spectrum, it is prone to interference from Wi-Fi, Bluetooth and physical structures. It also typically has higher power consumption than other LPWAN options.

The unlicensed **LoRa**, specified and backed by the LoRa Alliance, transmits in several sub-gigahertz frequencies, making it less prone to interference. A derivative of chirp spread spectrum (CSS) modulation, LoRa allows users to define packet size. While open source, the underlying transceiver chip used to implement LoRa is only available from Semtech Corporation, the company behind the technology. LoRaWAN is the media access control (MAC) layer protocol that manages communication between LPWAN devices and gateways.

Weightless SIG has developed three LPWAN standards: The unidirectional **Weightless-N**, bidirectional **Weightless-P** and **Weightless-W**, which is also bidirectional and runs off of unused TV spectrum. Weightless-N and Weightless-P are often more popular options due to Weightless-W's shorter battery life. Weightless-N and Weightless-P run in the sub-1 GHz unlicensed spectrum but also support licensed spectrum operation using 12.5 kHz narrowband technology.

Narrowband-IoT (NB-IoT) and **LTE-M** are both 3GPP standards that operate on the licensed spectrum. While they have similar performance to other standards, they operate on existing cellular infrastructure, allowing service providers to quickly add cellular IoT connectivity to their service portfolios.

NB-IoT, also known as CAT-NB1, operates on existing LTE and Global System for Mobile (GSM) infrastructure. It offers uplink and downlink rates of around 200 Kbps, using only 200 kHz of available bandwidth.

LTE-M, also known as CAT-M1, offers higher bandwidth than NB-IoT, and the highest bandwidth of any LPWAN technology.

Some vendors, including Orange and SK Telecom, are deploying both licensed and unlicensed technologies to capture both markets.

Other LPWAN technologies include:

- **DASH7**[4], a low latency, bi-directional firmware standard that operates over multiple LPWAN radio technologies including LoRa;
- **Symphony Link** from Link Labs Inc.[5];
- **ThingPark Wireless f**rom Actility[6];
- **Ultra Narrow Band** from various companies including Telensa, Nwave and Sigfox.

Different LPWAN technologies offer varying levels of security. Most include device or subscriber authentication, network authentication, identity protection, advanced encryption standard (AES), message confidentiality, and key provisioning.

The LPWAN landscape is constantly changing and far from mature. With many players in the market, it is unclear who the winner(s) will be, especially as the speed of market

[4] Dash7 Alliance. https://dash7-alliance.org
[5] Link Labs. https://www.link-labs.com/symphony
[6] Actility. https://www.actility.com/enterprise-iot-connectivity-solutions

expansion is also unknown. Long-term performance of each LPWAN variation is also uncertain, as many are still in their initial rollouts and real-world testing at scale is still underway.

Edge Computing

Edge computing is a distributed computing paradigm that brings computation and data storage closer to the sources of data. This is expected to improve response times and save bandwidth. A common misconception is that edge and IoT are synonymous. Simply stated, edge computing is a topology- and location-sensitive form of distributed computing, while IoT is a use case instantiation of edge computing. The term refers to an architecture rather than a specific technology. It is true, however, that edge is ideally suited for IoT applications. Actually, edge and IoT are more than just good partners: they are likely to increasingly depend on each other.

Note that edge does not replace the cloud, but extends or complements it. In fact, edge is a key enabler for unlocking the full power of data in the cloud. It does this by bringing data processing and other computing needs as close to the sensor or other device as possible, which reduces latency, among other potential benefits. Rather than sending data to be processed on external [cloud] servers or at data centers, costing precious seconds and additional resources, the computation takes place on the device or in the network itself. From there, the processed data can be delivered to its destination sooner. Edge computing reduces the cascade of potential bandwidth bottlenecks and processes the data that matters, keeping it close to the source. In fact, this is currently the most common architectural pattern: an edge computing environment placed near the sensors - or the “things” in IoT - that generate data.

IoT is about “things” that are distributed and connected with each other or a centralized infrastructure (like cloud) through a compute environment. These compute environments come in various forms: e.g., an embedded computer in a traffic signal cabinet, a webcam with processing capability, a connected car. They have the capacity to perform certain processing, analysis, and decision tasks, sending only end results to the cloud or data center.

Latest IoT Trends

The 2021 trends in IoT can be summarized as follows:

- The choice of connectivity options keeps growing. Satellite communication and LPWAN adoption gains speed.

- There is scrutiny on the security and privacy of IoT, as complex safety challenges crop up. IoT providers are pressed to highlight their security measures and handling private data.

- Edge computing in IoT becomes commonplace. As a result, cloud and edge providers explore partnerships. The systems that organizations can leverage to enable real-time analytics are already expanding beyond traditional data centers and deployment locations. Devices and computing platforms closer to end customers and/or co-located with real-world assets become an increasingly critical component of organizations' IT systems.

- Artificial Intelligence (AI), including Machine Learning (ML), becomes an integral part of IoT. Edge computing heavily relies on AI. Until recently, pre-processing of data via near-edge technologies or gateways had its share of challenges due to the increased complexity of data solutions, especially in cases with a high volume of events or limited connectivity. Now, AI/ML-optimized hardware, container-packaged analytics applications, frameworks such as TensorFlow Lite[7] and tinyML[8], and open standards such as the Open Neural Network Exchange (ONNX)[9] are encouraging machine learning interoperability and making on-device machine learning and data analytics at the edge a reality.

- IT and operational technology begin to converge. IoT-enabled devices (and other connected equipment) drive the adoption of edge solutions where infrastructure and applications are being placed within operations facilities. This approach is critical for real-time inference using AI models, which can detect changes in operating conditions and automate remediation.

High-Precision GNSS

Among the many sensors aboard a vehicle, the GNSS[10] receiver has a unique role: it provides absolute positioning in the form of latitude and longitude coordinates. A traditional GNSS receiver listens to satellites to measure distances to each satellite at a given time, and thus its location by simple geometry calculation. At a minimum, three satellites are needed for 2-D positioning; four satellites are required for identifying position and altitude. The distance is measured by the time it takes for a GNSS signal to travel from a satellite to the receiver. Traditional GNSS provides localization accuracy of 15-30 feet. Sufficient for destination navigation (navigation from point A to point B), this

[7] TensorFlow Lite. https://www.tensorflow.org/lite
[8] Tiny Machine Learning. https://www.tinyml.org
[9] Open Neural Network Exchange. https://onnx.ai
[10] GNSS stands for Global Navigation Satellite System. GPS is a U.S. implementation of GNSS.

accuracy is not enough to identify either a lane where the vehicle is driving or a distance between two vehicles on the road. To achieve a high accuracy in localization, additional corrections are needed.

The four approaches to high-precision GNSS are: Differential GNSS (D-GNSS), Satellite Based Augmentation System (SBAS), Real-Time Kinematics (RTK) and Precise Point Positioning (PPP). How do these work?

D-GNSS relies on (fixed) base stations whose locations are known exactly. Receiving signals from GNSS satellites, these bases deduce localization errors. Then, they transmit corrections to moving GNSS receivers by radio. With a distance of up to 600 miles between a base and a moving GPS, D-GNSS delivers localization accuracy of 1.6-16 feet. In the US, D-GNSS is known as NDGPS (National Differential GPS) and is maintained by the Coast Guard. Today, this technology is effectively obsolete.

SBAS builds upon additional, stationary satellites and a network of ground base stations, and provides wide-area or regional augmentation of GNSS through the use of broadcast messages from those stationary satellites. Using measurements from the ground stations, correction messages are created and sent to one or more satellites for broadcast to end users as a differential signal. In the US, the SBAS is known as the Wide Area Augmentation System (WAAS) and is operated by the Federal Aviation Administration (FAA). It is used to provide navigation in the national airspace.

The **RTK** concept is similar to that of D-GNSS. The improvement over D-GNSS is that it uses measurements of the phase of the signal's carrier wave in addition to the information content of the signal and relies on a single reference station or interpolated virtual station to provide real-time corrections. The system is commonly referred to as carrier-phase enhancement, or CP-GNSS. In practice, RTK systems use a single base station and a number of mobile receivers. The base station re-broadcasts the phase of the carrier that it observes, and the mobile units compare their own phase measurements with the one received from the base station and with each other. RTK provides accuracy enhancements up to about 20 km from the base station This allows the units to calculate their relative position to within millimeters, although their absolute position is accurate only to the same accuracy as the computed position of the base station. The typical nominal accuracy for these systems is one centimeter horizontally and two centimeters vertically. The biggest problem of contemporary RTK systems is that they are not robust: the phase calculation may be distorted under the bridges and near large buildings, and resetting may take up to a minute or more. The other problem of RTK is the relatively high cost of the GNSS-RTK receivers. Prices start at $250 per unit. To address this, Swift Navigation[11], a hardware and software developer, is building a global corrections network. To make it suitable for the automotive market, the company is aiming to make its corrections service affordable and scalable.

Precise point positioning (**PPP)** relies on two general sources of information: direct observables and ephemerides. Direct observables are data that the GPS receiver can measure on its own. One direct observable for PPP is a carrier phase, as in RTK. Another important direct observable is the differential delay between GNSS signals of different frequencies. By measuring the difference in the delays between signals of different frequencies, the receiver software (or later post-processing) can model and remove the delay at any frequency. This process is only approximate, and non-dispersive sources of delay remain, but it improves accuracy significantly. Ephemerides are precise measurements of the GNSS satellites' orbits, made by the geodetic community[12] with global networks of ground stations. Satellite navigation assumes that the satellites' positions at any given time are known, but in practice orbits are not perfectly predictable. The ephemerides that the satellites broadcast are earlier forecasts, up to a few hours old, and are less accurate (by up to a few meters) than carefully processed observations of where the satellites actually were. Therefore, if a GNSS receiver system stores raw observations, they can be processed later against a more accurate ephemeris than what was in the GNSS messages, yielding more accurate position estimates than what would be possible with standard real-time calculations. This post-processing technique has long been standard for GNSS applications that need high accuracy. More recently, projects such as APPS, the Automatic Precise Positioning Service of NASA JPL (National Aeronautics and Space Administration Jet Propulsion Laboratory), have begun publishing improved ephemerides over the internet with very low latency. PPP uses these streams to apply in near real-time the same kind of correction that used to be done in post-processing. Today, all automotive manufacturers are using a form of PPP corrections, which is a one-way broadcast, as opposed to the two-way communication between a base station and a moving receiver required for RTK. This means that a single correction stream can serve an entire continent, but its accuracy is 20 centimeters - worse than that of RTK. Trimble[13] and Hexagon[14] are the leading PPP providers.

Both RTK and PPP make the GNSS receiver a unique sensor providing a precise location in absolute (global) coordinates. The precise localization enables:

[11] Swift Navigation. https://www.swiftnav.com
[12] The International GNSS Service and other public and private organizations.
[13] Trimble. https://positioningservices.trimble.com/services/rtx
[14] Hexagon. https://hexagonpositioning.com/autonomous-x/automotive-positioning/serial-production/correction-services

- Identification of traffic speed on a lane level;
- Augmenting or even replacing camera- and lidar-based localization with GNSS, which is weather agnostic;
- Using GNSS sensor for distance and speed measurement in vehicle control;
- Tying the precise vehicle position with time, which allows more accurate processing and fusion of other onboard sensor measurements - e.g. cameras, LiDAR;
- Unmanned docking and package delivery by mobile robots (including drones).

Electrification

The electrification of transportation relies on availability of 1) efficient batteries; and 2) “on-the-go” (away from home or workplace) charging/fueling infrastructure.

Batteries

Presently, **lithium-ion (Li-ion)** batteries have taken over the world. Tesla has bet big on them and built a Gigafactory that is now knocking out their car batteries, as well as Powerwall and Powerpacks for homes and business. Many other manufacturers are working on their own supply chains of Li-ion batteries.

Li-ion is a type of rechargeable battery, in which lithium ions move from the negative electrode through an electrolyte to the positive electrode during discharge, and back when charging. Li-ion batteries use an intercalated lithium compound as the material at the positive electrode and typically graphite at the negative electrode. Li-ion batteries have a relatively high energy density, no memory effect (other than in lithium-iron-phosphate batteries)[15] and low self-discharge, and their cells can be manufactured to prioritize either energy or power density. They have their own limitations, however, which include:

- Slow recharging process;
- The danger of the electrolyte catching fire or exploding;
- Limited number of power cycles;
- Cost of utilization of used batteries.

Battery technology is cutting-edge, though, and there is a lot of research being conducted to improve the current state of the art. Li-ion technology providers mainly focus on:

[15] The memory effect is a deviation in the working voltage of the battery, caused by incomplete charging or discharging, that can lead to only part of the stored energy being available and an inability to determine the charge level of the battery reliably.

- Modularizing the battery to speed up the charging process;
- Increasing battery capacity for the same size and weight;
- Modifying the electrolyte to make it more durable - to increase the admissible number of power cycles;
- Modifying the electrolyte to make it less hazardous - less flammable and explosive.

At the same time, there are numerous competing power technologies that could replace or complement Li-ion. These are listed below.

Lithium–sulfur (Li-S) batteries may succeed lithium-ion cells because of their higher energy density and reduced cost due to the use of sulfur instead of cobalt, which is commonly used in lithium-ion batteries. Some Li-S batteries offer specific energies of the order of 550 Wh/kg, while most lithium-ion batteries are in the range of 150-260 Wh/kg. Li-S batteries with up to 1,500 charge and discharge cycles were demonstrated in 2017, but cycle life tests at commercial scale and with lean electrolyte are still needed. As of today, none are commercially available.

Conventional battery electrode materials (and prospective ones) are significantly improved when enhanced with **graphene**. Graphene is an allotrope of carbon consisting of a single layer of atoms arranged in a two-dimensional honeycomb lattice nanostructure. Its name is derived from “graphite” and the suffix -ene, reflecting the fact that the graphite allotrope of carbon contains numerous double bonds. A graphene battery can be light, durable and suitable for high capacity energy storage, as well as shorten charging times. It will extend the battery’s life, which is negatively linked to the amount of carbon that is coated on the material or added to electrodes to achieve conductivity, and graphene adds conductivity without requiring the amounts of carbon that are used in conventional batteries. Graphene can improve such battery attributes as energy density and form in various ways. Li-ion batteries (and other types of rechargeable batteries) can be enhanced by introducing graphene to the battery’s anode (negative electrode) and capitalizing on the material’s conductivity and large surface area traits to achieve morphological optimization and performance. For the same size and weight, a graphene battery has 6 times as much capacity as the present Li-ion, while its recharging time is 10 times less. The main challenge of graphene technology today is the cost of production.

Solid-state batteries use solid electrodes and a solid electrolyte, instead of the liquid or polymer gel electrolytes found in lithium-ion or lithium polymer batteries. They avoid the use of dangerous or toxic materials found in commercial batteries, such as organic electrolytes, but deliver higher energy densities (2.5x) by enabling lithium metal anodes. These batteries recharge faster than Li-ion, provide higher voltage and more power cycles. Moreover, they are non-flammable. Their disadvantages are their high production cost and poor performance in low temperatures.

New generation solar batteries are called **perovskite solar cells (PSCs)**. PSC is a type of solar cell which includes a perovskite-structured[16] compound, most commonly a hybrid organic-inorganic lead or tin halide-based material, as the light-harvesting active layer. Perovskite materials, such as methylammonium lead halides and all-inorganic caesium lead halides, are cheap to produce and simple to manufacture. The cost of PSCs is estimated at 25 cents per square foot. This technology is an enabler for the whole surface of the car to be a solar panel in the years to come, which could make vehicles self-sufficient.

Leading automakers are heavily investing in **hydrogen fuel cells**. Fuel cell electric vehicles (FCEVs) use a propulsion system similar to that of electric vehicles, where energy stored as hydrogen is converted to electricity by the fuel cell. Unlike conventional internal combustion engine vehicles, these vehicles produce no harmful tailpipe emissions - they only emit water vapor and warm air. The most common type of fuel cell for vehicle applications is the polymer electrolyte membrane (PEM) fuel cell. In a PEM fuel cell, an electrolyte membrane is sandwiched between a positive electrode (cathode) and a negative electrode (anode). Hydrogen is introduced to the anode, and oxygen (from air) is introduced to the cathode. The hydrogen molecules break apart into protons and electrons due to an electrochemical reaction in the fuel cell catalyst. Protons then travel through the membrane to the cathode. The electrons are forced to travel through an external circuit to perform work (providing power to the electric car) then recombine with the protons on the cathode side where the protons, electrons, and oxygen molecules combine to form water. FCEVs are fueled with pure hydrogen gas stored in a tank on the vehicle. Similar to conventional internal combustion engine vehicles, they can fuel in less than four minutes and have a driving range over 300 miles. FCEVs require hydrogen fueling infrastructure that presently is still in its infancy in the US.

On-the-Go Charging/ Fueling Infrastructure

Presently, electric vehicles (EVs) account for 1.5% of the 28 million vehicles registered in California. For the US, the share of EVs is even less.

Right now, there are about 60,000 public EV chargers in California. By 2025, the state expects to have 250,000, including 10,000 so-called fast chargers that can refuel a battery electric vehicle to 80% full in 25 minutes. That is enough to support 1.5 million

[16] A perovskite is any material with a crystal structure similar to the mineral called perovskite, which consists of calcium titanium oxide.

zero emissions vehicles. Meanwhile, Governor Newsom's executive order requires that all passenger vehicle sales be zero emission by 2035. To get to the number of EVs the governor's executive order calls for, one has to increase the existing infrastructure tenfold. The goal is to have roughly 15 million zero emissions vehicles on the road by 2035.

EV chargers are defined by the amount of energy delivered to the vehicle's battery per unit of time. There are four levels, with Level 4 being the fastest. As the table below shows, different levels of chargers have very different power ratings and charge times for typical EVs.

Charger Level	Power Rating (KW)	Typical Installation	Charge Time (80%)
1	1	Standard electrical outlet in a domestic garage	20 hours
2	5	● Specialized domestic charging apparatus (often sold as optional extra with vehicle or supplied by specialist third parties) ● Workplace parking lot installations ● Public access charge points at retail stores, parking lots, etc.	4 hours
3	80	Specialized fast chargers designed for users on-the-go	40 minutes
4	120	● Ultra-fast chargers for users on the go ● To-date installed by vehicle OEMs as proprietary support for their customers or owners	25 minutes

The capital and operating costs of higher-capacity chargers (excluding the electrical power delivered) is higher than those of lower-capacity chargers. Level 1 chargers - like a residential 110-volt outlet - are nearly free to operate. Level 2, 3 and 4 chargers, however, are more expensive and complex. The capital expenditure (CapEx) depends on the number of chargers per charging station. The larger is the number of chargers, the smaller is the CapEx per charger. According to a PricewaterhouseCoopers report[17]

[17] PricewaterhouseCoopers. Electric vehicles and the charging infrastructure: a new mindset? https://www.pwc.com/us/en/industries/industrial-products/library/electric-vehicles-charging-infrastructure.html

on EVs and the charging infrastructure, for six charges per station the CapEx is estimated at:

Level 2 - $6,000 per charger, $1,200 per KW;

Level 3 - $49,000 per charger, $613 per KW;

Level 4 - $96,000 per charger, $800 per KW.

Operating costs for charging stations are quite low, because they are typically unattended, self-service and require less maintenance than gas pumps. However, charging locations connected to a local electricity grid are subject to “capacity charges” used to offset the costs of utility upgrades needed to accommodate them. Capacity charges are linked to the highest expected load a given charging station may require. In principle, then, a Level 4 station could have twice the peak load of a Level 2 station - and a higher capacity charge. This capacity charge is fixed and recurring (typically monthly), regardless of whether or not the peak load is reached in any given month.

Higher level chargers have higher costs than lower level ones, unless the utilization is very high - see **Figure 1**.

The four implications of this cost structure are:

1. There are relatively modest returns-to-scale at the location level.
2. Once an EV charging station is built, all costs are essentially fixed, so utilization is key to achieving efficiency.
3. Although fast chargers are significantly more costly to build, their greater capacity

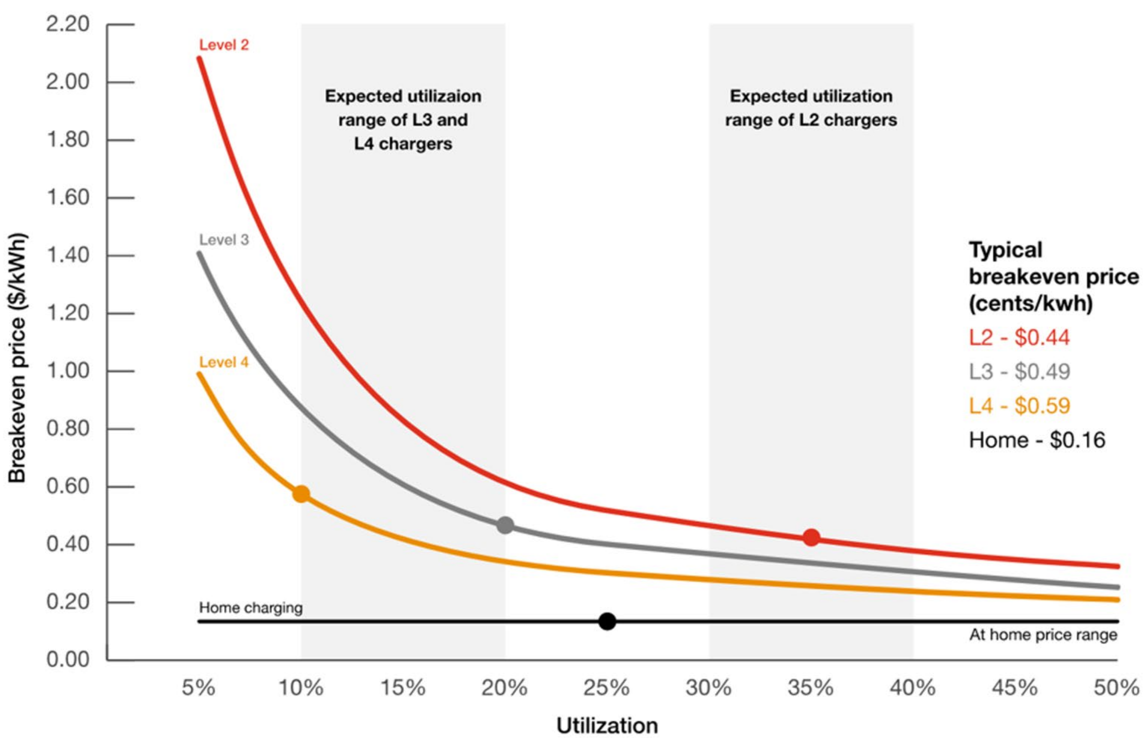

Figure 1. Charging economics: breakeven price by charger type. 4 charger types for various utilization levels. All-in breakeven price is a selling price required to earn 10% return on capital invested with a wholesale power cost of $0.16 per kwh. Source: PwC Report.

means that their overall cost-to-serve (per kwh delivered) can compete with lower-capacity chargers even at lower utilizations.

4. Third-party charging stations are significantly more costly than much slower charging at home or at work.

The main stakeholders of EV infrastructure are:

1. **Standalone EV charging providers**. These companies are playing the long game. They're building charging capacity in advance of demand, absorbing losses and counting on the eventual rise in EV adoption. These players often seek to play in numerous EV-charging use cases such as installing home chargers (not unlike installing a home appliance) and at-work charging networks, setting up charging points at commercial locations (such as parking lots at hotels or retail centers) and building roadside, fast-charging stations.

2. **Vehicle OEMs (including consortia)**. Vehicle OEMs (Original Equipment Manufacturer) are investing in charging networks across the U.S. to drive demand for their vehicles. Some may offer charging for free (at least for some limited period) as an incentive to buy the car. These players typically negotiate deals with other businesses - a gas station or restaurant, for example - that pay fees in exchange for the prospect that the charger will draw customers. It seems to work: One convenience-store retailer that installed an OEM's chargers reported an 80% conversion of charges to store visits, with those store visits resulting in roughly twice the value of purchases compared to regular fuel customers.
3. **State and local governments**. EV adoption feeds into numerous state and municipal government environmental and sustainability goals. Some are either investing directly or subsidizing EV charging infrastructure development within their jurisdictions. Typically, though, these charging networks are Level 2 roadside or parking lot solutions which, as we've discussed, are a challenging use case. The upside, of course, is that these slow chargers are less costly to install.

The availability of stations providing reasonably priced hydrogen in places where vehicles will be deployed remains a key challenge to the adoption of this technology. To address this challenge, the U.S. Department of Energy (DOE) launched H2USA[18], a public-private collaboration with federal agencies, automakers, hydrogen providers, fuel cell developers, national laboratories, and additional stakeholders. H2USA is focused on advancing hydrogen infrastructure to support more transportation energy options for U.S. consumers.

As of mid-2021, there are 48 open retail hydrogen stations in the United States. Additionally, there were at least 60 stations in various stages of planning or construction. Most of the existing and planned stations are in California, with one in

[18] Hydrogen and Fuel Cell Technologies Office. https://www.energy.gov/eere/fuelcells/h2usa

Hawaii and 14 planned for the Northeastern states. As the market expands, hydrogen fueling stations will be matched with vehicle rollout as both grow together. Customers are expected to have similar experiences at hydrogen fueling stations as at gasoline stations, with most hydrogen dispensers being added at existing gasoline stations.

Artificial Intelligence

The term Artificial Intelligence (AI) refers to software and systems that mimic the cognitive functions of humans. Machine Learning (ML) is a subset of AI and refers to algorithms that can learn to iteratively improve at a certain task. At a theoretical level, machine learning can be considered as the mathematical modeling of how "teaching and learning" works. At the practical level, it is the study of how to build applications that exhibit this iterative improvement. There are many ways to formulate this idea, but largely there are four major recognized categories: supervised learning[19], unsupervised learning[20], semi-supervised learning[21], and reinforcement learning[22] (RL).

Supervised learning requires a labelled set of data, and is commonly used to train systems to solve classification problems. Unsupervised learning does not require labelled data, and is commonly used to find anomalies or to find features for clustering. Semi-supervised learning involves training with both labelled and unlabelled data. One application is in facial recognition, where data can first be clustered and then labeled pictures used to identify particular faces. Reinforcement learning involves creating a feedback loop between an agent and an environment. The environment is able to reward or penalize actions of an agent. The agent attempts to optimize the long-term reward. This approach has resulted in programs that have mastered games such as chess, Go, and poker.

Deep learning is synonymous with "deep neural networks" involving networks with multiple layers between the input and output. The Convolutional Neural Network (CNN) is inspired by the organization of the human visual cortex. It consists of layers to perform operations such as convolution, and pooling to process features of the input. A Recurrent Neural Network (RNN) is structured so that connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior. It is used for processing sequential data like sound, time series data, and written natural language. Other examples of neural networks include the Multilayer Perceptron Neural Network (MLPNN), Long-short Term Memory (LSTM), and Generative Adversarial Network (GAN). These algorithms can be used alone or

[19] Supervised learning: https://en.wikipedia.org/wiki/Supervised_learning
[20] Unsupervised learning: https://en.wikipedia.org/wiki/Unsupervised_learning
[21] Semi-supervised learning: https://en.wikipedia.org/wiki/Semi-supervised_learning
[22] Reinforcement learning: https://en.wikipedia.org/wiki/Reinforcement_learning

combined with other algorithms according to the properties of the target application problem.

There are some well-known algorithms that are often used as a part of AI solutions. For example the Support Vector Machine (SVM) finds an optimal separating hyperplane in the feature space of the data so as to perform classification. Ensemble methods combine multiple weak classifiers to build a strong classifier. One common example of an ensemble is a random forest made up of many decision trees.

Applications of Artificial Intelligence

AI finds myriad applications in areas such as computer vision, natural language processing, robotic operation and automated driving.

Tasks in computer vision are commonly defined as follows:

- Image classification involves predicting the class of one object in an image.
- Object localization refers to identifying the location of one or more objects in an image and drawing a bounding box around them.
- Object detection combines these two tasks for one or more objects in an image - see **Figure 2**.[23]

[23] K. Potdar, C. Pai, S. Akolkar. A Convolutional Neural Network based Live Object Recognition System as Blind Aid. November 2018. https://www.researchgate.net/publication/329217107_A_Convolutional_Neural_Network_based_Live_Object_Recognition_System_as_Blind_Aid

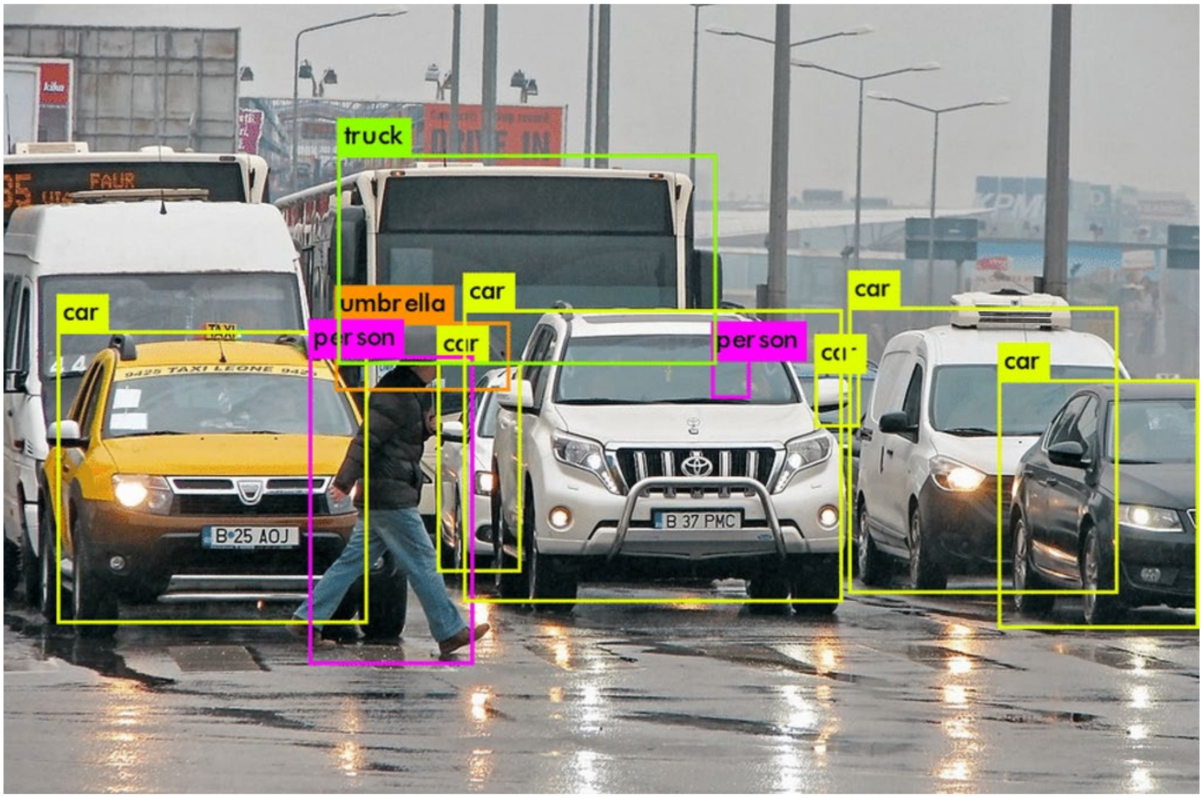

Figure 2. Object detection in a dense scene. Source: DOI:10.13140/RG.2.2.34494.54085.

Self-driving technology enables an AV to operate by perceiving the environment and executing maneuvers. **Figure 3** below depicts a machine learning based self-driving architecture. The architecture can be implemented either as a sequential perception-planning-action pipeline (a); or as an End2End system (b). In the sequential pipeline case, the components can be designed either using AI and deep learning methodologies or based on classical non-learning approaches. End2End learning systems are mainly based on deep learning methods. A safety monitor is usually designed to ensure the safety of either module.[24]

[24] S. Grigorescu, B. Trasnea, T. Cocias, G. Macesanu. A Survey of Deep Learning Techniques for Autonomous Driving. March 2020. https://arxiv.org/abs/1910.07738

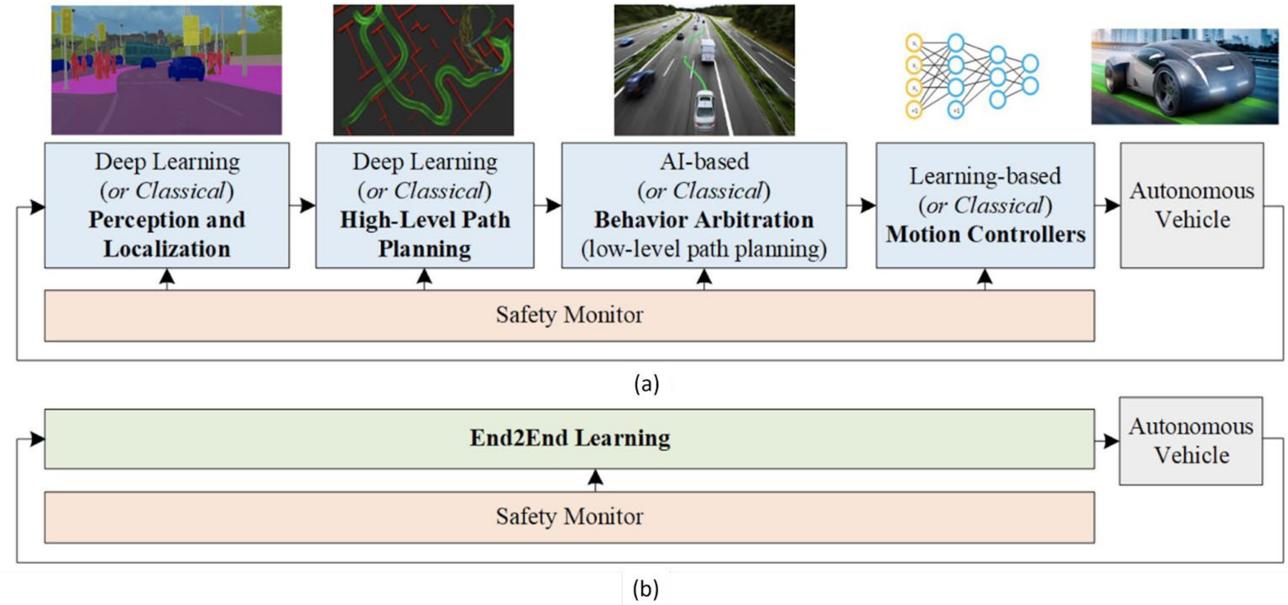

Figure 3. Self-driving architecture approaches: (a) sequential pipeline; (b) End2End.
Source: https://arxiv.org/abs/1910.07738

High-Definition (HD) Map Construction and Update

HD maps provide layers of data to overcome sensing limitations of AVs. These HD maps need to represent the world at an unprecedented centimeter resolution[25], and be updated frequently as road conditions in the real world constantly change. Aspects of this can be accomplished with crowd-sourced data collection via vehicle fleets. In several proof-of-concept projects, the Toyota Research Institute-Advanced Development (TRI-AD) built map information for self-driving via map data derived from the cameras of ordinary vehicles as well as satellite imagery, without the use of conventional means of collecting data such as survey vehicles.[26]

Mobileye's map creation process is called Road Experience Management – REM™. Mobileye gathers data for its maps through consumer vehicles equipped with Mobileye's EyeQ4 driving assistance system. It turns mapmaking into a byproduct of human driving without the need to add any expensive hardware or additional cars on the road.[27]

[25] Medium.com. Rethinking Maps for Self-Driving, Oct. 2018. https://medium.com/lyftlevel5/https-medium-com-lyftlevel5-rethinking-maps-for-self-driving-a147c24758d6
[26] Toyota.com. TRI-AD enables creation of HD maps, March 2020. https://global.toyota/en/newsroom/corporate/31898884.html
[27] NewsroomIntel.com. The Why and How of Making HD Maps for Automated Vehicles, Nov. 2019. https://newsroom.intel.com/articles/why-how-making-hd-maps-automated-vehicles/#gs.dlx828

EyeQ4 camera systems identify and process lane markings, curbs, landmarks, traffic signs, telephone poles and other infrastructure. Data is translated into coordinates in compact, compressible data – about 10 kilobytes per kilometer captured. In a year of typical North American driving, this equates to approximately 200 megabytes per vehicle.

Driver Monitoring and Driver-Vehicle Interaction

AI/ML has enhanced the sensing capabilities inside a car to monitor drivers and potentially to detect distraction and drowsiness. Additionally, the AI/ML technologies have been introduced into late model cars to enable new modes of interaction with drivers, such as voice-activated and gesture-based interface control.

A paper on driver behavior recognition[28] proposed a dedicated Interwoven Deep Convolutional Neural Network (InterCNN) architecture to tackle the accurate classification of driver behaviors in real-time. Basic driver monitoring features include head tracking, gaze tracking, eye state analysis – blink rate, blink duration, eye open/close all of which can be used to implement driver safety applications like driver distraction and driver drowsiness detection.

For automobiles, it is important to ensure that human-machine interface can be understood and effectively used by occupants. Voice assistants like Alexa, Siri, and Google Assistant may play a larger role in the driver experience in coming years. Gesture control makes it easier to use built-in systems and reduces the probability of causing a distraction-related accident by enabling drivers to better keep their eyes on the road.[29]

Traffic Data Processing

Advancements in image processing are giving roadside sensors an increasingly important role in traffic monitoring and surveillance.[30] Supervised machine learning algorithms are widely applied in this domain.

---

[28] Zhang, C. et al. Driver Behavior Recognition via Interwoven Deep Convolutional Neural Nets with Multi-stream Inputs, Nov. 2018. https://arxiv.org/pdf/1811.09128.pdf

[29] Automotive Gesture Recognition—The Next Level in Road Safety, Mar. 6, 2019, https://www.electronicdesign.com/markets/automotive/article/21807672/automotive-gesture-recognitionthe-next-level-in-road-safety

[30]Datondji, S.R. et al. A Survey of Vision-Based Traffic Monitoring of Road Intersections, April 2016. https://ieeexplore.ieee.org/document/7458203

Object Detection

Machine learning-based processing capability supports traffic management and helps us gain insight on traffic conditions and infrastructure status - facilitating alerts to drivers and vehicular control assistance. This is a very fast-moving field.

New approaches are being applied to improve the automated detection of vehicles and pedestrians. An approach using Generative adversarial networks (GANs) learned a multi-scale metric to distinguish multiple objects at different distances in one image.[31]

Researchers at Russia's South Ural State University (SUSU) developed a system for monitoring traffic flow.[32] The Artificial Intelligence Monitoring System (AIMS) classified 10 categories of vehicles, measured speed, the current load level of each direction of the intersection, and determined the further direction of vehicles.

Additional studies advance CNN-based object detection solutions to handle various nighttime environments, such as urban nighttime conditions with insufficient illumination.[33]

Vehicle Classification

Active research continues in the area of vehicle classification. In one example study using a CNN-based approach,[34] three types of vehicles like bus, car, and bike were considered for classification and detection. The approach used the entire image as input and created a bounding box with probability estimates of the feature classes as output. The results of the experiment showed that the projected system could considerably improve the accuracy of the detection.

Incident and Event Detection in Social Media

Incident detection is vital for traffic control and management and is an important prerequisite for quick restoration of smooth traffic flow in urban networks. Some agencies have started implementing processes to change signal timings in real time based on observations of incident and traffic conditions at congested locations.

---

[31] Dinakaran,R. Deep Learning based Pedestrian Detection at Distance in Smart Cities, May 2019. https://www.researchgate.net/publication/329387607_Deep_Learning_based_Pedestrian_Detection_at_Distance_in_Smart_Cities

[32] Traffic Technology Today. AI-based traffic monitoring system developed by researchers, Aug. 2019. https://www.traffictechnologytoday.com/news/congestion-reduction/ai-based-traffic-monitoring-system-developed-by-researchers.html

[33]Leung, H.K. et al. A Deep-Learning-Based Vehicle Detection Approach for Insufficient and Nighttime Illumination Conditions, Nov. 2019. https://www.researchgate.net/publication/337120612_A_Deep-Learning-Based_Vehicle_Detection_Approach_for_Insufficient_and_Nighttime_Illumination_Conditions

[34] Vijayaraghavan, V., Laavanya, M. Vehicle Classification and Detection using Deep Learning, Dec. 2019. https://www.ijeat.org/wp-content/uploads/papers/v9i1s5/A10061291S52019.pdf

One study[35] employed deep learning in detecting the traffic accident from 1-year social media data over 3 million tweet contents in two metropolitan areas: Northern Virginia and New York City. Two deep learning methods: Deep Belief Network (DBN) and Long Short-Term Memory (LSTM) were investigated. The accident-related tweets were compared with both the traffic accident log on freeways and traffic data on local roads from 15,000 loop detectors. It was found that nearly 66% of the accident-related tweets could be located by the accident log and more than 80% of them could be tied to nearby abnormal traffic data.

Another study[36] developed a machine learning algorithm to separate the alerts for potential high-impact incidents from the alerts related to congestion and other non-critical stops or slowdowns. The results indicated a massive reduction in non-critical alerts without a significant reduction in detection rate or time-to-detect the incidents.

Signal Control

Available real-time traffic data together with recent advancements in computational methods has the potential to make significant improvements in traffic signal control. In particular, the successful applications of reinforcement learning in solving complex problems suggests a promising opportunity to endow traffic signals with a higher degree of intelligence.

In a recent work by our team at California PATH, we adopted a neural networked RL approach that took into consideration the spatial-temporal characteristics of urban traffic flows. A Q-network was proposed to extract discriminative information from the state space to derive an optimal signal control policy. The results demonstrated the superior performance of our RL approach under different traffic patterns against three benchmark methods in terms of queue length and wait time.

Traffic System Management

One study[37] explored the data mining and machine learning technologies adopted in research and industry. The study comprehensively reviews around 165 studies, criticizes, and categorizes all of them into a chronological category. The study finds that there is no standard traffic management approach that the community of traffic

---

[35] Zhang, Z. et al. A deep learning approach for detecting traffic accidents from social media data, Jan. 2018.
https://www.sciencedirect.com/science/article/pii/S0968090X1730356X

[36] Kim, H.G. et al. Using Machine Learning to Identify High Impact Incidents in Automatic Incident Detection (AID) System Generated Alarms, Jan. 2020. https://smartech.gatech.edu/handle/1853/62451

[37] Alserehin, N. et al. Intelligent Transportation and Control Systems Using Data Mining and Machine Learning Techniques: A Comprehensive Study, April 2019.
https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8681028

management agrees on. Additionally, the study draws general attention to a new traffic management proposition approach.

Next Steps for AI Applications in Transportation

As evidenced by the numerous potential solutions made possible by AI and ML, the landscape of transportation has been radically transformed. Various functions in transportation can be enhanced and facilitated by AI to enable smart traffic management solutions.

A USDOT report[38] introduces the potential applications in transportation management system (TMS) and transportation management center (TMC) operations, showcases successful AI applications for TMSs, and provides a list of important issues to consider in developing AI applications. In US, a variety of AI technologies have been deployed for transportation systems management and operations (TSMO) applications:

- Several State and local DOTs (Nevada, Florida, and Iowa)[39] began deploying neural network technologies for incident detection using video image analysis and traffic prediction.
- Fuzzy logic has been used by Washington State DOT for more than 20 years[40] and California DOT (Caltrans) has begun deployment of fuzzy logic metering in a pilot corridor.[41]
- Delaware DOT has piloted several AI applications for traffic congestion and incident prediction.
- The Metropolitan Transportation Commission of the Bay Area and several other agencies have light integration of 511 with Alexa. Several arterial management agencies are piloting use of Google Assistant.
- More than 20 State DOTs have active UAS programs that may be enhanced with AI in the near to medium term.
- Several DOTs are piloting use of automated vehicles for crash abatement.

Going forward, AI technologies are likely to come closer and closer to "plug and play," but currently there is still a reasonably large barrier between the dreams of AI-enabled

[38]FHWA. Raising Awareness of Artificial Intelligence for Transportation Systems Management and Operations, Dec. 2019.
https://ops.fhwa.dot.gov/publications/fhwahop19052/execsumm.htm

[39] FHWA. Raising Awareness of Artificial Intelligence for Transportation Systems Management and Operations.
https://ops.fhwa.dot.gov/publications/fhwahop19052/chap4.htm

[40] USDOT. Evaluation of a Fuzzy Logic Ramp Metering Algorithm: A Comparative Study Among Three Ramp Metering Algorithms, Feb. 2000. https://www.wsdot.wa.gov/research/reports/fullreports/481.2.pdf

[41] FHWA. https://ops.fhwa.dot.gov/publications/fhwahop19052/chap4.htm

TSMO applications and the need for significant expertise and investment to make those dreams a reality. As fast as the pace of development of AI tools and technologies is progressing, AI applications should find their way from research experiments and pilot demonstrations to fully scalable applications in the near term.

Suggested Approaches for AI Projects

The Federal Highway Administration (FHWA) report further comments on the need for a holistic program plan that should be developed. The agencies in the early planning stages may take the following steps:

1. Convene an interdepartmental workshop to educate stakeholders, partners, and potential partners on AI and brainstorm potential applications and synergies.
2. Discuss priorities, opportunities, and barriers to AI applications in each of the TSMO areas.
3. Determine a short list of high-priority applications and a longer list of secondary-priority functions that address regional issues, challenges, and goals. While many goals are generic, tailoring the AI strategy to regional hot-button issues is typically helpful in gaining broader buy-in from decision makers and associated departments of State DOTs and local infrastructure owner-operators.
4. Review the list of general and detailed questions in the FHWA report and consider the responses of your organization to each.
5. Develop a project plan to implement the actions.

Research Planning for AI Deployment

The following factors should be considered for the selection of suitable projects.

- AI/ML use cases
    - AI developments have migrated from scientific explorations to engineering exercises.
    - The core capability of AI/ML resides mostly in software, which can be trained and tested on most computing platforms, and do not necessarily require high computing costs except for in cases with large volumes of video and image data.
    - Current strength of AI/ML is in pattern recognition, but increasingly, decision-making systems are being engineered based on AI/ML.
    - Methodology is mostly transferrable, from one problem domain to another, if the formulation of data structure is similar.

- Resources and data support

- Open-sourced resources, including algorithms and datasets, should be leveraged.
- For training and testing of machine learning algorithms, real-world data sets are preferred.
- Simulated and synthetic data can also be utilized to augment the existing data. Techniques for generating valuable data augmentation are largely existent.

- Project selection
  - The ideal project should be scalable, with potential to grow after pilot validation.
  - Selection of candidate projects can be prioritized by the projected benefits in operational efficiency versus the project costs.
  - UC Berkeley has a highly ranked AI group and several programs (BAIR[42], BDD[43], BAIR Commons[44]) that conduct collaborative research with industries.
  - The knowledge base and implementation experience of the AI community at Berkeley can help incubate project ideas.
  - Caltrans can leverage the AI expertise at Berkeley via a research program such as an AI Center for Transportation at California PATH.

Transportation Information Systems Infrastructure, Data Infrastructure and Management, and the Opportunities for a True Intelligent Transportation System

California's Transportation Information Technology Systems and Data Infrastructure

California's transportation technology infrastructure is a large, complex, and often fragmented system. It is composed of local, regional, and state systems with little or no connectivity. Caltrans' systems infrastructure is often defined at a district level meeting district needs with limited state level standardization and limited connectivity with local and regional partners. Caltrans has several efforts underway to standardize some of its infrastructure across its districts, for example its Correspondence and Task Management System (CATMS) program. Such programs provide not only the opportunity to standardize the systems that Caltrans utilizes and lower overall lifecycle

[42] https://bair.berkeley.edu/
[43] https://deepdrive.berkeley.edu/
[44] https://bair.berkeley.edu/blog/2019/03/24/bair-commons/

costs of its systems infrastructure, but also opportunities to standardize and modernize its data collection, exchange, and analysis capabilities. It provides the opportunity to enable real-time data exchange with regional and local partners, creating the California transportation data infrastructure necessary to enable a modern, intelligent transportation system capable of meeting California's goals and priorities.

Each of the national and state priorities mentioned in the beginning of this report as well as the goals of the Caltrans strategic plan and CTP 2050 depend upon a modern transportation systems infrastructure and the ability to collect, exchange, and utilize both existing and new sources of transportation data. The technologies and transportation applications identified in this report require significant advances in Caltrans' systems infrastructure, data acquisition, data management, and data governance capabilities to achieve their objectives and measure their effectiveness. For a variety of reasons including legacy transportation infrastructure, organizational issues, workforce limitations, and cost, the transportation industry and Caltrans have been unable to take full advantage of the advances in information technology. Significant change, effort, and funding will be required to address these challenges. The modernization of transportation, the new solutions required to address California's transportation challenges, the move toward truly intelligent transportation systems and automation, and the limited ability to add more pavement require modernization of California's systems and data infrastructure.

Importance of Data in the Present and Future of Transportation

Today's transportation problems do not respect jurisdictional boundaries. Rather, they require solutions that cross jurisdictional boundaries, incorporating and integrating state, regional, and local transportation systems and data. Integration between traffic management centers, public transportation, regional systems and other cross-jurisdictional transportation systems that have eluded past projects will be critical to achieving lower vehicle-miles traveled (VMT), climate goals, and other critical transportation objectives.

Tomorrow's transportation solutions will require new data sources and new integrations with non-traditional data sources. These new data sources and integration efforts will also include some of the traditional data integrations that have been difficult to accomplish in the past. Some examples of these data integrations include:

- Transit data, including real-time status, real-time ridership, schedules and routes, capacity, parking availability;
- Private transportation providers (traditional and non-traditional), including geo-specific capacity and trip information;
- Real-time traffic information providers;

- Micro-transit providers (bike, scooter, others);
- Parking providers;
- Connected and autonomous vehicles and their supporting systems; and
- New sensing technologies.

Simply put, a modern connected transportation system requires connected data systems across public and private transportation solution providers.

Challenges of a Connected Systems and Data Infrastructure

Challenges can be characterized as:

- **Data Standards and Removal of Roadblocks to Information Portability**. The usability of data is dependent upon the ability to share and transfer data with minimal friction, ensure its quality, and have a common understanding of its semantics and temporal state. Data standards such as J2735[45], J2945[46], Traffic Management Data Dictionary (TMDD), General Transit Feed Specification (GTFS), and others exist to enable interoperability of systems and portability of data. Standards require updates and improvements, as well as adequate funding to support those updates. Implementation guidance and support by an active community of users and implementers is critical to their success. In addition, the pace of technology change often necessitates their replacement with newer standards. Without updates, improvements, and at times, replacement, standards can become roadblocks to information exchange as opposed to enablers. Investment in data standards creation and maintenance is critical to information exchange. Standards with active development in areas of significant commercial interest and importance are better positioned for sufficient funding. Older standards or those with smaller potential commercial markets are at greater risk for a lack of community interest and insufficient funding regardless of their importance in ensuring a connected transportation system.

- **Data management and governance** are often underappreciated, creating roadblocks to effective usage. Data management includes the activities of collecting, inventorying, storing, analyzing, and utilizing data effectively, efficiently, and securely. Data governance includes the activities of defining and following the processes, roles, policies, standards, and metrics used to control how data is collected, utilized, secured, documented, and managed effectively. These practices are often overlooked and receive limited attention and investment in transportation, resulting in lost opportunities to ensure investments in our transportation system are as effective and efficient as possible and

[45] SAE. J2735 (sae.org)
[46] SAE. J2945 (sae.org)

increased potential for privacy and security incidents and associated economic and disruptive impacts.

- **Data privacy** is a critical and relatively new challenge for transportation data. The rise of the collection of data from individual and personal devices and connected vehicles has created greater demand and discussion regarding data privacy requirements. Ensuring data privacy while ensuring maximum usability and effectiveness of data analysis and results creates challenges for users of this private data.

- **Cybersecurity, External Dependencies, and System Resilience**. Cybersecurity and resilience of our transportation system have become a critical issue with the increases in the sources and volumes of data and the advances in information technology to operate our transportation system. Recent cyber attacks of our transportation infrastructure have exposed key vulnerabilities, including in other industry sectors upon which our transportation sector depends. The recent Colonial Pipeline ransomware attack has highlighted such vulnerabilities and the significant impacts such cybersecurity events can have on our transportation infrastructure. It also highlights the lack of resilience in our transportation infrastructure and the industries upon which it depends.

- **Security of Devices, Including Vehicles**. With the rise in the Internet of Things new vulnerabilities have been introduced. As devices have been connected to the internet and private networks, the security of those devices is often overlooked. In the transportation space, these devices may include road control assets such as intersection signals or signs, sensors, cameras, and even vehicles. In addition, vulnerabilities are introduced in other devices such as the mobile devices of transportation personnel. Even microphones and cameras that have become prolific in our traffic management centers, personal devices, vehicles, and infrastructure elements can be targeted. This vastly expands the footprint of vulnerabilities within the transportation sector and complicates the effort to secure the transportation infrastructure.

Much of the effort to address these challenges can be achieved by implementing known practices and technologies and by affecting the necessary changes within organizational and management operations. However, there are also advances in technology and improved technology practices that can provide significant benefit and address some of these challenges, including:

- **Infrastructure modernization**. A revolutionary change from data center to cloud computing has been occurring for some time now, but with limited adoption in the transportation sector. Cloud computing provides significant advantages over traditional data centers, including:

  - Reduced system development and deployment time and cost;
  - Improved access to infrastructure (IaaS), software (SaaS), and platforms (PaaS) on demand;
  - Advanced deployment automation;
  - Improved security automation, monitoring, and services;
  - Automation of system updates;
  - Scalability on demand;
  - Reduced physical security cost;
  - Reduced or eliminated data center utilization and costs;
  - Access to tools and services on demand that otherwise would be cost prohibitive;
  - Ease of provisioning redundant and resilient services and capabilities;
  - Automated failover and recovery;
  - Increased granularity of security, control, and access.

- **Zero-trust security models**. Past enterprise security practices involved creating barriers between what was classified as “inside” an enterprise security perimeter and what was classified as “outside” the enterprise security perimeter. The goal of these practices was to prevent an attack. In general, these barriers consisted of networking barriers such as firewalls that served as gatekeepers between the “outside” and “inside”. Authentication and authorization of users also created a barrier to services, generally via usernames and passwords. Strict device limitations, with no personal devices allowed were not uncommon. Communication between hardware and services that was contained within the enterprise perimeter was generally unrestricted at the network level and security controls between such hardware and services was limited with more trust granted between such internal services. Limits to trust relationships were generally scrutinized much more when interacting between internal and external endpoints.

  Individuals and organizations that attack such systems take advantage of this trust within the enterprise security perimeter. Once an attacker penetrates the enterprise security perimeter, they utilize these internal trust relationships to extend their attack within the perimeter. While the goal of these traditional methods was to prevent all attacks, the weakness exposed is that once an attack is successful, the ability to detect the attack, limit the damage, and recover was limited, difficult, and expensive at best.

Zero-trust practices begin with the acceptance that you cannot prevent all successful attacks. The goal of zero-trust practices is to not only prevent the majority of attacks, but to quickly identify a successful attack, limit the penetration of the attacker along with the damage potential of the attack, and stop the attacker as quickly as possible. To do this, zero-trust establishes a practice of reducing or eliminating the trust between systems and services within the enterprise security perimeter. Combining zero-trust with additional resiliency provided by technologies such as cloud services, increased user security such as two-factor authentication, increased monitoring, automated security practices, and user behavior analysis techniques can significantly reduce the impact of attacks when they occur, reduce the recovery time from such attacks, and reduce the cost of damage and recovery from such attacks.

- **Standards modernization**. As mentioned earlier, to enable a connected transportation system, the movement and portability of data between organizations and systems is of paramount importance. This requires data standards that reflect not only the needs of transportation, but also the current, and sometimes future state of technology. In addition it requires standards that address not only data relationships and formats, but also semantics and meaningful implementation guidance. Current transportation standards, for example the Traffic Management Data Dictionary, often specify legacy technologies that are not capable of meeting current or future transportation demand and requirements. Ensuring modern standards and ongoing maintenance and lifecycle management of data standards is critical to the future of transportation.
- **Modernization of systems design and implementation**. Technology systems design has undergone significant transformation with an increasing speed of innovation, driven by and feeding the demand for more data. This transformation has resulted in a number of improvements in systems design that are built for the scale of data and processing required for tomorrow's automated and connected vehicles, new transportation technologies, improved transportation services, advanced data analysis, automated decision support and artificial intelligence, and improved transportation planning. Many of these technologies are available today, but given limitations of many governmental agencies and budgetary constraints have limited adoption in the transportation sector. The advancements in information technology systems design and implementation include:
    - Cloud computing technologies mentioned earlier;
    - Analytics and streaming engines such as Apache Spark, Google BigQuery, and Apache Storm;

- High speed messaging and real-time processing systems such as Apache Kafka, Apache Flink, or Amazon Kinesis;
  - Machine learning and artificial intelligence platforms;
  - Improved automated scaling in cloud services, such as AWS EC2 autoscaling;
  - Improved purpose specific, high speed data stores, such as Apache Cassandra, MongoDB, cloud-based data lake solutions, graph databases such as Neo4j.

- **Information systems practice**. These advancements in design and implementation have been accompanied by improvements in technology practices that are better suited to the additional complexity of today's information systems solutions and increase the speed and efficiency of teams that provide these solutions and services. Examples of these practice improvements include:
  - Agile development practices that emphasize results over process and documentation;
  - DevOps and SecDevOps practices that combine traditionally siloed development, operations, and security practices into a combined practice, taking advantage of cloud and automation technologies.

- **Blockchain**. Blockchain use in transportation is in its infancy, but its use in the Internet of Things realm, especially in the areas of ensuring security and sourcing of data, has significant potential. Blockchain is a form of digital ledger that is difficult to hack or alter. The use of blockchain technology within a datastream can provide significant assurance of the source of the data as well as limiting the ability to tamper with the data.

- **Artificial Intelligence and User Behavior Analytics in Information Security**. Artificial intelligence has applicability in the systems security realm, providing the ability to identify attackers in real time and notify administrators of the attack. User behavior analytics is another possible security method, utilizing observation, measurement, and understanding of typical individual user behavior to identify and alert security personnel when anomalous behavior is detected.

- **Resilient systems design and implementation**. Many of these technology advances provide the opportunity for increased resilience in systems design and implementation. Cloud technologies, automation, and infrastructure as code provide for system redundancy, failure detection, automated recovery, and even complete automated redeployment of systems in the event of failure if necessary. Implementation of these capabilities, along with improved security monitoring, fault and attack detection, and automated security actions coupled with human security actions can create systems with significantly improved fault tolerance and resiliency.

Advances in information technology have significantly improved the landscape for a much more capable intelligent transportation system. Advances in the data available within our transportation system are just starting to accelerate and the volume and complexity of the data available will likely increase in the near future. These advances present both opportunities and challenges for Caltrans, California, and the nation. Taking advantage of the opportunities and managing and addressing the challenges will require significant attention and resources.

Connected and Automated Vehicles

In recent years, the amount of attention and funding being focused on connected and automated vehicles (CAVs) has grown immensely, particularly in the private sector, as CAVs have gone from being a "pie-in-the-sky" concept to a reality. Some observers consider CAVs to be a tipping point in transportation, of a magnitude only seen at intervals of many decades. With this increased level of attention and advancement in CAVs comes great opportunity for Caltrans to leverage this technology to fulfil their strategic goals.

Background on CAVs

The terminology surrounding CAVs has evolved over the years and It is important to recognize that there are some differences between connected vehicles (CV), automated vehicles (AV) and their combination as connected and automated vehicles (CAV). These concepts are likely to have different implications for the transportation system and they are likely to develop along different timelines.

CV technology refers to robust, standardized vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-everything (V2X) communication, broadly representing wireless communication between vehicles, infrastructure and other road users (such as pedestrians and cyclists). AV technology provides automated driving control in relation to steering, acceleration and braking. The Society of Automotive Engineers (SAE) has defined standard definitions for six levels of driving automation,[47] ranging from no automation (level 0) to full automation, all of the time (level 5).

Readiness of CAVs

Currently, CV and AV are pursuing parallel technological and policy paths and their relationship is evolving. The major national stakeholders, including USDOT and state

[47] SAE. http://standards.sae.org/j3016_201401

transportation agencies, are actively considering the impact of both CVs and AVs and are encouraging a supportive relationship between these technologies. DOTs are carrying out pilot deployments of CV through the provision of roadside communication technology with the understanding that this technology will also support future AV deployments. While there has been some progress in the deployment of CVs and AVs separately, there has been less progress made with respect to true CAV deployment.

CV Readiness

CV technologies include a variety of wireless communications systems, such as DSRC (Dedicated Short Range Communications), that operate on a dedicated band of radio frequency known as the 5.9 GHz band, as well as the newer cellular technologies, such as IoT approaches - cellular V2X (C-V2X) and the forthcoming fifth-generation wireless network, 5G. Much of the CV technology is well researched, and robust standards are available. Some auto manufacturers started deploying CV technology in limited models as early as 2017 but most have decided to delay CV deployment while waiting for a federal mandate. The FHWA has developed standards for the connections between roadside units and traffic controllers, and controller companies have developed products based on these standards. In addition, many states have deployed CV technology, both DSRC and C-V2X, on key corridors throughout the country in a variety of geographic environments, roadway scenarios, and weather conditions.

At the federal policy level, the push to mandate CV equipment in all new vehicles has stalled in recent years. In addition, a recent action by the Federal Communications Commission (FCC) to share the previously dedicated 5.9 GHz band with non-licensed users is likely to increase uncertainty in the CV marketplace. This uncertainty is likely to cause a slow rollout of CV technologies in the coming years. Nevertheless, the auto industry is committed to deploying CV technology so transportation agencies such as Caltrans need to be prepared.

AV Readiness

From the perspective of the OEM, many of the basic challenges associated with enabling Level 1 and 2 driving have been solved as extensions of driver-assist technologies. In fact, Level 2 vehicles are commercially available and becoming more common. Such vehicles are often privately owned and still under the control of a human driver. In contrast, AVs capable of driverless operation and intended to underpin a new on-demand mobility system are still in development, including all Level 3-5 AVs intended for fleet operation in defined environments. While there is a growing number of automated test fleet deployments, such AVs are far from mature or commercially viable, and these technologies are highly competitive and closely guarded.

Current efforts by OEMs and AV manufacturers are oriented around developing AI software for vehicular control. Developers have achieved meaningful advancements in two categories of AI that are best positioned to solve the complex challenges associated with higher levels of automated driving: machine learning and deep learning. It is intended that ADS (Advanced Design System) software will have a “self-learning capability” that recognizes and corrects driver misperceptions and misbehaviors.

There have also been major developments in AV sensor technology. Many view Light Detection and Ranging (LiDAR) technology, which is able to accurately gauge a vehicle’s surroundings in many complex situations, as essential to enabling higher levels of automation. Major efforts continue to reduce the size and cost of historically bulky and expensive units. The sensor suites used in AVs are changing rapidly and will continue to do so. In the future, AV technology will be built on integrated sensors (where measurement and processing both happen at the edge) and IoT communication that will allow efficient data sharing.

Infrastructure

The continued engagement of infrastructure owner operators (IOO) is crucial to preparing roads for CAVs. While AVs can be tested in off-roadway facilities, all companies engaged in AV development place high value on miles logged on public roadways. Most manufacturers maintain a position of non-reliance on infrastructure assistance, although improved road maintenance in relation to markings, signage, and road surface condition (and their national consistency) is an important factor in the deployment of AVs. There is an emerging sense that IOOs can apply “no regrets” actions in the infrastructure, such as wider, higher-contrast lane markings, with the understanding that it benefits drivers of both AVs and non-AVs. Important infrastructure issues include real-time information regarding work zones and other road conditions, pick-up and drop-off areas (including curb design and access), use of dedicated or managed lanes, and dynamic lane assignment. IoT technology is the enabler for this type of information collection, processing and exchange.

FHWA has awarded multiple ADS Demonstration Grants to specific combinations of IOOs, universities, and private companies. These collaborative demonstrations include data architecture development on the part of the IOOs and focus on the safe integration of higher levels of AVs in the transportation system. The grants provide a good opportunity for IOOs to engage in CAV development and testing.

Human Acceptance of CAVs

Vehicle manufacturers, technology companies, and policymakers are continuing their effort and investment in the development and deployment of automated driving systems. In addition, policymakers in the United States have been advancing the related legislation. Policymakers' efforts are continued to put in place federal regulation, beyond testing purposes on public roads. To achieve the widespread adoption of CAVs, it requires not only addressing technological and political challenges, but also resolving psychological and perceptual challenges. Various studies have investigated public opinion regarding full driving automation. A survey of 5,000 individuals in 109 countries found that driving with full driving automation would be easier than driving conventional vehicles for users. However, driving with partial automation would be more difficult than driving conventional vehicles. To thoroughly understand what makes the public accept or reject CAVs, theoretical and psychological research is needed to identify the factors which influence public acceptance.

On the other hand, CAVs are expected to save lives and provide enhanced mobility for all people. CAV technologies also allow the emergence of novel business models such as shared automated vehicles (SAVs), which could provide more accessible on-demand mobility services and play a vital role in sustainable transportation. In particular, SAVs could be an attractive means to provide transportation access for those who cannot drive or face significant barriers to driving, including the elderly and people with disabilities. Individuals with disabilities are a unique but sizable demographic in the United States. According to the 2018 disability status report, nearly 13% of the population has a disability. Unfortunately, many people with disabilities have limited transportation opportunities and adverse transportation experiences. We know that CAVs have the potential to transform personal mobility. Whether disabled travelers can benefit from this transformation depends on to what extent the accessibility is taken into consideration in the development and planning process of the deployment.

Unmanned Aerial Vehicles (UAVs)

The future certainly seems bright for the unmanned aerial vehicles (UAV) industry, helped in huge part by a thriving commercial market. From the global market size of $4.4B in 2018, the commercial UAV market is expected to grow to about $63.6B by 2025, when the U.S. market is expected to reach a value of $13.6B. Drones equipped with cameras are now backed with machine learning and deep learning algorithms, leading to drastic changes in the drone industry. Computer vision algorithms learn from captured images while using the drones for various purposes, including traffic monitoring.[48]

[48] Vsinghbisen, V. How AI Based Drone Works: Artificial Intelligence Drone Use Cases, Feb. 5, 2020. https://medium.com/vsinghbisen/how-ai-based-drone-works-artificial-intelligence-drone-use-cases-7f3d44b8abe3

In the context of a transportation system, UAVs can be categorized by application:

- **Package delivery**. In November 2020 the FAA proposed airworthiness criteria for type certification of delivery drones with an intent to initialize commercial operations. There are ten companies that received this type of certification for their aircraft: 3D Robotics (3DR-GS), Airobotics (OPTIMUS 1-EX), Amazon (MK27), Flirtey (F4.5), Flytrex (FTX-M600P), Matternet (M2), Percepto (System 2.4), Telegrid (DE2020), Wingcopter (Model 198) and Zipline (Zip).

- **Mapping and surveying.** This is when UAVs are used to capture aerial data with downward-facing sensors, such as RGB or multispectral cameras, and LiDAR payloads. During a drone survey with an RGB camera, the ground is photographed several times from different angles, and each image is tagged with coordinates. UAVs used for this purpose must have long battery time and be equipped with high-precision GNSS (Global Navigation Satellite System) and interchangeable sensors, such as LiDAR, 4K camera, thermal imaging and multispectral imaging. The top vendors of such UAVs are: DJI (Phantom 4 RTK, Mavic Air 2, Matrice 210 RTK), Yuneek (H520 RTK), senseFly (eBee X).

- **Infrastructure inspections**. They are particularly useful for inspections that are dangerous for human surveyors, such as bridges, tunnels, or disaster areas. Generally, it is required that UAVs can operate in a GNSS-less environment. Vendors that provide inspection UAVs include: Flyability (ELIOS 2), DJI (Matrice 210, Inspire), Lockheed (Indago 3) and Intel (Falcon 8+).

- **Surveillance**. Aerial surveillance of large areas, including road patrol, is possible with low-cost UAVs. Object detection and tracking capability drives the expansion of this application. Skydio 2 and Skydio X2 are the state of the art UAVs for this purpose.

- **Passenger transport**. Also known as flying taxis, passenger UAVs are developed by Volocopter, Kittyhawk, Workhorse Group and Hyundai.

As the use of UAVs keeps expanding, the topic of UAV fleet management draws a lot of attention. Companies that provide fleet management solutions include: Skyward, Aloft owned by Kittyhawk, Drone Harmony, and vHive.

According to the American Association of State Highway and Transportation Officials, 17 state transportation departments have actively researched, tested, or are using unmanned aerial vehicles to inspect bridges, assist with clearing vehicle crashes, monitor traffic flows, and survey areas susceptible to landslides, rockslides and flooding.[49] Another 16 agencies are studying the possibility of using the technology because of potential safety and cost benefits.

Transportation Application Areas

Traffic Monitoring and Performance Measurement

Accurate, reliable, and timely information is the life-blood of traffic management, for both real-time decision-making and future planning.

There is still no substitute for dedicated point sensors to collect samples of vehicular flow, density, and truck classification information. Third-party vendors provide travel time based on data from smartphone applications, navigational devices, and connected vehicles. However lane-by-lane, disaggregated speeds are not available. Speeds on High Occupancy Vehicle (HOV) facilities are not reported separately from speeds on mainline freeway lanes. Truck classification data from third parties are limited, and not defined consistently with standard Caltrans or FHWA practices. Nonetheless, as technologies and the use of big data in the transportation industries evolve, new products and services continue to be introduced. For example, level-of-service measures are now available on major arterials from connected vehicle data to improve signal synchronization.

For the foreseeable future, hybrid approaches that utilize both point sensors and third-party data are likely to provide the best value and most accurate information for traffic monitoring and performance measurement purposes. Recent research[50] has found that a hybrid approach provides the best estimates of performance measures for both freeway-mainline and freeway-freeway connectors.

In the past, a perceived risk of using hybrid data for traffic management systems was its dependence on external vendors. However, the new risk is that mobile devices are so prolific that drivers are now being influenced by the apps, and traffic management systems lack direct access to this influential, and useful, information.

In an increasingly interconnected world, the future of transportation management will require better and more complete data that can only be obtained through greater connectivity to the data feeds of private vendors as well as increased cooperation and collaboration with local stakeholders. The first step in this direction is to adopt computing tools and infrastructure that have already been tried and implemented at

[49] American Shipper. States to use drones for monitoring highway systems, March 2016. https://www.freightwaves.com/news/states-to-use-drones-for-monitoring-highway-systems
[50] Khan et al. Hybrid Data Implementation, 2021. https://escholarship.org/uc/item/32b6s0fk

scale in the private sector. Of course, data quality would need to be monitored on a continuous basis, and attention should be paid to costs when selecting a portfolio of data sources.

Data Quality

High quality, actionable information is the life-blood of traffic management. Recent advances in Artificial Intelligence and Machine Learning have provided some promising tools to expand the nature of the available data, to combine multiple data sets in new ways, and to improve the basic quality and reliability of the data.

Classification and Anomaly Detection

There are a number of learning methods used in Artificial Intelligence that are appropriate for identifying anomalies in data. In supervised learning, a model is developed to predict or estimate an output based on one or more inputs. In unsupervised learning, only features (independent variables) are observed and there are no measurements of the outcome. In contrast to supervised learning, unsupervised learning does not use labeled data and identifies patterns in a dataset. Therefore, identifying anomalies using unsupervised models does not require ground truth labels.

Ensembles combine results of different models to produce more powerful models to identify outliers. In an ensemble, a weighted combination of base models is learned. An ensemble can include different types of models or different versions of the same model. Random forest is a well-known ensemble model which has been widely used in literature.

Supervised classification models function by attempting to assign the data into known classifications. These methods rely on known classification labels as a dependent variable used to train the model, and thus are deemed “supervised” learning methods due to this a priori data available for training.

- K-Nearest Neighbors (KNN) – A non-parametric model which uses stored training data to classify the target variable based on nearest neighboring data points in the training set.
- Logistics Regression – A probabilistic model which estimates the probability of the target variable belonging to one category or another using a sigmoid function to represent the distribution.
- Decision Tree – Uses a conjunction of rules to organize the data into a tree structure where each branch is a binary decision with a threshold for each variable (i.e., belongs to one group or another). Following the branches along each decision point will determine final classification.

- Random Forest – An ensemble of decision trees where a single training data set is randomly divided into M samples and each used to train the classifiers before being aggregated (i.e., bootstrapped) in a final combined result.

Unsupervised anomaly detection functions by attempting to identify statistical outliers in the data. Rather than training the data using a label as a predicted variable (i.e., HOV or mainline), the label is part of the data itself and not the dependent variable, thus making the approach unsupervised. As a result, unsupervised methods do not need to be trained and are run directly on the data.

- Isolation Forest – An unsupervised method utilizing decision trees, but instead of profiling normal values into classifications it attempts to isolate anomalies.
- Local Outlier Factor (LOF) (Density-based detection) – Uses KNN to compute local reachability for each point to its nearest neighbors, enabling statistical outliers to be detected.
- One-class Support Vector Machine (SVM) – SVMs find a decision boundary between points that maximizes the margins, dividing the data into discrete classifications. A one-class SVM is a special case of SVM where no labels are provided, making it an unsupervised method.
- Robust Covariance (Distance-based detection) – Another variation of KNN which relies on the median distance to the nearest neighbors. The median value, rather than the mean, provides a more “robust” calculation of distance less sensitive to anomalous data.

Congestion Management

Congestion management will always be one of the main priorities of Caltrans, even if unspoken. To address the congestion problem, one first has to pinpoint traffic bottlenecks and identify their high-level cause:

1. Poor traffic control; or
2. Excessive traffic demand.

For each of these causes, there exists its own arsenal of tools.

Traffic Flow Control

A symptom of poor control is when a traffic flow passing through some critical point (a road segment or a junction) does not exceed the throughput capacity of that critical point, but nevertheless it cannot be accommodated causing vehicle queues’ buildup and travel delay. Generally, the problems of poor control can be localized. For example, weaving at merge sections is a classic problem that should be resolved by traffic flow control.

The means of traffic flow control are limited to:

1. Flow metering by means of a traffic signal (intersections and merge points).
2. Flow prioritization by means of signal or changeable message signs.
3. Lane management by means of opening or closing certain lanes for selected traffic modes during demand spikes.
4. Variable speed limit by means of road signs, or connected vehicles and infrastructure.[51]
5. Increase or decrease of vehicle headway at critical junctions by means of adaptive cruise control (ACC) or cooperative ACC (CACC) that would rely on connected vehicles and infrastructure.[52]
6. Reduce vehicle sizes. This, however, largely depends on vehicle manufacturers and can formally lead to an increased VMT.

Solutions whose implementation implies the use of connected vehicles and infrastructure will have to employ IoT and high-precision GNSS technologies.

During the planning phase, any solution that promises to eliminate a local bottleneck should be considered in the context of a wider road network to understand how demand could possibly shift. We discuss this in the next section.

Demand Management

Excessive demand is a more common and difficult problem, which cannot be resolved locally around some single problem spot. One has to first identify the inputs to the selected traffic network that contribute to the excessive demand upstream of the given bottlenecks. The next step is to compute how to reassign traffic within this traffic network so as to accommodate traffic demand while maximizing traffic throughput and minimizing delays.

It is well known that drivers choose their routes according to the principle,[53] which states that the journey costs in all routes actually used are equal and less than those that would be encountered by a single vehicle on any unused route. Each driver selfishly minimizes his cost, and the traffic pattern that satisfies this principle is said to be **user equilibrium (UE)**. Path finding in TomTom, Waze, Google and Apple Maps is based on UE. Today, the widespread use of routing apps leads to unexpected traffic jams in

---

[51] Berkeley Institute of Transportation Studies. CIRCLES Project. https://its.berkeley.edu/news/circles-project-kicks-multi-campus-traffic-smoothing-study

[52] National Science Foundation. Traffic Operating System (TOS) for Smart Cities. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1545116

[53] Ebrary.com.Wardrop's First Principle and its Dynamic Extensions. https://ebrary.net/10792/economics/wardrops_first_principle_dynamic_extensions

previously quiet neighborhoods because these apps flag traffic bottlenecks ahead and reroute users to shortcuts through residential areas to save a minute or two in travel time.[54] The inefficiency of selfish routing is particularly acute when a surge in demand occurs due to an evacuation, end of a concert or sporting event, or routine rush hour traffic. A common case of this type of inefficiency is known as Braess Paradox.[55] This paradox shows that the addition of links to a road network may in fact increase the total travel time experienced by the drivers.

The goal of the traffic network control is to achieve the equilibrium, at which average journey cost is minimal.[56] This type of equilibrium is called **system (or socially) optimal (SO)**. Generally, SO traffic assignment is different from UE, and forcing drivers to cooperate requires network-level traffic control tools that consist of:

1. Closure of lanes or road segments to selected traffic modes and/or during certain time periods. This essentially reduces routing options. For effective lane or road closure, one has to develop a methodology for finding road segments that enable Braess paradox. Information about such closures and their schedules should be made publicly available to be reflected in the navigation apps.
2. Readjust signal timing at critical intersections so as to reduce green intervals for maneuvers that direct vehicles to undesired routes.
3. Dynamic congestion pricing. This option includes express lanes as well as area tolls - when fees are paid by drivers entering restricted areas. Presently, there is no area toll implementation in the U.S.
4. An alternative to area tolls is a token system to allow cars into congested areas. There is a limited pool of tokens, and they are issued on the first-come, first-served basis through a mobile application that uses geo-location information. Once the car leaves the critical area, its token is released back into the pool. The advantage of the scheme is its flexibility: the number of tokens can be changed by area and time of use. One could treat exceptions, e.g. residents could be guaranteed a token. One can also imagine implementing dynamic cordon pricing, in which tokens may be acquired for a fee that depends on the number of tokens left in the pool.

To effectively implement these mitigation strategies, one requires reliable traffic pattern data and information about available storage and and throughout capacity of individual road segments within a given road network. The token scheme implies the use of IoT

[54] J. Macfarlane.When apps rule the road: The proliferation of navigation apps is causing traffic chaos. It's time to restore order. IEEE Spectrum. Oct. 2019. doi: 10.1109/MSPEC.2019.8847586.
[55] Braess Paradox:The Braess Paradox (harvard.edu)
[56] This is known as Wardrop's second principle: Wardrop's Principles - System Optimal (bowenwen.com)

and presupposes that a non-token holding car is easily identified, making enforcement costs low.

In parallel with the listed measures one should encourage a mode shift that will be discussed below, in the section Public Transportation.

Incident Management

Far from being rare, incidents are estimated to cause somewhere between 52% and 58% of total delay experienced by motorists in all urban area population groups. Over two thirds of incidents do not cause delay. However, among those that do cause delay the so-called 10-90 rule applies: 10% of incidents account for 90% of incident-induced delay. The delay caused by the same type of incident varies substantially depending on location (e.g. relative to ramps, availability of shoulder) and, more importantly, on the state of traffic. Similar incidents can cause no added delay when traffic is either very light or already congested, or significant additional delay at other times. Ultimately, serious delay-causing incidents are too rare for statistical data to predict, and historical data cannot be used to design effective response strategies.

An effective incident response is a combination of optimal resource allocation, adaptive traffic control and real time demand management. Resource allocation implies proper placement of patrol services to minimize the incident clearance times. Adaptive traffic control, besides the already discussed traffic flow metering and intersection signal timing, includes opening designated lanes (e.g. HOV, HOT) and shoulders to general traffic as a means of temporary capacity increase. Demand management in these situations is mostly about traveler information, route and mode choice recommendations. In certain cases, these route recommendations may become mandatory and, thus, implement direct traffic control.

Capturing Incident Scene

When crashes occur, the responder who makes it to the scene first is in charge. Incident clearance time depends a lot on the efficiency of their orders. As responders rush to the rescue, it would be very helpful for them to know ahead of time:

1. The exact location of the incident;
2. Parties involved and their physical condition;
3. The look of the incident scene - photos or videos.

There is a need for an online facility that allows reporting of incidents by the bystanders, fellow road users or the involved parties themselves, including geo-tagged and timestamped picture or video upload. The uploaded material becomes immediately available to the responders, even before they arrive at the scene, allowing them to be prepared for a specific scenario and start acting as soon as they arrive.

To help reconstruct the incident, it is now possible to use UAVs to help reconstruct accident scenes by collecting visual data that can then be processed using mapping software into 3D models or orthomosaic maps of the scene.[57]

Using these data, it is possible to build a knowledge base that will provide accurate estimates of incidents' clearance time and serve as training material for responders.

Ad-Hoc Traffic Detection System

The cornerstone of an effective incident response is a robust and adaptive traffic control, which relies on dense volume/occupancy/speed (VOS) measurements to provide a full picture of the changing traffic patterns. Presently, detectors (loops and magnetic sensors) on urban streets in the U.S. are used primarily for signal actuation. We know of no city in the U.S. that systematically (say, once a month) measures VOS on its main streets. In the rare cases when traffic measurements are collected and processed, they end up in closed proprietary databases.

In the absence of sufficient sensing infrastructure, it is possible to deploy a group of UAVs for ad-hoc traffic monitoring in the area of interest.[58] Such a mobile traffic detection system, which could be deployed on demand in real time and provide VOS data collection at specified locations, would help pinpoint newly emerged bottlenecks and decision points for traffic routing in case of unusual traffic demand patterns. It would also be handy during large scale emergencies, such as wildfires, providing observability of the traffic network for the emergency responders.

Safety Management

Activities related to road safety can be categorized as follows:

- Scalable standardized assessment;
- Approaches and technologies to improve safety;
- Outreach and education.

[57] UAVCoach.com. Drones in Accident Reconstruction: How Drones Are Helping Make Traffic Crash Site Assessments Faster, Safer, and More Accurate. https://uavcoach.com/drones-accident-reconstruction

[58] Outay, F., Mengash, H.A., and Adnan, M. Applications of Unmanned Aerial Vehicle (UAV) in Road Safety, Traffic and Highway Infrastructure Management: Recent Advances and Challenges. Transportation Research, Part A: Policy and Practice. Nov. 2020. https://doi.org/10.1016/j.tra.2020.09.018

Scalable Standardized Assessment

The Federal Highway Administration (FHWA) recommends the online tool known as U.S. Road Assessment Program (usRAP)[59] for the safety assessment of the highway system. The purpose of usRAP is to help its users:

- Identify roadway sections with high risk of fatal and serious injury crashes;
- Develop maps for crash density, crash rate, crash rate ratio, and potential crash savings;
- Track the safety performance of highways over time;
- Direct resources toward systematic improvement of the road system;
- Develop a safer roads investment plan (i.e., recommendations of mitigation measures to improve safety at specific locations).

There are three functions that usRAP provides:

1. **Risk maps** are used to document the risk of fatal and serious injury crashes and show where risk is high and low. Five color coded risk levels are used for the development of maps for crash density, crash rate, crash rate ratio, and potential crash savings. The categories include highest risk (5 percent of system), medium-high risk (10 percent of system), medium risk (20 percent of system), medium-low risk (25 percent of system), and lowest risk (40 percent of system). The maps are presented based on crash type (e.g., roadway departure crashes).

2. **Star ratings** are based on inspection of roads to examine how well they protect users from crashes and from deaths and serious injuries when crashes occur. Road protection scores (RPSs) are used to derive star ratings to identify important differences in road design or management which are likely to lead to different probabilities of fatal or serious crashes and RPSs can be determined relatively quickly, with low cost. Roadway data can be used without detailed crash data to estimate star ratings.

3. **Performance tracking** compares the safety performance of highways over time and relates those changes to ongoing safety improvement programs.

AI tools, such as object/feature detection and classification allow us to use usRAP not just for individual road segments, but at scale - for the whole California State Highway System. The approach would include requesting and downloading satellite and street view images along state highways and analyzing each of them those for:

- Median type and width;

[59] U.S. RAP: http://www.usrap.org

- Shoulder type and width;
- Lane width;
- Speed limit;
- Ramp type;
- Driveway type;
- Passing zones;
- Presence of guardrail;
- Presence of drop-offs;
- Clear zone width.

Intersection locations can be identified using OpenStreetMap (OSM),[60] and their characteristics can be extracted, including:

- Intersection type;
- Turning lanes;
- Intersection control;
- Presence of pedestrian crossings and bike lanes;
- Types of admissible maneuvers.

Combining these parameters with traffic flows - measured, where possible, or estimated from the annual average daily traffic (AADT) - and repeated periodically with the subsequent upload to usRAP, would create a safety information system with spatio-temporal analysis functionality for the California State Highway System.

Approaches and Technologies to Improve Safety

Safety management in California cities generally follows the **Vision Zero (VZ)**[61] paradigm implemented via **Complete Streets** program. VZ is a strategy to eliminate all traffic fatalities and severe injuries, while increasing safe, healthy, equitable mobility for all. VZ recognizes that people will sometimes make mistakes, so the road system and related policies should be designed to ensure those inevitable mistakes do not result in severe injuries or fatalities. This means that system designers and policymakers are expected to improve the roadway environment, policies (such as speed management), and other related systems to lessen the severity of crashes. First implemented in Sweden in the 1990s, Vision Zero became popular across Europe, and now it's gaining momentum in major American cities. Presently, 13 California cities committed to VZ goals - Alameda, Berkeley, Fremont, Los Angeles, La Mesa, Monterey, Sacramento, San Diego, San Francisco, San Jose, San Luis Obispo, Santa Barbara and Watsonville.

[60] OpenStreetMap: https://osm.org
[61] Vision Zero Network: https://visionzeronetwork.org

VZ approach implies road diet - it reduces vehicular speed and throughput, but we do not know how and where congestion problems are shifted from streets where VZ is implemented. Moreover, sometimes VZ results are directly opposite of what was promised.[62] Thus, there is a strong need for a well-defined and understandable criteria of VZ proposal evaluation and before/after assessment. Not just streets and intersections with VZ modifications, but the **whole surrounding street network** must be analyzed in terms of changing traffic patterns, risks and throughput capacity.

Next, we talk about some technologies that could be employed to improve safety. These include risk analysis tools, connected vehicles and infrastructure that can signal about possible danger, driver monitoring and data sharing about work zones.

Depth Analysis

In the previous section about safety assessment, we talked about analysis at scale involving hundreds of roads and intersections - the so-called breadth analysis. Contrary to that, depth analysis zooms into a particular location - a junction or a road segment - and studies its road geometry, admissible maneuvers, observed traffic modes and their flow/frequency. Companies, such as Street Simplified[63] and Derq[64], conduct video surveillance for a given amount of time (typically, 24 hours) at requested locations and employ computer vision techniques for object classification and tracking. The subsequent processing of resulting trajectories enables identification of conflicts and potential crashes (i.e., near misses).

One Berkeley Deep Drive (BDD) project[65] developed a toolchain for automatic generation of hazardous scenarios that arise because of shortage of information (e.g., occlusions) and incorrect interpretation of the other object's actions. Some of these risky scenarios can be avoided by employing connected vehicles and infrastructure.

Connected Vehicles and Infrastructure

In complex environments, such as intersections, challenges arise from complicated vehicle trajectories; occasional absence of lane markings to guide vehicles; split phases that prevent determining who has the right of way; invisible vehicle approaches; illegal movements; simultaneous interactions among pedestrians, bicycles and vehicles. Intelligent intersection improves awareness of drivers, bicyclists and pedestrians when

---

[62] Rebuilding From the Failures of Vision Zero: https://www.planetizen.com/news/2021/05/113209-rebuilding-failures-vision-zero

[63] Street Simplified: https://www.streetsimplified.com

[64] Derq: https://derq.com

[65] Varaiya, P., Seisha, S. Generation and Mitigation of Hazardous Scenarios at Traffic Intersections. https://deepdrive.berkeley.edu/project/generation-and-mitigation-hazardous-scenarios-traffic-intersections

they maneuver through the junction.[66] Intelligent intersection technology has four critical components: (1) sensing; (2) interpretation of sensor measurements; (3) anyone-to-infrastructure (X2I) connectivity, extending the concept of vehicle-to-infrastructure (V2I) connectivity; and (4) signal phase prioritization. It has the potential to remove spatial and temporal uncertainties that lead drivers, cyclists and pedestrians to wrong decisions resulting in accidents.

This research is directly related to AV safety. It was recently shown that even with perfect onboard sensing AVs without connectivity cannot be safe.[67]

Safety of Rail Crossings

Safety at highway-rail crossings is an issue that affects many travelers using various modes in various locations, from unsignalized crossings in rural locations to signalized crossings in dense urban areas where both light and heavy transit rail are operated in the same right-of-way as vehicular traffic.

In 2018, more than half of public at-grade crossings had automatic warning systems (flashing lights and/or gates) but very few private at-grade crossings did. Still, more than 60% of collisions occurred at crossings with automatic warning systems. A study on driver behavior using cameras installed in cars found that 46% of the time, drivers were likely engaged in secondary tasks and 35% of the time, drivers failed to look either left or right on approach to passive grade crossings. In general, 94% of train-vehicle collisions could be attributed to driver behavior or poor judgement. An advanced warning system that provides another source of information to travelers, perhaps an in-vehicle notification to immediately and directly alert the driver, could help further improve safety. This could also be a valuable data source for automated vehicle systems.

While rail intersection safety applications have been recently demonstrated in operation, they have rarely included connected vehicle technology. There are also limitations on the ability to expand existing systems due to the lack of data sharing agreements and other coordination with private railroad companies.

Due to the difficulty of coordinating with the railroad industry, it is unlikely that connected rail crossing safety applications will be based on direct vehicle-to-train communication. The most feasible solution would be to extend the intelligent intersection and X2I

---

[66] Varaiya, P. Which Intersections Need I2V and When? https://deepdrive.berkeley.edu/project/which-intersections-need-i2v-and-when

[67] Shetty, A., Yu, M., Kurzhanskiy, A., Grembek, o., Tavafoghi, H., Varaiya, P. Safety challenges for autonomous vehicles in the absence of connectivity.
Transportation Research. July 2021. https://doi.org/10.1016/j.trc.2021.103133

concept to include rail crossings. Thus, the communication between vehicles and trains would occur through the infrastructure.

Driver Drowsiness and Distraction

Half of U.S. adult drivers admit to consistently getting behind the wheel while feeling drowsy. About 20% admit to falling asleep behind the wheel at some point in the past year – with more than 40% admitting this has happened at least once in their driving. According to a Governors Highway Safety Association report,[68] 5,000 people died in 2015 in crashes involving drowsy driving.

Distracted driving is another challenge; several studies have identified three main types of distraction: visual distractions (driver's eyes off the road), manual distractions (driver's hands off the wheel) and cognitive distractions (driver's mind off the driving task).[69] According to the National Highway Traffic Safety Administration (NHTSA), distracted driving claimed 2,841 lives in 2018 alone.[70]

A study on early detection of driver drowsiness[71] used a driving simulator and investigated the feasibility of classification of the alert states of drivers, particularly the slightly drowsy state, based on hybrid sensing of vehicle-based, behavioral, and physiological indicators. The results showed that the ensemble algorithm could obtain 82.4% classification accuracy using hybrid methods to identify the alert and slightly drowsy states, and 95.4% accuracy classifying the alert and moderately drowsy states. These results represented the feasibility of highly accurate early detection of driver drowsiness and the feasibility of implementing a driver drowsiness detection system.

Recognizing driver behaviors is becoming vital for in-vehicle systems that seek to reduce car accidents rooted in cognitive distraction. For example, a recent article provides a review of Uber's patents issued in the Summer 2019 that aim at driver surveillance and safety ranking.[72]

Work Zone Safety

Work zone information is becoming increasingly important for data-driven agency operations and maintenance. It is also critical information that has an impact on the safe deployment of CAVs. Work zones typically involve many stakeholders to manage and

---

68 Governors Highway Safety Association. Drowsy driving. https://www.ghsa.org/issues/drowsy-driving
69 CDC. Distracted Driving. https://www.cdc.gov/motorvehiclesafety/distracted_driving/index.html
70 NHTSA. Distracted Driving. https://www.nhtsa.gov/risky-driving/distracted-driving
71 Gwak, J., Hirao, A., Shino, M. An Investigation of Early Detection of Driver Drowsiness Using Ensemble Machine Learning Based on Hybrid Sensing. Apr. 2020. https://www.mdpi.com/2076-3417/10/8/2890
72 Theintercept.com. Uber Patents Reveal Experiments With Predictive Algorithms to Identify Risky Drivers: https://theintercept.com/2021/10/30/uber-patent-driver-risk-algorithms

maintain them. Many transportation agencies face the challenge of deciding how to gather and share work zone activity information. Work Zone Event Data (WZED) collection is currently largely ad-hoc and limited in scope to address specific agency needs. Besides, it is not easily shared outside of agency-specific systems. As a result, FHWA's Work Zone Data Initiative (WZDI), launched in 2017, is an effort to enable sharing and application of WZED across the country. WZDI is developing a standard approach for collecting, organizing, and sharing data on the WZED including the **what**, **where** and **when** of work zone activities. As a result of this initiative, there will be a structure for organizing the next generation of WZED and data systems to enable local, regional, and national data sharing.

There is a need for a standardized approach to facilitate sharing critical work zone-related information seamlessly across multiple jurisdictions, regions, and information delivery platforms. Standardized WZED will enable effective coordination of activities and enhanced mobility and safety in and around work zones. It meets specifications for content, structure, and format that ensure data are accurately interpreted when communicated to other jurisdictions and stakeholders. It is a critical step to advance the work zone management for the deployment of connected and automated vehicles. Specifically, the standardization of WZED supports and enables analysis of likely impacts and resulting enhancements of work zones on traffic safety and mobility, as well as sharing work zone information with connected and automated vehicles.

Outreach and Education

Human factor is the most common aspect in all the incidents that occur on our roads. Therefore, outreach and education about road safety for various population groups are very useful for reducing injuries and fatalities. Examples of such traditional programs include:

- Pedestrian safer journey;[73]
- Bicycle safer journey;[74]
- Be work zone alert,[75] run by Caltrans Office of Traffic Safety (OTS).

The new way of outreach and education involves mobile apps. Examples of such apps include:

- Arility - augmented reality safety education platform for school children.[76]

[73] PedestrianSaferJourney.com. https://www.pedbikeinfo.org/pedsaferjourney
[74] PedestrianSaferJourney.com. https://www.pedbikeinfo.org/bicyclesaferjourney
[75] BeWorkZoneAlert.comhttp://beworkzonealert.com
[76] Arility.com. https://www.arility.com

- VIA - road safety education for the next generation.[77]
- Virtual Real World - a mobile game for children of age 6-8 years, teaching them about road safety.[78]

Climate Change and Environment Protection

Caltrans' climate action is twofold:

1. Create and maintain sustainable practices to reduce greenhouse gas emissions from transportation operations and projects. This consists mainly of making sure that Caltrans projects do not lead to the increase of vehicle miles traveled (VMT), and, where possible motivate mode shift towards active transportation (walking, cycling, paddling), public transportation and zero-emission vehicles (ZEV).

2. Implement adaptation measures to increase the resilience of the State Highway System to climate impacts and address vulnerabilities. This entails some level of disaster preparedness and anticipation of risks associated with wildfires, floods, soil erosion, buckling and rutting of roads, warping of rails, and difficulty in maintaining and constructing roads during the day because of heat.

Impacts of Specific Projects on Travel Demand and VMT

Following implementation of the 2003 Senate Bill 743, Caltrans policy and practice are changing, with a shift toward considering addition of new State Highway System (SHS) lane-miles as a "solution of last resort" rather than a standard response to operational and safety problems. This shift is based on the observation that increases in roadway capacity often lead to increased travel demand and thus, increased vehicle emissions and greenhouse gas emissions. To address this issue, Bill 743 changed the primary California Environmental Quality Act (CEQA) transportation analysis metric from level of service to reducing vehicle miles traveled (VMT). Governor Newsom's 2019 Executive Order N-19-19 further articulates the Administration's objective of reducing overall vehicle use, as does the 2021 draft Climate Action Plan for Transportation Infrastructure (CAPTI) from the California State Transportation Agency.

The above changes make it more difficult to promote capacity-enhancing projects to alleviate existing congestion. It does not prevent them but makes their approval subject to more onerous justification. The current directive is to assess the impacts on VMT of capacity-enhancing projects through the use a general elasticity formula that was developed by the National Center for Sustainable Transportation at the University of California Davis based on an analysis of past projects and which simply considers the

---

[77] GlobalRoadSafetyPartnership.com. https://www.grsproadsafety.org/programmes/via
[78] SugarCreative.com. https://www.sugarcs.com/projects/virtual-road-world

total number of lane miles added by the project within the county where the project is implemented. Projections from transportation planning models are only accepted if it is demonstrated that the models can adequately predict changes in travel demand and if the resulting projections do not deviate significantly from the formula projections.

Although an evaluation framework exists, it suffers from deficiencies in its ability to correctly estimate the impacts of capacity-enhancing projects on VMT for certain types of projects. Firstly, the elasticity formula developed by the National Center for Sustainable Transportation can only be applied to urban counties. Secondly, the impacts associated with several types of currently popular projects cannot be considered due to a lack of observations. This includes the implementation of managed lanes or hard shoulder running and the conversion of an HOV 2+ lane into a 3+ or 4+ lane. Research is needed to determine the potential of various types of projects to increase VMT and, if possible, to determine specific elasticity factors that could be used to improve the current evaluation framework.

On the transportation planning modeling side, research is also needed to better enable models to consider induced land use changes, i.e., land use changes that might indirectly result from the capacity improvement project.

Encouraging Modal Shift

Presently, many initiatives directed at mode shift propose some punishing pricing schemes or removal of parking at presumed destinations to **force** people to choose public transit over their car. However, this approach is unfair: it punishes the disadvantaged, allows the wealthy to continue current practices and discourages transit agencies from becoming competitive. We believe that an approach promoting competition between travel modes is more promising.

Programs informing and incentivizing people to explore transportation modes other than their own cars, need reliable and up to date information about alternatives. Travelers should be able to see the benefit of these alternatives in terms of travel time (getting to destination faster), cost (cheaper), health (sometimes cycling could be considered a safe and healthy workout in a beautiful weather), or convenience (can sleep, read, listen to or watch something during the commute). For that, we need:

1. Data about alternative modes in a standard format, such as General Transit Feed Format (GTFS) specifying public transit schedules. There can be GTFS-RT (real-time) or GTFS-Flex (on-demand routes) data available. There could be information about occupancy of parking lots at park-and-ride facilities, and so forth.

2. Secure data management system that keeps the information up to date, validates new data and maintains usage statistics.
3. Open Application Programming Interface (API) for making these data available to trip planner providers.
4. Mobile apps that make it interesting to compare the alternatives. Gamification can be a powerful tool to incentivise people to change their means of transportation.

Encouraging drivers to switch to ZEVs should be done with caution. California does not have sufficient infrastructure not just for charging (see the section On-the-Go Charging/ Fueling Infrastructure), but even for power generation. Thus, any shift toward ZEVs must be commensurate with the development of the corresponding infrastructure.

Disaster Preparedness

Wildfires can be predicted as far as 30 days ahead based on the weather forecast. Potential hot spots can be identified using comparative satellite imagery. Areas with high vegetation in winter months and dry in the summer are of high risk.

It must be noted that California is a flood-prone state. Most of California is vulnerable to floods. Every county has been declared a flood disaster area multiple times. The extent of floods can be forecasted using predictive modeling. For rivers and places with anticipated intensive rains or snowfalls, the flood models take into account soil-moisture conditions, ground temperature, snowpack, topography, vegetation cover, and impermeable land area, and, combining this information with the weather forecast, compute the extent and duration of floods.

Public Transportation

In the previous section we said the modal shift towards public transportation can be enabled by informing the public about the competitive alternatives. Now, the main question is: how to make public transit an attractive option? To become a service of convenience out of the service of necessity, the following criteria must be met:

- Provide sufficient reachability. That is, the desired destination should be reachable within reasonable time from the moment one wants to start their journey.
- Be accessible. It should be within reach, easy to use and affordable.
- Be safe. That includes public health concerns, such as the spread of COVID-19, road safety and crime prevention.
- Be clean, comfortable, connected to the internet and friendly towards passengers.

Transit Performance Assessment

Improving public transit starts with assessment of its performance. Presently, most transit agencies consistently report their schedules through GTFS. That alone, allows to answer the following questions:

1. On the area, agency and regional level, what percentage of vehicles is on time?
2. What percentage of vehicles provides regular location updates?
3. How many vehicles are off course at a given time?

More elaborate questions can be answered for agencies providing data about their bus locations via GTFS-RT:

4. Is there a regular service within, e.g. a quarter of given parcels X, Y, Z? Or, what service is there within a quarter of a mile from parcels X, Y, Z? And, what is the service frequency?
5. How regular is the service near given communities? What is the frequency of service in these areas? Specifically,
    a. What is the service reliability?
    b. What are the frequencies derived both from schedule and real time, and what is the deviation between the two?
    c. Does this service provide direct connectivity to the most popular destinations, or are transfers required?
    d. What is the reliability of timed transfers between communities of concern and popular destinations?
6. How many buses are within a given distance of a specified area and what are their locations at current time? This may be useful for emergency response and evacuation.

Combining automatic vehicle location (AVL) and automatic passenger count (APC) data produces ridership on individual route segments. With the ridership data, the following additional questions can be answered:

7. What is the transit ridership on given routes and between specified origins and destinations?
8. Are buses on given routes too crowded to safely accept passengers?
9. What is the passenger-miles traveled (PMT) on given routes or within given areas?
10. What is the bus utilization over different time periods?
11. What is the cost of transit route operation relative to trips and PMT?
12. What is the PMT in given areas? What are the most/least utilized routes starting/ending in these areas?

13. What is the correlation of PMT and VMT on neighboring freeways? This can help estimate the modal split.

There is a need for a Transit Performance Dashboard, where these questions could be explored for multiple transit agencies. Presently existing transit dashboards include:

- MBTA On-Track – Massachusetts Bay Transportation Authority's dashboard for reliability, ridership, financials and customer satisfaction;[79]
- Texas Transit Performance Dashboard – covers 27 urban and rural transit districts;[80]
- CapMetro Dashboard – Capital Metro, an Austin Public Transit dashboard for reliability, safety, ridership, financials and customer satisfaction;[81]
- NYMTA Performance Metrics – New York Metropolitan Transportation Authority publishes metrics of public transit performance;[82]
- SFMTA - COVID – San Francisco Municipal Transportation Agency's data sheet about transit ridership and revenue during COVID-19;[83]
- TriMet Dashboard – Tri-County Metropolitan Transportation District of Oregon dashboard for efficiency, safety, ridership and budget.[84]

Out of these, only the Texas Transit Performance Dashboard covers multiple agencies.

Traditionally, ridership has been the main performance metric for transit agencies. Money is at stake in addition to pride: ridership partly determines the amount of federal funding allocated to each system. The national transit industry scrutinizes passenger numbers too, celebrating a national uptick and bemoaning a downturn.

But the coronavirus pandemic has shown that ridership is an imperfect measure of transit's importance. With offices shuttered and agencies themselves advising passengers to avoid all but essential trips, passenger counts collapsed by as much as 90% on rail service in San Francisco. Ridership on buses has slumped as well. And these numbers are not coming back quickly to pre-COVID levels, as work-from-home habits forged during the pandemic stand to leave lasting marks on commuting patterns in major cities.

---

[79] MASSDOT.gov. https://www.mbtabackontrack.com
[80] Texas Transit Performance Dashboard. https://www.texastransitdashboard.com
[81] Capital Metro Dashboard. https://www.capmetro.org/dashboard
[82] New York Metropolitan Transportation Authority Metrics. https://new.mta.info/transparency/metrics
[83] San Francisco Municipal Transportation Agency's data sheet. https://www.sfmta.com/covid-19-dashboard
[84] Tri-County Metropolitan Transportation District of Oregon dashboard. https://trimet.org/about/dashboard

Few people would argue that transit systems' societal value has fallen as steeply as their ridership. On the contrary, the pandemic has reminded us that cities cease to function if public transportation is unavailable. You might be able to drive to a hospital or grocery store, but that won't do you much good if transit-reliant workers can't get there to staff it.

There exists a better measure of transit success - **reachability**. The basic method of calculating reachability is intuitive: Pick a particular neighborhood, and then determine the time it would take for a person living there to use transit to reach jobs dispersed throughout the region (people take many other kinds of trips, but commute data is most readily available). Transit planners can then aggregate the data to calculate an overall regional measure of access, or they can create one for a subset of the population (say, minorities or low-income residents). Agencies are then able to use these calculations when designing a systemwide service map or when making a narrower decision, like choosing whether to open a new route or add frequency to an existing one. An agency can continue to monitor the accessibility of its service, defined as the share of jobs reachable for a given population within a given trip time, such as 60 minutes door-to-door.

It is important to note that this reachability analysis can only be performed on a system (regional) level - not per individual agency.

Mobility as a Service (MaaS)

The concept of Mobility as a Service (MaaS) has forever revolutionized the transportation industry. It is also changing the role of the transit providers in the mobility space - their responsibility has expanded to incorporate new and innovative approaches, integrating multiple modes of transportation and introducing new technology tools to keep communities moving. Additionally, MaaS changes the way riders and agencies interact, allowing for passengers to select among ride-sharing, bike-sharing, public transit services, and private transit services to fully customize their travel preferences and choose the solution that fits their needs through an integrated network. Transit agencies share the responsibility to facilitate that evolution.

The American Public Transportation Association (APTA)[85] recently undertook a study mission looking into MaaS systems in Europe to gather data and evaluate the readiness of American transit systems to fully embrace MaaS.[86] The study emphasized the importance of a solid, sustainable public transit system as the backbone of the MaaS

[85] American Public Transportation Association. https://www.apta.com

[86] APTA Report. Being Mobility-as-a-Service (MaaS) Ready. June 2019. https://www.apta.com/wp-content/uploads/MaaS_European_Study_Mission-Final-Report_10-2019.pdf

concept. According to the study, a clear vision for enhancing mobility, a high quality, well-integrated public transportation system and mobility partners/contractors who share the agency's mission, are all vital components of effective MaaS initiatives.

The study also identified that regulatory policy and lack of governmental support, not technology, is the key challenge when it comes to MaaS. It is clear that technology is outpacing policy developments with many options currently available for real-time mobility apps that integrate a community's transportation options.

One of the changes expected in the near future includes the regionality of mobility. As cities grow and begin to overlap, the traditional boundaries of public transportation are blurred. One can imagine a future where a city or metroplex releases a Request for Proposals for turnkey mobility solutions for all modes, looking for a partner to handle all aspects of mobility. Clients will look for the best overall solution for their community's complete transportation needs - covering fixed route, paratransit, university and private shuttle, on-demand, school bus, contact center, rail, fare payment, maintenance services and more. Additionally, clients will look for a provider who can aggregate all services together on one platform, through one payment channel, to create seamless journeys powered by artificial intelligence and data, while also providing additional revenue streams.

Mobility on Demand (MoD)

Mobility on Demand (MoD) is one of the more popular approaches to modal shift, improving system flexibility, efficiency and convenience. On-demand options, which take many forms, are successfully incorporated into both fixed route and paratransit services, sometimes including the services of Transportation Network Companies (TNCs) and other third-party providers. Some of our first on-demand transit operations included late night services at colleges and universities, designed to ensure students have a safe ride home after traditional transportation services end for the evening. On-demand services have also been successful as a first-mile last-mile solution to increase ridership in fixed route operations and have also shown to improve passenger satisfaction and efficiency at paratransit operations, facilitating the ability to take same-day trip requests. More recently, on-demand services have expanded again with service for passengers needing wheelchairs or other mobility devices.

One specific type of on-demand transportation - microtransit - has proven to better connect various communities and/or zones to increase the use of public transit. Information technology is a vital component of microtransit operations, and there is no shortage of passenger apps available for this purpose. Through these apps, or via a designated phone number, passengers can request trips. Passengers and dispatchers can see the real-time location of their vehicle, making commutes and trip planning much easier and convenient on the part of the passenger.

Micro-Mobility

An important component of MaaS that transit agencies must embrace moving forward are micro-mobility solutions, including non-traditional technology modes such as scooters and bike sharing. These micro-mobility services can operate as first-mile, last-mile solutions for public transit systems and, to be successful, must reflect the mobility preferences of the community. For example, successful micro-mobility services could include bike sharing in a downtown environment where biking is popular, or incorporation of transportation in open-air golf carts to fixed route pick up points in a beach community.

According to the Mobility as a Service Alliance,[87] a public-private partnership dedicated to creating the foundation for a common approach to MaaS, micro-mobility solutions can help fill in the gaps in public transportation for commuters, making public transit a more attractive option for many. These solutions can also have tangible benefits for transit agencies, including reducing parking congestion and improving ridership on fixed route systems.

In many cities, commuters are able to rent bikes, scooters and other micro-mobility vehicles on a convenient app. In some cases, customers can use their mobile phones to locate a vehicle, unlock it, and pay their fee. This level of convenience is designed to appeal to today's commuter and offers services available on-demand, at any time the passenger is ready.

As agencies look to provide more complete mobility management services to the public, the future will likely mean more subcontracts with micro-mobility providers like scooter and bike sharing companies, taking accountability for their performance and adherence with applicable regulations.

Multi-Modal Connectivity and Dynamic Transit

Transit agencies in suburban regions in California and in the United States face enormous challenges for delivering cost effective transit operations, due to the fact that traveling populations are distributed in large geographic areas. Operators typically assign a limited number of buses on the maximum number of routes possible in order to offer wide geographic coverage. As a result, the headway of each bus route is long and many passengers need transfers among such routes in order to reach their destinations. Consequently, travel using transit takes much longer than the driving

---

[87] Mobility as a Service Alliance.eu. https://maas-alliance.eu

alternative. The longer travel time significantly discourages choices to take transit. Rural and suburban transit agencies need tools to be able to innovate transit operations.

The conventional transit system was established upon the framework of fixed routes, fixed schedules and almost fixed ticket prices. To improve the transport service in order to attract travelers from private driving, transit systems are substantially required to be flexible and dynamic. The concept of Dynamic Transit Operation can be interpreted as flexible operations, multi-mode linkages, traveler-centric, and demand-responsive.

Integrated Dynamic Transit Operation (IDTO) applications are expected to demonstrate great potentials of improving transit service. The IDTO algorithms and user apps are developed to allow holding at bus stops to meet with connecting passengers (Connection Protection or T-CONNECT), to facilitate first-mile and last-mile shared riders (Dynamic Rideshare or D-RIDE), to provide more accurate and reliable ETA/ETDs in real-time for entire-trip (RT-ETAD), to offer incentives for encouraging frequent and accumulative ridership (T-REWARD). Therefore, IDTO can offer significant reduction of travel time and connectivity improvement for a greater number of travelers, change travel demand which in turn helps congestion relief, run a more cost-effective operation, and create positive changes in inter-organizational cooperation. The implementation of IDTO may potentially change the way in which transit services are operated in suburban American cities.

Research efforts also aim to significantly improve the reliability of intramodal and intermodal connections through the integration with MaaS and MoD platforms and first/last mile mobility solutions. The function of Dynamic Connection has been designed as an add-on function to any trip planning/navigation application (such as HaCon, Google Transit, etc.) as long as the trip requires a real-time connection between successive segments.

Passenger Rail

There are three main benefits of rail transportation:

- **Mass transportation**. Rail is the only transportation mode capable of moving a massive number of people reliably at high speed. This unique capability, which is essential to alleviating a big surge in demand over a short period, makes rail a critical component of MaaS.
- **Energy efficiency**. Rail is three times more efficient compared to cars when it comes to per-capita resource consumption. It helps reduce road congestion and greenhouse emissions by taking vehicles off the road. The industry has made strides to adopt more efficient engines, cleaner fuels and electrification

to further reduce the carbon footprint. By choosing cutting edge technologies such as regenerative braking, which helps to capture power from braking, the industry continues to drive higher efficiencies.

- **Economic growth**. Studies show that rail systems generate two to four dollars equivalent economic growth per every dollar invested. The construction and extension of rail systems stimulate the local economy and provide a source of working-class jobs. People can work, study, and seek entertainment in large urban areas while living in more affordable houses outside the city, knowing that they can rely on rail transportation day in and day out.

Economic and cultural factors drive the development of rail networks in each region. For instance, population density determines the type of service and the scale required, and policy often influences the direction and pace of innovation. By studying the characteristics of a country one can understand why rail systems between countries may not be comparable in volume, modes or passenger experience. Aging infrastructure and a capital-intensive operation are top challenges in the rail industry, but particularly in the U.S. a significant car culture, traditional funding practices and a lack of customer focus have been critical reasons why our passenger rail system lags compared to other systems worldwide.

Paying for the Ride

COVID-19 triggered transit agencies all over the world to adopt contactless payment technologies. For example, The Greater Dayton Regional Transit Authority (RTA)[88] and app-developer Transit[89] cooperated to allow riders purchase reusable smartcards to tap and ride without needing to buy a ticket, or select a fare before traveling. The payment platform is being rolled out in partnership with Masabi[90], a provider of mobile ticketing technology for public transport.[91]

Apps from Transit and the regional transit EZfare consortium[92] serve 13 transit agencies across Ohio, Michigan and Kentucky, allowing riders to purchase tickets through their phone and access real-time trip planning. Other cities like New York, Washington D.C., Chicago and Los Angeles have also rolled out tap-to-pay

---

[88] The Greater Dayton Regional Transit Authority.org. https://www.i-riderta.org
[89] Transotapp.com. https://transitapp.com
[90]Masabi.com. https://www.masabi.com
[91] Greater Dayton Regional Transit Authority, Masabi and Transit app Launch Account-Based Ticketing for Tapp Pay on Mobile. Press Release, Feb. 2021. https://www.i-riderta.org/news/greater-dayton-regional-transit-authority-masabi-and-transit-app-launch-account-based-ticketing-for-tapp-pay-on-mobile
[92] EZfare consortium. https://www.ezfare.us

functionality for their transit systems to provide quicker, easier, and safer payment methods for riders and commuters.[93]

In Moscow, Russia, the contactless initiatives went further. In October 2021, Moscow subway system launched a Face Pay system - when a rider's credit card is charged as a result of facial recognition upon their entering the subway.[94]

Equity

Transportation equity means accessible and affordable transportation for everyone in the community resulting in fair distribution of transportation resources, benefits, costs, programs and services based upon differences in income, ability and other factors affecting transportation choice and impact. In that respect, public transportation and mobility services for elderly and disabled are the pillars of equity. Public transportation was discussed in the previous section. We can add that the reachability and accessibility analysis, described in the section Transit Performance Assessment, should be combined with the Census data on income, number of cars per household, housing cost, education, unemployment, and the land use data characterizing job-rich areas. In this section we will focus on mobility options for the elderly and disabled.

Possible Future of Paratransit

Private companies that are managing health care are finding that the lack of access to transportation is costing them big money; therefore, they are looking for ways to improve this situation. Some are even investing substantial funding into coordinated transportation partnerships because the cost of fulfilling an evening trip to take a patient home from the hospital is a fraction of the cost of an additional night's stay. Additionally, the cost of missed appointments alone is in the billions of dollars, again prompting health care organizations to take more than a passing interest in how medically related transportation services are provided. They are beginning to seek out public transit as a partner to remedy this situation, and slowly but surely, we are beginning to see integration of paratransit with non-emergency medical transportation (NEMT) resources.

In Richmond, Virginia, the Greater Richmond Transit Company (GRTC)[95] has established one such arrangement in which a private-sector broker, RoundTrip,[96] was

---

93 Transit payment systems revamped for post-COVID commuters. PaymentsDive. Apr. 2021. https://www.paymentsdive.com/news/transit-payment-systems-coronavirus-pandemic/598827

94 The Guardian. Privacy fears as Moscow metro rolls out facial recognition pay system. Oct. 2021. https://www.theguardian.com/world/2021/oct/15/privacy-fears-moscow-metro-rolls-out-facial-recognition-pay-system

95 Greater Richmond Transit Company.com. http://ridegrtc.com

96 RoundTrip.com. https://roundtriphealth.com

hired to integrate health care related transportation with ADA[97] paratransit. RoundTrip enables transportation for medical appointments to be arranged through third parties, which include health care organizations. RoundTrip in turn distributes the ensuing trips to various fleets which are inclusive of vehicles owned by the transit agency and others that are not. These are the seeds of coordination being planted, and the future of paratransit beckons us to continue in this direction to include referrals to services that are operated separately from the transit agency altogether. Meanwhile, GRTC built a highly successful bus rapid transit service, Pulse,[98] which has achieved national recognition for generating ridership increases as other agencies struggle to retain ridership.

MoD for Disabled Riders

MoD with wheelchair-accessible vehicles (WAVs) not only increases mobility options for individuals with disabilities, but also allows for better overall service delivery by freeing up resources and increasing productivity.

The expertise of transit providers assists clients in identifying areas that can benefit from microtransit services, design parameters of service delivery, choose appropriate technology tools, and develop successful marketing of new initiatives to increase public awareness.

In one successful example in Redding, California, First Transit[99] and Far Northern Regional Center (FNRC),[100] a nonprofit organization, worked together to improve access to jobs for individuals with disabilities. In 2018, First Transit began providing an on-demand program, Redi-Ride, for FNRC clients needing transportation to and from work or employment-related functions. As job opportunities increased for the community, the need for safe and reliable on-demand transportation quickly arose. The program now provides 600 rides per month, opening up opportunities for FNRC clients that did not previously exist.

Digital Companion for Disabled Riders

Spatial mismatch is the mismatch between where low-income households reside and suitable job opportunities. Disabled people who can and want to work are dramatically affected by the spatial mismatch. Generally of lower income, these people have to compromise on choosing where to work or find an appropriate, and sustainable for them, means of transportation. For most of them, public transit, and occasionally paratransit, is the only way to get to and from work. The main pain points of the disabled

[97] Americans with Disabilities Act: https://www.ada.gov
[98] GRTC Pulse: http://ridegrtc.com/brt
[99] FirstTransit.com. https://www.firsttransit.com
[100] FarNorthRegionalCenter.com. https://www.farnorthernrc.org

travelers include the inability to get a door-to-door ride, long waiting and travel times, risk of ending up stranded away from home due to changing services or lack of after-hours service, and COVID-19 related safety concerns due to rising crime and rioting.

One way to help is to provide an app that would serve as a personalized digital companion to disabled travelers providing information about mobility options in trip planning, given travelers' circumstances. The goal is to build a working prototype system that would help generate a travel plan for a disabled person using available mobility options that include transit and paratransit but is not limited to those. In addition to the prototype, we will produce a concept design for the product that will describe how travelers and mobility services are added to the system and a path to deployment.

The core technology will be building a Knowledge Graph (KG) of disabled travelers with their needs, restrictions, preferences and points of interest (POIs); and mobility services including transit, paratransit and private companies such as Uber WAV with their schedules, coverage areas, cost, etc. As the KG grows, new relationships between existing entities may be discovered. A knowledge update engine will be a distinguishing feature of the system. A review-like function will be the core feature of the KG built on the semi-automatic collection of feedback, reviews and surveys. This will update user preferences. The information contained in the KG will be accessible through an API by a functional prototype user interface (UI).

Transportation Accessibility and Adoption of Shared Automated Vehicles (SAV) by Disabled Travelers

Automated vehicles (AVs) are one of the most significant technological advances to happen to our transportation systems - they are expected to save lives and provide enhanced mobility for all people. AV technologies also allow the emergence of novel business models such as shared automated vehicles (SAVs), which could provide more accessible on-demand mobility services and play a vital role in sustainable transportation. In particular, SAVs could be an attractive means to provide transportation access for those who cannot drive or face significant barriers to driving, including the elderly and people with disabilities.

Individuals with disabilities are a unique but sizable demographic in the United States. According to the 2018 disability status report, nearly 13% of the population has a disability. Unfortunately, many people with disabilities have limited transportation opportunities and adverse transportation experiences. We know that AVs have the potential to transform personal mobility. Whether disabled travelers can benefit from this transformation depends on to what extent the accessibility is taken into consideration in the development and planning process of the deployment.

Research is needed to understand the transportation accessibility for people with disabilities, their future adoption of SAVs, and the potential change to their transportation mode choices. There is a knowledge gap in the literature that prevents us from understanding the specific transportation needs, preferences, and accessibility issues for these disadvantaged groups who could benefit from AV technologies' deployment. Therefore, to better understand this population's transportation needs and unique challenges, firstly, we need to carry out a thorough analysis of disabled travelers' transportation experience (particularly while using public transit and ride-hailing services). Research on user acceptance and interest in SAVs by the disabled population is also limited, so it is unknown what the key determinants would be for them to use SAVs. Therefore, we also need to investigate the technology acceptance of SAVs and what factors (e.g., safety, compatibility, trust, accessibility, and ease-to-use) will contribute to the intentions of using SAVs by disabled travelers.

Parking and Curb Management

Urban congestion is largely a result of parking shortage. Drivers arrive at their destination and then start circling around blocks in search of street parking - this adds to VMT, emissions, delay, and safety risks, as drivers searching for parking are distracted. Another big problem facing cities is double parking by delivery vehicles and Transportation Network Companies (TNCs) such as Uber and Lyft. Trying to solve these problems, cities embark on experiments that lead to a transformation in urban curb management. Curb space is increasingly becoming a dynamic interface between people and vehicles, bikeways, bus lanes, street vendors (e.g., food trucks), and paratransit picking up and dropping off passengers with disabilities.

How cities manage this public asset and resource for local economies is a timely problem, whose solution relies on first understanding the emerging usage patterns over time and space. Currently, there is no systematic method for identifying these patterns and there is great untapped potential for doing so with emerging machine learning technologies and nontraditional data sources, such as dashboard cameras mounted on shuttle buses, parking enforcement and other vehicles that routinely travel a geographic area of interest. When combined with stationary cameras, these provide an excellent data source covering broad ranges of time and space. Recent advances in deep learning provide an opportunity to learn usage patterns from such data.

Smart Parking

Smart parking is a technological approach to improve the parking process and the cars' positioning in a city with a shortage of space. Its main purpose is to spare time for drivers who want to leave their cars close to their destinations. It effectively reduces the number of vehicles in the streets, which in turn improves the environment.

Any smart parking initiative implies the use of additional smart devices starting from a regular smartphone to special sensors and cameras in the parking areas. Also, AI takes its place among smart parking solutions. In other words, smart parking is a highly diverse branch using the latest tech advancements. Smart parking technologies include:

- IoT smart parking sensors. The sensor is positioned in a parking space, underground. Before the work starts, the sensor should be properly calibrated. The vehicle located in a parking space changes the earth's magnetic field, and the sensor tracks this change. The occupation status is then sent to the cloud.

  Users can track the sensors' signals in real-time from a PC or smartphone. First and foremost, drivers see all parking areas around them along with free spots they can take - it looks like an interactive map. The interface of such a tool is usually simple and user-friendly. Quick navigation is essential when driving: no extra taps on the smartphone screen - only minimalistic and understandable navigation symbols (big icons, vivid arrows, etc.). Smart cities parking usually involves third-party integrations. For instance, an app developed for a certain company should be easily integrated with the devices of the city's authorities.

- The counter system is another smart device that detects when a car appears in the parking lot and when it leaves. This data is then processed by the IoT device, and it can identify free parking spots at any given moment. The data from the counters is stored in the cloud, so parking managers can use it later for analysis and demand predictions.

- Control systems have several features that make them more complex than sensors or counter systems. A control system detects if someone breaks parking rules, then it registers the violation and stores the evidence (a photo or a video), plus, it can issue a ticket and inform the violator. This way, smart parking solutions exclude the human component and contribute a lot to reducing the city's expenses and increasing safety as they often have to work tirelessly 24/7.

- Parking lifts are aimed to optimize parking spaces by increasing the capacity of the parking lot. Lifts help optimize even the tiniest spaces by exploiting vertical space, allowing the placement of up to three cars on the same parking spot.

- Parking robots serve as valets that take your car and place it with maximum efficiency. Robotic car parking is complex and involves several components. It works similarly to AI robots but involves lifts, shelves, and other similar tools. In other words, this is a robotic system that simplifies and organizes the parking process.

These solutions require significant capital spending. The return on investment should come from parking fees. To maximize it, flexible pricing can be used:

- The pollution level. High pollution is an enormous challenge for lots of big cities, so authorities opt for modern technologies to encourage people to drive "greener" cars. Pollution-based fees is another way smart parking technology helps us to take care of the planet. The system checks the car's year and engine, then calculates parking fees based on this information. As a result, the owners of hybrid and diesel vehicles pay a 20% lower price, while for electric cars it is completely free. This should encourage citizens to drive eco-friendly cars in the city area.
- The car's length. Do you prefer bigger cars? You should be ready to pay more for their parking. This is a democratic way to support the owners of smaller cars, whose vehicles take less space. In other words, you pay not only for the time but also for the space occupied. And, of course, smaller cars are another eco-friendly option.
- The driver's citizenship. If you live in this city, you pay taxes for its maintenance, including parking facilities, so citizens do not have to pay the same parking fees as tourists. Governors can create "residential parking" areas throughout the city to support taxpayers and help them in a constant battle for a parking spot with tourists.

Curb Management

In Las Vegas, Cox Communications[101] has entered a pilot partnership with the city to introduce its Cox2M curb management system to help manage downtown congestion. Curbside kiosks collect video data, making vehicle type and license plate information to kick off a countdown clock alerting drivers when they've spent too much time on a curb.[102]

Another project in Aspen, Colorado, established "smart zones" where delivery drivers use an app to book small delivery reservations. The technology, provided by Coord,[103] has made for a smoother orchestration of goods coming into the city's small downtown. In Aspen, 11 popular loading zones were converted to the new smart zones, which the

[101] Cox Communications. https://www.cox.com

[102] City of Las Vegas Launches Smart Curbside Management Corridor with Cox. Nevada Business, Mar. 2021. https://www.nevadabusiness.com/2021/03/city-of-las-vegas-launches-smart-curbside-management-corridor-with-cox

[103] Coord.com. Curb management for fast-changing cities. https://www.coord.com

city charges $2 per hour to use. Nearly 120 delivery drivers have signed up to use the zones. The move has ostensibly reduced congestion and double-parking.[104]

Other cities like Columbus, Ohio and Omaha, Nebraska, have also embarked upon curb management initiatives that use technology.

Logistics

Transportation is an important function in logistics industry operation, and its automation plays a significant role in the technological transformation of logistics. In this section, we discuss truck monitoring, vehicle automation and last mile delivery.

Monitoring of Truck Movements within the State

Within California, traditional truck data sources include Weigh-In-Motion and Traffic Census stations operated by Caltrans, and truck scales operated by the California Highway Patrol. While these sources provide valuable information about the physical characteristics of passing trucks, they provide limited data about their origins, destinations, and route choices between the existing set of fixed observation points. This limitation particularly comes into play when trying to determine truck movements in urban areas or in rural areas outside the heavily intercity or interstate traveled corridors. This data gap may potentially be closed by using information from data aggregators relying on fleet tracking, such as Streetlight Data[105], INRIX[106], Replica[107] and others. However, while the information provided by these data aggregators has already proven valuable at determining typical route choices, there is still significant uncertainty on the quality of the volume estimates provided. This is because volume estimates are based on samples of trucks that may vary from location to location, and time to time. There may also be an overrepresentation or underrepresentation of some truck classes. Research is needed to assess the quality of these emerging data sources and the ability of using them to supplement current data sources. There is also a need to develop guidelines on the number and placement of fixed data points needed to ensure that the collected sample data adequately represent desired truck movements.

---

104 Coord.com. Smart Zones Deliver: Full Aspen Case Study. Coord Blog, Mar. 2021. https://www.coord.com/blog/aspen-smart-zone-case-study

105 Street Light Data. https://www.streetlightdata.com

106 INRIX. https://inrix.com

107 Replica. https://replicahq.com

Automated Trucks

Level-4 automated Class-8 trucks are powered by self-driving systems for trucks from Aurora[108], Plus AI[109], Waymo[110] and Kodiak Robotics[111]. AV technology is fundamentally agnostic when it comes to the powertrain of the vehicle, so these companies test with traditional diesel trucks on the road today. As the industry moves toward electric trucks, these technology providers intend to be part of the transition; however, they don't plan to wait around for large scale charging infrastructure or to hold off on the other benefits that can be realized from autonomous technology on any type of powertrain. They believe that once electrification will envelop the trucking industry, it would be easy to adapt the AV technology to the electric powertrain.

PATH has been working on development of a CACC system for heavy trucks to enable platooning in collaboration with the Volvo Group since 2015, under the sponsorship of the FHWA and Caltrans. Platooning is supposed to reduce fuel consumption, increase truck speed and throughput. Operation of truck platoons was demonstrated in Richmond, San Jose, Los Angeles - in California; Centreville, VA; and in Blainville, Quebec, Canada.[112]

Last Mile Delivery

Last-mile delivery refers to the stage when a carrier picks up orders from a distribution hub or warehouse and delivers it to the final destination. The goal of last-mile delivery is to deliver orders to customers quickly and effectively. Faster delivery is no longer a thing of the past but expected from your customers due to the standards set by names like Amazon and Walmart Marketplace.

To ensure customer satisfaction, it is important to make sure that your company has a handle on its delivery process. The sooner a customer receives their package, the happier they will be. Cost is also a big factor. An estimated 28% of an online brand's bottom line comes from last-mile delivery costs. To offer customers incentives to shop from you over competitors, you will need a shipping strategy that enables 2-day shipping or even free shipping. The key is to find ways to optimize last-mile delivery so you can offset shipping costs.

---

108 Aurora. https://aurora.tech
109 Plus AI. https://plus.ai
110 Waymo. https://waymo.com
111 Kodiak Robotics. https://kodiak.ai
112 Shladover, S. et al. Partial Automation for Truck Platooning. The Exploratory Advanced Research Program - Research Summary Report FHWA-HRT-19-02. Nov. 2019. https://www.fhwa.dot.gov/publications/research/ear/19028/19028.pdf

There are numerous startups, e.g., Starship[113], Nuro[114], Udelv[115], Kiwibot[116], Eliport[117], TeleRetail[118], Postmates[119], Robby[120], BoxBot[121], as well as established companies, such as Amazon[122], Yandex[123], Continental[124], that produce robots for last mile delivery. These robots are battery-powered IoT devices with built-in computer vision capabilities, precise positioning, and communication systems that can support control by a remote operator.

Emergency Response and Evacuation

Emergencies happen often, but scenarios never repeat themselves. Therefore, effective emergency response requires real-time data collection, separating signal from noise and decision support - what orders to issue, to whom and when; and what outcome to expect. This decision support must be based on predictive models.

When an evacuation is necessary, residents are asked to leave their homes in an orderly fashion maintaining a priority sequence defined by the district authorities. The major pain point is that the reaction of local residents to the recommendations of district authorities is unknown. On the one end of the reaction scale there is rush and panic, whereas on the other end there is olympian delusion and procrastination. Disregard for recommendations leads to unexpected traffic bottlenecks and necessitates planning adjustments by the authorities. For example, during a wildfire evacuation, the goal is to get people from the endangered areas to the freeway as soon as possible, which may turn impossible due to road blockages. The solution would be to redirect a certain portion of traffic to safe refuge areas of limited capacity (e.g., shopping mall parking lots) so as to unblock the road for the others. Timely reaction to traffic events is crucial to avoid injuries and/or casualties.

---

[113] Starship. https://www.starship.xyz
[114] Nuro. https://www.nuro.ai
[115] Udelv. https://www.udelv.com
[116] Kiwibot. https://www.kiwibot.com
[117] Eliport. https://eliport.com
[118] Teleretail. https://teleretail.com
[119] Postmates. https://serve.postmates.com
[120] Robby. https://robby.io
[121] Boxbot. https://www.boxbot.io
[122] Amazon.com. Meet Scout. Jan. 2019. https://www.aboutamazon.com/news/transportation/meet-scout
[123] Yandex.com. https://sdg.yandex.com
[124] ContintentalAG.com. Package Robot for the Last Mile. https://www.continental.com/en/products-and-innovation/innovation/our-vision-for-tomorrow/package-robot-for-the-last-mile

According to the comprehensive literature survey on disaster response,[125] several factors are consistent predictors of evacuation behavior. Gender, age, and ethnicity are demographic factors associated with response to recommendations. Although traffic incidents are common during evacuations, their effect is in general neither factored into traffic forecast models nor accounted for in evacuation planning.[126]

To help the districts in their assessment of the traffic situation during the emergencies and their decision making, one should use a tool chain depicted in **Figure 4**. Here, the hazard model (e.g., a model of fire propagation), and the household information are the exogenous inputs to the system. A control input is the response, a set of informed actions initiated by the district authorities that include recommendations to residents with respect to sheltering in place or leaving, which route to take to what destination, as well as active traffic management through road closures and/or opening extra lanes in a given direction.

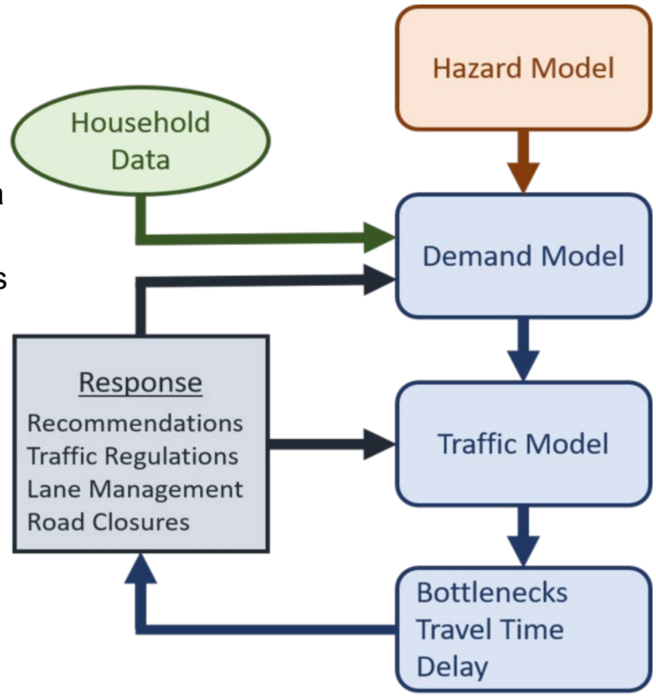

Figure 4. Decision support workflow in emergencies.

Population games[127] can be incorporated into the demand and traffic models. The demand model computes how vehicle flows sent into the road network by given residential areas are distributed in time. The compliance with recommendations is modeled by a population game. In the traffic model, population games determine destination and route choices based on traffic conditions. Ultimately, the traffic model predicts bottleneck locations and severity together with expected travel times and delays.

---

125 Thompson, R., Garfin, D. and Silver, R. Evacuation from Natural Disasters: A Systematic Review of the Literature. Risk Analysis. 2017. https://cpb-us-e2.wpmucdn.com/faculty.sites.uci.edu/dist/2/870/files/2020/01/Thompson-Garfin-Silver-Risk-Analysis-2016-ONLINE-FIRST.pdf

126 Robinson, R. et al. Role of Traffic Incidents in Hampton Roads Hurricane Evacuations. 88th Annual Meeting of the Transportation Research Board. Jan. 2009. https://trid.trb.org/view/881133

127 Springer. Population Games. Game Theory: Decisions, Interaction and Evolution. 2007. https://doi.org/10.1007/978-1-84628-636-0_8

The proposed modeling tool chain will be utilized in offline and online modes. In the offline mode, an ensemble of simulations will be run for a variety of population parameters to determine the spectrum of outcomes and to identify worst cases. In the online mode, population parameters will be adjusted in real time based on vehicle counts at key junctions, and simulation models will be constantly re-run for the best prediction of the outcome.