Berkeley Institute of Transportation Studies

The Transforming Transportation Ecosystem — A Call to Action

Jane Macfarlane, Institute of Transportation Studies, University of California, Berkeley

November, 2019

UCB-ITS-RR-2019-01

DOI:10.7922/G28C9THN

Berkeley UNIVERSITY OF CALIFORNIA

Contact Information: Jane Macfarlane, UC Berkeley Institute for Transportation Studies.
janemacfarlane@berkeley.edu

The Transforming Transportation Ecosystem—A Call to Action

The transportation landscape is in transition. Rising congestion, failing infrastructure, changing behaviors, adapting to a more inclusive definition of mobility, the desire for cleaner and more efficient engines, and grappling with the role of autonomous vehicles and drones, to name just some of the factors, demands that we take a fresh approach to designing for mobility. Yet the rapid pace of technology development is creating emerging trends that are driving change faster than our ability to model, design, and manage them. This could potentially result in undesirable economic, environmental, and societal outcomes. The speed in which technology is remaking transportation and introducing new business models is leaving policy makers and government systems at a disadvantage, and the data needed to frame policy and new social infrastructure is becoming increasingly privatized.

This white paper shares the expertise and collective wisdom of leading researchers and practitioners who are engaged in the development of next-generation mobility systems and the built environment. It summarizes the presentations and discussions conducted in a workshop in May 2017 in which participants addressed how to bridge the gap of the data and models needed to adequately and intelligently design infrastructure and systems for a cleaner, safer transportation network while expanding and reinventing the notion of mobility.

A common reflection among all participants is the need for urgency. The proverbial train has left the station, but who will be the conductors and who will decide the route we take? We must begin now and bring together the many sectors that contribute to our transportation network and infrastructure.

Our goal must be to guide these innovations to a social optimum, rather than let technologies drive us to respond with a patchwork of policies to address unintentional consequences.

We need data to drive our theory, policy, and models. Today, with the world's pervasive use of mobile devices and the Internet of Things that can track our interests, choices, and locations, private companies have a wealth of data that can help us infer behaviors, inform and drive our models, and add significant insight into mobility demands. Yet government entities have limited access to this data.

In today's information economy, can we create private-public partnerships to share data to realize a safe, sustainable, and equitable transportation system?

Achieving a Maximum-Mobility, Minimum-Energy Future

After electric power usage, transportation is the second-leading sector in energy consumption. Seventy percent of petroleum consumed in the United States is for transportation, of which 85 percent is for on-road vehicles. For many households, transportation costs are the second-highest budget item.

Previous transportation models and metrics have favored the use of individual vehicles, leading to urban decentralization and sprawl and, with the internal combustion engine being the primary mode of transport, energy inefficiencies and environmental impact. Congestion is stressing our infrastructure, slowing economic growth, and requiring fundamental changes in how we address mobility in our cities.

Mounting congestion and lack of sufficient public transit is leading people to embrace emerging mobility solutions, such as dynamic routing and on-demand driving services. Without policy and direction, these changes to mobility habits will lead to increased energy consumption and wear to the infrastructure.

To address the complexity of our transportation transformation, the DOE Energy Efficiency and Renewable Energy office has initiated these projects.

- Smart Mobility Consortium—A five-lab effort focused on driving transportation solutions toward a maximum-mobility, minimum-energy future
- Technologist in Cities—An interagency collaboration between DOT and DOE to support the DOT's Smart City Initiative
- Living Labs—Pilots that feed real-time data and learnings to researchers generated by field-ready solutions
- ARPA-E's NEXTCAR/TRANSNET
- Virtual Ultra-High Efficiency City—Exploring the integration of mobility, technology, decision science, and urban planning
- Research and Development—Applying big data, AI, and machine learning to mobility challenges

The Department of Energy Vehicle Technologies Office's mission is to find transformative technologies that drive significant energy improvements across the transportation sector. The transportation ecosystem is complex, and it is difficult to partition for specific aspects, such as minimizing energy use. The Energy Efficient Mobility Systems group focuses on a broad range of metrics to reduce consumer and business costs, maintain industry competitiveness, improve energy efficiency, and increase domestic energy security to create a maximum-mobility, minimum-energy future.

The DOE's recent report, *The Transforming Mobility Ecosystem: Enabling an Energy-Efficient Future,*[1] evaluates a range of possible future scenarios of how to achieve positive outcomes for the economy, safety, affordability, accessibility, and energy-efficient mobility.

[1] https://energy.gov/eere/vehicles/downloads/transforming-mobility-ecosystem-report

The report was created "to bring a much-needed focus to the range of possible impacts this transformation may have on energy, while acknowledging the economic, safety, and accessibility implications. In addition, it highlights the impacts that the mobility system of the future will have on our built environment, and how these interactions could change our cityscapes, as well as suburban and rural areas."

This transformation is being driven by a variety of trends, such as urbanization, increasing use of technology by individuals, and innovative business models facilitated through new technology.

Will This Be a Happy Marriage?

Today, we are witnessing the marriage of technology, vehicles, and the roadway infrastructure, each of which have different cultures and time frames. How will we merge rapid, agile software development with long-term infrastructure planning?

- Information technology operates in product life cycles of months. Products are generally low-capital investments, and often the customer does the beta testing.
- Consumers are increasingly engaged in applications on mobile devices that deliver transportation information, such as traffic routing, multi-modal route planning, parking reservation systems, and ride hailing and sharing.

On the other hand:

- Vehicles have life cycles of years, a high capital cost, and extensive safety testing must be conducted before release.
- Roadway infrastructure has a life cycle of decades with a very high capital cost, typically publicly funded. Safety is a key factor, and planning and construction are folded into 50-year planning activities for the region or state.

This mismatch in life cycles, costs, and planning makes integration a challenging proposition.

Future Mobility Requirements: What Assumptions Can We Make?

Planning the next-generation infrastructure and built environment intelligently requires understanding mobility needs. However, we have limited visibility into the demand side of the issue. Human behavior is the most challenging to predict.

On Automatd Vehicles

"How do we certify safe enough? Aerospace V&V represents 50% of the development costs, which is orders of magnitude simpler than on-road vehicles." – Steve Shladover

Mobility decisions are often contextual. Data captured from transit systems and road usage do not provide the causality associated with the observed patterns. Behavior is affected by past experiences, emotion, and social influences.

With new technologies, we tend to jump to the ideal situation. Many have great hopes for what autonomous vehicles could do for sustainability, congestion, safety, accessibility. The hype cycle takes control, and massive investment occurs. Certainly, automation might improve energy efficiency and safety, but given the massive congestion in many cities, it won't be enough to relieve the congestion. Automation might offer a 1.5 to 3 times capacity increase because of changes to vehicle density, but we have increasing population, urbanization, and vehicle miles traveled (VMT) per capita.

Moving 100 people with a bus or bikes requires much less space than moving them with cars. Whether using a personal vehicle, a ride service, or an autonomous car, the amount of road space occupied remains the same.

On Shared Rides

"We need to start now!" – Joan Walker

Automation could potentially add vehicles to the road if people see them as a more convenient option to public transit. Cities are grappling with this question today. San Francisco, for example, registered over 45,000 Lyft and Uber drivers. As congestion increased in the city, San Francisco demanded driver data to determine to what extent these drivers were generating congestion. On-demand delivery is escalating, which also has the potential to generate even more vehicles on the road.

The following assumptions are important to consider as we predict future vehicle usage patterns and ownership.

- **People won't own cars anymore**. This premise anticipates that it will be cheaper not to own a car. However, this assumes a rationality that people currently don't have. Most current car owners do not consider the entire cost of car ownership and do not own a car that just meets their basic mobility needs. Vehicle purchase choice involves a myriad of decision constraints, many of which are emotionally driven.
- **Convenience will drive people to use ride-hail, ride-share, and autonomous services**. Counter to this assumption, owning an autonomous vehicle would be just as easy, if not more convenient.
- **Flexibility will allow users to choose a vehicle specific to the activity they engage in**. Analysis of technology adoption and diffusion shows that when a technology is introduced, the price is high, so people share. But when the price drops, people buy a vehicle that covers their personal needs.
- **Adoption of alternative fuel vehicles will increase.** It is assumed that consumers want more efficient vehicles and that alternate fuel vehicle (AFV) use will increase when the fueling infrastructure expands. But survey results show that 92 percent of the U.S. car-buying population places miles per gallon far down their priority list when choosing a vehicle.

- **Sharing will be common**. How many people want to share cars with a stranger or take it to a transit station? Past and current behavior contradicts this assumption. Other considerations include: How will connected automated vehicles (CAV) fit into the urban environment? And how will the urban environment need to change to accommodate them? What are the implications of a heterogeneous fleet of CAVs and human-piloted vehicles?

As with any new technology, unintended consequences accompany the advancements. Zero-occupant trips, such as someone sending a car home unoccupied and then back for a pickup, could potentially double VMT and dramatically increase congestion. Regulations and pricing will protect us from these effects and move us toward changes in behavior and decision-making.

It's important to begin pricing new technologies, such as drones and zero-occupancy AVs, because pricing is more difficult after free access has been established.

Equity for the elderly, disabled, and young is another important consideration. If the widespread dissemination of AVs leads to a downward spiral for public transit, it could have a negative impact on accessibility. Private transport services do not have a good record of broadly serving all populations. However, AVs could offer new mobility options for these populations. Pricing could impact this if revenues are used to target the equity issues. It will be up to the public sector to determine how disadvantaged communities can be served in all areas and at an affordable price.

With these considerations in mind, we must find new ways to measure mobility. Are longitudinal studies the most effective method of gathering the information needed? With technology changing so rapidly, how do we implement these studies?

Who is Managing Our Traffic?—Ad Hoc Traffic Management Systems

The unbridled influence of technology is clear when it comes to congestion management. Dynamic traffic-aware routing applications available on smartphones and navigation systems are becoming ad hoc traffic management systems as consumers willingly provide private companies data to get an understanding of the instantaneous state of the road network. These private sector solutions interfere with the established public traffic management system, because traffic managers have no insight into the providers' rerouting activities.

Yet, experimental validation of rerouting scenarios has shown that congestion is not avoided. Instead, cars are pushed from highways onto local streets that are not equipped to handle the additional traffic. And, is it in our best interest to use every inch of asphalt for vehicular traffic rather than promote emission-free travel modes? Modeling indicates that dynamic routing can actually increase travel time and the energy footprint.

This bottom-up approach exemplifies a multi-player, non-cooperative gaming situation because the companies involved do not share information. The process uses a learning algorithm that improves daily by monitoring how well it performed that day. Knowledge of how any company is improving its algorithms is not available, so each model continually makes assumptions in a vacuum.

On Design Paradigms

"A lot of dystopias are made by well-intentioned engineers " Referencing how much personal information can be derived from simple utility usage and the congestion 'Frankenstein' created by independent routing engines. Roy Dong Alex Bayen

To control traffic flow and the resulting energy impact requires an understanding of the system's dynamics. Because there is no insight into the routing companies' algorithms, and the feedback loop is constantly changing due to dynamic routing, the outcomes are uncertain.

We must understand the broader ecosystem of all stakeholders and enable a collaboration between the private and public sectors if we are to control traffic patterns and the consequent energy footprint.

It is estimated that more than 80 percent of this mobility data is behind corporate firewalls, and it tends to stay behind firewalls because of "privacy concerns."

Fueling Demand for AFVs

Although alternative fuel vehicles have been available for many years, the lack of a fueling infrastructure has been thought to be the dominant inhibitor to their deployment. Policy has made some headway in changing the demand profile, such as by allowing single-occupant HOV lane usage and offering pricing incentives, but range anxiety continues to inhibit growth.

However, we can't plan an infrastructure for a fleet if its role in the mobility ecosystem is unknown. Do we wait and slowly build to the demands that emerge? What if the urban landscape dramatically changes? Would this designed infrastructure that reflects today's car-driving population atrophy? How does the value proposition and fleet energy consumption vary for different AFV technologies in the near term and mid to long term as shared mobility increases?

This information is necessary for the private sector to make informed estimates of shared mobility fleet AFV adoption, thus lowering the investment risk and increasing economic stability, energy efficiency, and affordability.

To date, travel behavior and traffic patterns have been the basis for determining ideal locations for charging stations. Would this change if the shared mobility market grows? For example, if vehicles are not consumer owned, fueling could occur at central charging locations. Companies that operate charging stations collect data on the number of charges, how long they are, and more. How will this data inform siting and power ratings for 10–15 years from now?

Transportation Network Companies: Privatizing Public Transit

Consumers have a latent demand when it comes to transportation mobility. Every time we make it faster, easier, more convenient, we use it more. To date, TNCs have not put more people into one vehicle and eased congestion. Instead, many cars are on the road that wouldn't have been there previously. The ease of requesting a ride has encouraged people to abandon other modes of transportation or even walking short distances. Yet on-demand vehicle services could have a role in supporting public transit use by addressing the first-last mile issue and reducing car ownership. Policy and pricing could promote accessibility to neighborhoods that lack access to other means of transportation.

Connected Automated Vehicles: Safety First

The environment in which CAVs must operate is complex and randomly determined. Dynamic external hazards as well as varying environmental conditions contribute to the complexity.

Our current mode of self-driving is remarkably safe on a per mile traveled basis.

The current U.S. traffic safety record sets a high bar: 3.4 million vehicle hours between fatal crashes (which equates to 390 years of nonstop 24/7 driving) and 61,400 vehicle hours between injury crashes (which equates to 7 years of 24/7 nonstop driving). Our expectations are that CAVs must at least match these levels of safety. A RAND study recommends that many more hours than these baselines need to be tested before CAVs could be proven safe.

To address this qualification issue, state DMVs are requiring companies engaged in public road testing of CAVs to generate safety-related disengagements. For 2016, safety disengagement reports showed that some companies report numbers as low as once per mile, while others ranged in the hundreds of miles. In comparison, human drivers in the U.S. traffic safety statistics show about 2 million miles per injury crash, and 100 million miles per fatal crash.

As vehicles with varying levels of automation are placed on public roads, qualifying information to determine the safety information must be generated. Many new-generation vehicles have excellent embedded passive systems, and no doubt, safety numbers will improve. The number of fatalities on our road network had a significant downward trend as automakers made the vehicles more safe. A key challenge going forward will be in educating vehicle owners on the capabilities of these systems. Recent CAV accidents reflect a misunderstanding of the vehicle's level of automation. Today, no publicly available vehicles on the road allow the driver to disengage from the driving process. It may be many decades before vehicles in the wild become completely automated.

The DOT Connected Vehicle program maintains a record of the collective safety of OEMs engaged in CAV testing. NHTSA is incorporating the information in its vehicle-to-vehicle and vehicle-to-infrastructure communication efforts. By 2021, all vehicles must be equipped to broadcast a basic message of what they are doing and what they encounter, creating an environment in which all vehicles in the vicinity can learn from and be aware of one another. The technology is codified in a set of standards developed by the Institute of Electrical and Electronics Engineers and Automotive Service Excellence.

With all these efforts in place, we have yet to determine how to certify that a new vehicle is safe enough. What conditions does it need to be driven in to validate safety? What will be the time and cost? Experience from the aerospace industry suggests that we have a long way to go, given that software validation and verification represents 50 percent of a new aircraft's cost, and aerospace challenges are orders of magnitude simpler than driving a CAV on public roads.

Creating a Vision Through Policy

Policy has an instrumental role in the speed and style of this mobility transformation. Consider mobility as a holistic ecosystem of moving people and goods. We must focus on land use, not only in cities and dense core areas, but also in suburban and rural areas.

In principle, we can use policy and public funding investments to shape the kind of world we want to live in and use emerging technologies as an opportunity to correct the things we don't like today. Cities formed in the streetcar era were compact and walkable, but with the advent of automobiles, they quickly became less dense and spread out. This change was accelerated by investments in a transportation system and an infrastructure that supported single-occupancy vehicles, thus carving up cities, splitting neighborhoods, and damaging their vibrancy. In recent years, investment in roadway infrastructure has slowed, and the goal has refocused to optimize highway performance and improve multi-modal options, creating a resurgence of city cores.

The Urban Landscape

Transportation costs in terms of time and money have a bearing on how far away from cities people are willing to live. Consequently, policy can influence demand through pricing and cost structures.

City governments are having an increasingly important role in developing transportation solutions because they govern access to right of way and are stewards of the infrastructure. Going forward, they will need to leverage their authority to institute policies.

Dynamic Data Requires Dynamic Policy Making

A key challenge is that technology and new economic models are moving much faster than public policy**.**

> ***In this rapidly changing technology environment, policy needs to be more adaptive and dynamic.***

Massive infrastructure projects always create some uncertainty. We must be more dynamic with our predictions and continually revisit the questions. We will need a more incremental and agile form of planning that is infused with better and more informed data at different decision points to update and rethink our planning.

The pace of technology creates debate in the community as new solutions emerge. For example, some decision-makers are asserting regulation for automated vehicles, while others have called for open innovation and no regulation. The rapid growth of TNCs has spurred activity to understand their part in the ecosystem and how to regulate their impact on mobility. If the public sector's response is too slow, will industry step up to set the agenda? Private industry is focused on their own business agenda. It would be unwise to expect something different. Yet what do we do in the absence of data, what do we do to protect privacy, and what do we do about proprietary-related considerations?

Policy also needs to address consumers equitably, making mobility affordable and accessible to all, and understand consumer behavior. For example, to drive AFV adoption, is it enough to provide pricing incentives? In terms of regulation, we tend to lean toward incentives, but incentives often reduce equity. Furthermore, when the incentive is removed, it is unlikely that the same choice will be made the next time. Instead, policy must directly affect the customers and their demand.

THE CALIFORNIA COMPLETE STREETS ACT OF 2008 AB1358 REQUIRES THE CIRCULATION ELEMENT OF GENERAL PLANS TO BE MODIFIED TO: "PLAN FOR A BALANCED, MULTIMODAL TRANSPORTATION NETWORK THAT MEETS THE NEEDS OF ALL USERS OF THE STREETS, ROADS, AND HIGHWAYS DEFINED TO INCLUDE MOTORISTS, PEDESTRIANS, BICYCLISTS, CHILDREN, PERSONS WITH DISABILITIES, SENIORS, MOVERS OF COMMERCIAL GOODS, AND USERS OF PUBLIC TRANSPORTATION, IN A MANNER THAT IS SUITABLE TO THE RURAL, SUBURBAN, OR URBAN CONTEXT OF THE GENERAL PLAN." THE ACT DECLARES THIS TO:

- **SUPPORT THE TARGETS IN THE CALIFORNIA GLOBAL WARMING SOLUTIONS ACT OF 2006, TO REDUCE GREENHOUSE GAS EMISSIONS FROM TRANSPORTATION.**
- **SHIFT TRANSPORTATION MODE SHARE FROM SINGLE PASSENGER CARS TO PUBLIC TRANSIT, BICYCLING AND WALKING, TO REDUCE VEHICLE MILES TRAVELLED.**
- **REALIZE ADDITIONAL BENEFITS FROM WALKING AND BICYCLING OF IMPROVING PUBLIC HEALTH AND REDUCING TREATMENT**

On Urban Change

"We have a long way to go ... it's interdisciplinary work, and we don't have much time." – Paul Waddell

Using Models to Inform Policy

The transportation ecosystem is complex, so modeling must reflect the range of metrics involved in pursuit of an optimal solution. It requires a multidisciplinary approach, diverse datasets, and multiple tools, incorporating the transportation network, energy consumption, economic impacts, life-cycle analyses, land use, safety, accessibility, equity, emissions, and behavior into a common framework. And the metrics need to constantly adapt as new technologies emerge.

Reflecting reality: What data will we collect and what data do the models need?

Today's modeling of cities to support regional planning captures the induced demand effect by coupling land use and transportation. To capture the true dynamics, it is necessary to model the real estate, housing, and nonresidential markets, how businesses and households make location choices, how developers decide what to build and where, and how zoning and the type and placement of new infrastructure affect those decisions. Having a unified insight into combinations of land-use policies and transportation investments will help us predict the consequent energy consumption profile.

Every system simulation needs to be multimodal, combining data to address transportation issues holistically. Using an integrated multiscale model with feedback loops determines how each element interacts with and affects the others, beginning with the physical system model and then feeding in various data sources, such as land use, population distribution, activity engagement, household resources, value of travel time, mode choice, and vehicle ownership and types. It must look at many different time horizons: long-term, mid-term, and within-day choices.

Urban Sim

The open-source UrbanSim model couples with a transport model system to enable bidirectional analysis of transport policies and land use policies aggregated up from the individual cities. ActivitySim, an activity travel model system based on the CTRAMP family of models that a number of MPOs use, is still in development.

However, the computational costs of current transportation models are high and do not deliver in the time frame that operational planners need. Data is commonly in a variety of formats, with no standardization for many aspects of travel modeling. A major challenge is to digitalize all the information.

Standard formats for mobility data will be essential to all transportation modeling.

Most metropolitan planning organizations (MPO) have neither the staff nor expertise to put the pieces together. Although the models are evolving, the models most agencies use do not address the next generation of mobility topics. While some cities have accelerated the amount of data that they are collecting and sharing, some datasets are incomplete due to personally identifiable information (PII) and other issues. Data collected by U.S. government agencies often focuses on particular issues (for example, vehicle-to-vehicle communication testing) and lacks the details required for a full suite of smart mobility applications. Private companies are the biggest repository of data, but their business models make data inaccessible or cost prohibitive.

Behavioral Modeling: Why Decisions Are Made

To predict demand and usage requires understanding behavior: Who uses it, when, and where. Behavioral models must focus on long-term decisions, such as vehicle purchase, housing, and workplace choice, as well as short- to medium-term decisions, such as activity generation, mode choice, destination choice, timing decisions, scheduling, and route choice.

Behavior is inherently contextual and varies substantially. We need to implement new behavioral experiments as new solutions develop, using the emerging technologies to our advantage, including simulation-based scenario analysis, survey responses to hypothetical scenarios, virtual reality and gaming, and field experiments using analogous modes and prototypes.

Designing Transportation System Solutions: It Starts with Data

When designing mobility solutions, modeling and simulation are essential to explore what-if scenarios, test policies, design infrastructure, predict performance, examine behaviors, and calculate energy impacts. But models are only as good as the data that drives them. And the data that researchers need to make valid conclusions is at a level of detail that private individuals fear to release.

The Sunshine Act and Freedom of Information Act limit government's ability to protect the data that drives their decision processes. Solutions have to rely on data aggregation and trusted individuals. For example, when comparing competitors in a duopoly, releasing aggregated data gives the companies a view into their competitor's data. Levels of aggregation and questions of proprietary data and privacy need to be woven into the process thoughtfully as we move forward.

Can we find ways to share data so that it does not impact a company's business model, trade secrets, and competitive advantages? What data are people are willing to share and with whom?

The open-source, free model has its advantages in that it puts ideas and solutions into a public forum and breaks practitioners out of siloed thinking. The many contributors increase the diversity of perspectives. Open and free is a good start, but in the long term, will the best approach be to establish collaborations that create harmonies to enhance the next stage of work?

We should seek to build standards in our software and data sharing. By standardizing the interface to the data, we can build an ecosystem that can be more easily integrated.

Data: You Can't Always Get What You Want

Many companies state that privacy issues prevent them from sharing mobility data. Next-generation data collection systems might solve these issues by building in privacy design models that ensure privacy is maintained by isolating data that can be used without impinging on privacy. One solution to the privacy issue is to determine what needs to be kept private, decouple the data, and make two independent factors—one that can be shared and one that's kept private. This assumes that the resulting data still has value once decoupled. As such, it will be a function of which analytics are being pursued.

On Finding Relevant Data

"Think global, act local." Aymeric Rousseau

Ownership of the data can be tricky as well. To share data with a third party, you must have the rights to the data. In general, machine-generated data belongs to the owner of the machine. If a car passing through picks up road data and you want to give that to another vehicle passing through, that's safety data, which can be considered public information. What are the implications of this data collection and merging process? Who then owns the data?

The National Research Council recommends that research involving detailed data combined with demographic information be performed in a secure data center that makes data available for legitimate research while preserving privacy.

Some algorithms come with theoretical proofs on how well they can protect privacy. Continuing research is moving us toward a "gold standard" that can be an agreed-upon metric. When everyone agrees on one definition of privacy, it will become a lot simpler, but we are not there yet. Differential privacy is a statistical technique that attempts to learn as much as possible about a group without learning details about the individuals. It is popular because of its modularity—as data moves from company to company, you just need to add the next company's parameters. Differential privacy is likely to dominate the market going forward. AI and powerful computing clusters are also being used to resolve interchange and overlapping issues.

Secure Access to High-Resolution Travel Data

The Transportation Secure Data Center (TSDC) at NREL allows trusted researchers to work with detailed travel data. The TSDC hosts publicly conducted travel surveys where high-resolution travel data has been collected and is valuable for research, but where misuse could violate individual privacy.

National Data

At a national level, many efforts are underway to archive and distribute data.

University of Maryland CATT Lab
The Center for Advanced Transportation Technology Laboratory focuses on information visualization, data fusion, performance measurement, user interface design, games and online training, and operations. It disseminates information to transportation agencies and develops tools to visualize and analyze transportation data to help understand how their services and infrastructure are performing. The Transportation Energy Analytics Dashboard is an online suite of tools for monitoring energy use and emissions in real time and historically. Examples include the ability to plot incidence response in parallel with congestion scans to see how a roadway is performing in terms of travel speed.

NREL Transportation Secure Data Center
The National Renewable Energy Laboratory is working with trajectory datasets in Columbus, Ohio, to analyze charging infrastructure siting to support Smart Cities activities. Real-time data shows if a station is occupied. The aggregated data can be used to understand station usage and how often vehicles are parked but not charging.

Data Depot: Creating a Sharing Economy for Data

In the information economy, data is a key asset for many companies. But often, some of the data collected is not a critical business differentiator. Can we create a sharing economy for data in which institutions are able to exchange or rent data? A transportation data marketplace could be run by a private entity to allow buyers and sellers to acquire or provide data.

This model offers several advantages. Data owners have a financial motivation to share data. Data buyers pay only for what they need, and as the market grows, pricing is more competitive. Sharing data improves overall market efficiency, because the process of collecting data is managed by few and used by many.

Creating good quality data creates value. In a data marketplace, cities could put data in an open data repository where someone can offer data enhancement services, or prospective users can buy it from the marketplace where it has been cleaned and processed. The fee would save hundreds of hours of staff time, staff that the buyer might not possess. It creates a value-added data product for those who need it.

The beginnings of a marketplace for real-time traveler information and travel time data analytics is already in place. Several state DOTs are buying and licensing data from various private providers. Some larger data collection companies also gather data, clean it, and sell it. However, a key difference between that model and the data marketplace is that there are currently only two or three companies, which allows them to control the price. In the proposed data marketplace, many participants would exchange data, making it more affordable and varied.

The proposed marketplace would encourage government agencies and businesses to keep and exchange data that they tend not to keep beyond its immediate use. If the data itself has value, either as something to sell or exchange, entities would benefit by participating.

One person's data exhaust could be another person's gold.

Cities have data that private companies want to use for product targeting or developing applications, and cities need data from the private sector to help them understand the mobility dynamics of their city.

Acquiring, storing, and maintaining data is costly, and pricing data isn't straightforward. The marketplace's ability to create a fair, revenue-sharing model rests on our ability to accept data as a commodity.

End users could also contribute to the marketplace. For example, utility data and cell phone data is end-user data captured by a separate entity. Should we create the mechanisms to shift the ownership of end-user data back to individuals who might be much more willing to take risks in the marketplace?

Uncensored Sensors

With sensors everywhere, IoT data sources are growing rapidly. Devices that were originally built for a single purpose are finding uses in a variety of different applications. Modern vehicles are rolling sensor platforms, carrying from 70 to 300 sensing devices. Some, like barometric pressure and temperature, can be sampled at high rates and can be used to calculate fuel efficiency, emissions, and power. We all carry mobile devices that essentially track our movements and capture our application usage. Telecommunication companies, ride-sharing networks, and application developers are constantly uploading this information as we make the decision to release our private data in exchange for their services.

Imagine how commercial interests could benefit from sensor information collected by truck fleets nationwide. Example data that a data exchange marketplace could provide includes regional meteorological and road condition databases, vehicle-derived sensor outputs, human-activity tracking, topical data streams, fleet operators, local last-mile deliverers, insurance carriers, and regional agencies.

Two Start-Ups

Terbine and Voyomotive have partnered to work with fleet data. They are conducting a trial in a U.S. city with a uniform fleet of 500 city-owned paratransit vehicles to demonstrate acquiring data from vehicles without driver participation. Data from wheel slippage, wheel turning, braking, and acceleration indicates that cars avoided an obstacle. Using AI, you can deduce or induce what the obstacle might be. The city's goal is to know whether to send a policeman, ambulance, fireman, road maintenance, or ignore the report. This information can also be used for signal management, and insurance companies are looking at this for insuring cities and fleets.

Building Effective Data Exchanges—Experiences from the Wild

Many cities have introduced technology-driven services and have strategic plans for exploiting technology for its citizens' benefit. As people become more dependent on smartphones for transit information, cities are making their fleet data public so that application developers can provide informational services for consumers. As IoT expands, sensorizing cities to generate a more complete understanding of their dynamics will no doubt impact government services design and urban planning.

To date, at least 30 cities have already established chief data officers whose role is to use data to drive operational efficiency and improve the quality of life in their cities. In the last three years, they have become more involved with transportation research, and in the next 5–10 years, more transportation research will be done at the city level than over the past 60 years, in part because of the large growth of data. Cities are challenged to staff this work going forward. Creative funding models need to be established to allow outside researchers to supplement the staff.

Cities are beginning to publish public data on data portals, which have become an effective mechanism to provide data to others, such as universities and national lab partners. Data can be accessed for free and partnerships can be established to address a city's problems and solutions. However, this requires a large amount of staff overhead to manage, because cities can't meet with all that are interested in one-on-one discussions.

Three Cities
Chicago, Illinois Chicago has 115 transportation-related datasets on its Open Data Portal (data.cityofchicago.org) and has prioritized making data available publicly. To maintain a unified structure, the city employs a chief data scientist to manage all the data from the moment it goes into the database through publishing it on the open portal. In contrast, San Francisco is highly decentralized—every agency has its own stand-alone CIO. In New York City, the CIO sits in the mayor's office and does not control any data but instead focuses solely on the analytics.
In Chicago, 28 percent of households do not own a vehicle. The bike share program is profitable and accounts for over 10 million trips during the past four years. The city has data for almost 5,000 bikes on line and also has daily data for vehicles licensed to operate as a taxi, including the vehicle type and the fuel efficiency, as well as taxi trips since 2013 until present, which are updated monthly. To protect privacy, the data is masked for 15-minute intervals and is rounded up to the census track or neighborhood.
Ridesharing companies that operate in the city must provide monthly data on origin, destination, timestamp, vehicle, and driver for every trip, but the data is not published due to company concerns.
Portland, Oregon Portland introduced PORTAL in 2004, provides transportation and intelligent transportation data. It collects data from transit agencies, signal systems, ITS data, and more for use by agencies, staff, consultants, and other stakeholders. It consists of speeds, volumes, travel time, vehicle lengths, transit ons and offs, on-time performance, GTFS, bike counts, traffic signal data, incidents, and more.
The Smart Cities program is collecting pedestrian data to understand intersection safety, and regional collaboration around Smart Cities in collaboration with a mobility needs assessment.
Oakland, California Oakland was identified as the kick-off city for the Bloomberg Associates' Equity Intelligence Platform. The city is using data and models to translate values and goals into visions.
The cities of Berkeley and Oakland recently partnered with a for-profit company and introduced the first multi-jurisdictional, free-floating car share program in California. The fleet consists of 250 Toyota Priuses. Every parking meter space has been geo-referenced and will true up with the vehicles' GPS to know how long and where each vehicle is in use. The data collected will drive the city's shared mobility initiatives.
In Oakland, an open invitation from the mayor to help the city of Oakland develop the proposal attracted responses from around the country. The city raised the question of whether cities should be competing for grants, entering into franchise agreements like New York City did with LinkNYC (now Google or Sidewalk Labs), or working together toward a third way, where data is in a public or public-private trusted, democratic environment. After the Smart Cities challenge, Transformation for America was sponsored by Sidewalk Labs to initiate an ongoing collaborative, and 16 cities are now participating.

Regional and State Solutions: Helping Government Keep Pace with Industry Changes

State DOTs are working with legacy systems and their shortcomings while preparing for the future. An added challenge is that the infusion of technology impacts a state's cities differently. Each city's budget defines its ability to move forward with new technologies. As a result, implementations can vary widely across city borders, conflicting with city planners' efforts to maintain a regionalized view to adapt to population and business changes.

However, many cities are evolving from decentralized organizations to connected communities focusing on a system-wide view of transportation. This transition requires transforming a tabular data structure to a geo-enabled system—or even more drastic, moving data from paper to digital solutions. However, a National League of Cities study found that of the 50 states, 50 large metropolitan areas, and 15 other areas studied, only 3 percent consider TNCs in their long-range plan, and only 6 percent consider AV technologies. To make good investment decisions, transportation plans must include these technologies.

The changing transportation environment is placing new demands on our institutions, such as motor vehicle departments and public utilities, to regulate these new and dynamic transportation trends with limited knowledge. Government organizations are not equipped to address the big data and machine-learning methods being used by the private sector to disrupt our current management processes. Research institutions can integrate the experience and knowledge of the regulatory institutions to enrich the analytics and modeling necessary to codify the dynamics.

Getting data of reasonable quality is difficult, and many different data sources are needed. In the field, a constant challenge is sensor validation. For example, in the Connected Corridors Pilot, a good percentage of the sensors are not positioned according to the specification or may be malfunctioning. But without good data, the end of the process will not provide actionable results. Data quality is a holistic challenge, involving hardware, software, the agency culture, organizations, personnel, and funding.

Some strategies that have been applied in the field include dynamic pricing for managed lanes, ramp metering, travel information from many different and sometimes conflicting sources, and parking management. A transportation manager can use these levers to modify performance. The challenging aspect is travel demand and people's behaviors.

We must move to an integrated view, where different agencies work together to deliver the services. So that when there's an accident, the transit lines can add cars, and the available parking garages and their prices are disseminated to mode shifters in real time. The result is that it is much more complex, but less balkanized.

Connected Corridors, I-210 Pilot
The Connected Corridors Program is a $16 million effort focused on a 20-mile integrated corridor. It involves 15 groups across multiple organizations and cities working together in Pasadena.

San Diego Integrated Corridor Management
This effort uses a real-time, multimodal decision-support system that consists of four steps.

1. Make a prediction every 5 minutes for 60 minutes.
2. Decide which meters to turn on and how to modify signals.
3. Evaluate the potential response strategies over a number of criteria.
4. Send control directives to the field.

Next-generation summation data will track vehicle activity on a stretch of road every tenth of a second. This information provides speed and acceleration data, and from that, emissions and energy consumption can be estimated. Most vehicles emit higher emissions when they accelerate quickly. As a consequence, most operational strategies and most of the connected technologies should have a positive effect on emissions and fuel consumption if they smooth traffic flow.

A key challenge is that many sensors fail in the field. For example, statistics for loop detectors show that approximately two-thirds of them are working at any given time, which is about the same as it was decades ago when there were far fewer sensors. To provide more reliable data, this maintenance issue needs to be addressed.

Improving Traffic Flow in the Near Term with Available Data

With the availability of high-resolution data, it's now possible to carefully assess existing traffics operations in terms of progression and capacity. A fully adaptive traffic control system is expensive and complex. An alternative pragmatic approach is to add a few detectors upstream of the intersection and a loop detector at the stoplight. With just these sensors, you can obtain a reasonable estimate of the status of the signal with the added benefit of a 24/7 data collection system. To get information about the signal, the conflict monitor, which is standard on every controller, can inform on the status and pace of the signal. An intelligent combination of these sensors provides a continuous block of data and traffic and approach volumes, which in turn can generate solutions to the signal timing. From this, one can empirically devise better timing plans.

These new timing plans can reduce intersection signal delay by 10 percent on average. These approaches have significantly improved by moving from point-based sensing to trajectory evaluation using probe data, such as GPS data from smartphones. Here again, a lack of data inhibits the progress of city governments. Cities must purchase the data from private entities, and the resources are simply not available.

State-of-the-art systems use machine learning, which require significant resources for calibration—an Achilles heel for most modeling and simulation tools. ALINEAR is a route metering system that optimizes the flow from on-ramps and keeps track of the freeway flow. Pacing traffic flow reduces the stop-and-go, which in turn reduces emissions and fuel consumption. However, often the goal is not clear. Should the system optimize throughput or minimize emissions? These two objectives are antithetical with one another, and there will always be a human in the loop to make those decisions as a function of external information, such as when it is a Spare the Air Day.

Tapping National Lab Experience and Capabilities

Metropolitan transit agencies would greatly benefit from computational systems that could generate a demand profile for their urban landscapes. However, the tools they have at hand are computationally limited, potentially taking days or weeks to run a simulation at scale. High-performance computing (HPC) could enable this capability with the ability to analyze large datasets and more complex problems.

The DOE National Laboratories offer hundreds of petabytes of storage, large-scale analysis, and a network that connects lab datasets to the Internet, as well as applied mathematics modeling and simulation. National Lab experts can be tapped to advise companies, governments, and other researchers on effective ways to approach transportation data and computing issues. HPC machine-learning capabilities could add new dimensions to reasoning about the dynamics of mobility.

Benefits of Test Facilities and Open Road Testing

Unlike testing on public roads, secure test facilities provide users with the opportunity to test their technology to failure. They provide a space where vehicles can travel at faster speeds and merge in ways not possible on public roadways. In addition, the testing can be done without a DMV permit.

Data sharing is important to identify the approach and handling across multiple car companies of the types of incidents that force failure.

Mcity
The University of Michigan's Mcity (formerly the Mobility Transformation Center) is a public-private R&D partnership focused on mobility and developing the foundations of a commercially viable ecosystem of connected and automated vehicles. It was specifically planned to not be a DOT-designated testbed.
Because data challenges must be addressed through a team approach, Mcity combines labs, research, education, and outreach with industry and government partnerships. Industry partners work with faculty to road-test their technologies in real time. The multidisciplinary team includes researchers from energy, public policy, medicine, business, and law. Partners include major automakers, data companies, tech companies, cell phone providers, insurance companies, toll collectors, modeling and simulation companies, and mobility providers.

Contra Costa Transportation Authority: How Small Organizations Use Data
The CCTA GoMentum program provides a hub where CV and AV technology, innovation, and business converge. The 2,100 acre testbed is a mini-city featuring old roads, two-bore tunnels, and many intersections. Safety is the top priority. Various vehicle, freight, and Internet companies make use of the facility.
CCTA shares data as a designated AV proving grounds, and it also sends data internationally through partnerships. CCTA is working to create the largest shared AV pilot to share information and data with other countries.
To meet California's GHG reduction goals, CCTA is updating how congestion is modeled. But to do so, agencies need better data to model transportation trends, future transportation systems need data that are interoperable across modes, and manufacturers need to share data to help the transportation system adjust and scale.

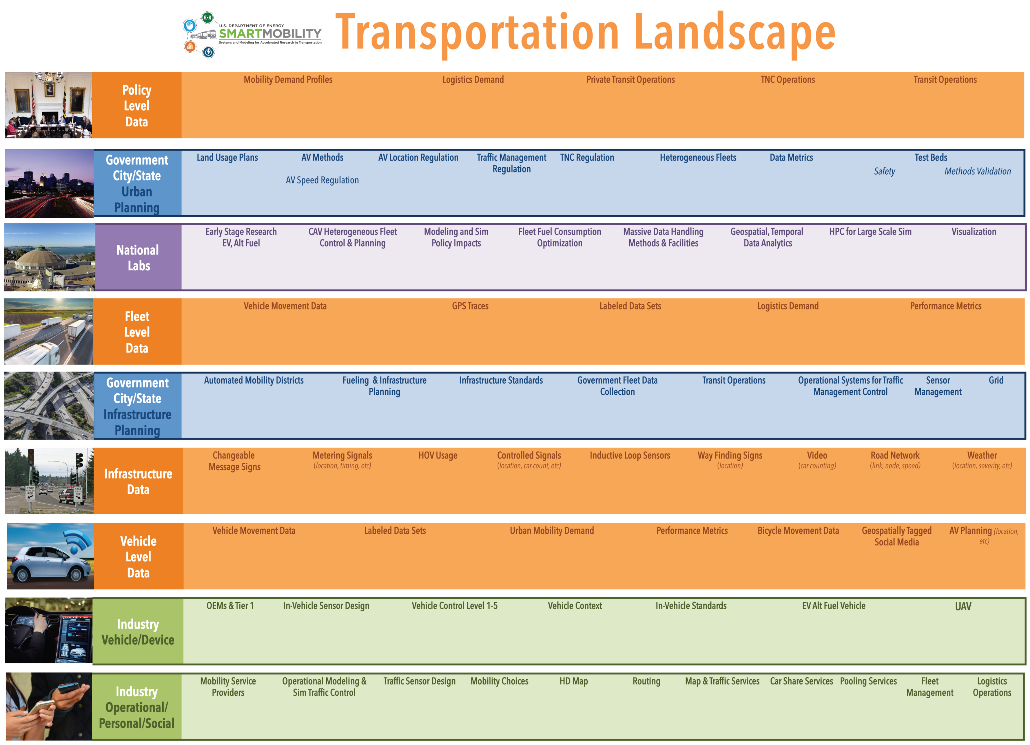
The Transportation Landscape

This paper presented the many perspectives of the workshop participants who represent the range of entities involved in delivering integrated transportation solutions in light of changing technologies, environmental goals, infrastructure demands, and shifting populations. As demonstrated, the system is complex, and all parts are under transition. The examples included are from the participants'

Contact Information: Jane Macfarlane, UC Berkeley Institute for Transportation Studies.
janemacfarlane@berkeley.edu

firsthand experience and are just a representative sample of the thousands of private companies, universities, and government agencies involved.

The Transportation Landscape chart shows the various entities engaged in transportation and some suggested roles as we make this critical transition in how we address our transportation infrastructure. It is layered by level of abstraction from the devices in the field. The orange layers represent the data collection requirements and opportunities at each level of abstraction.



Entities include:

- City and state governments involved in urban planning and infrastructure planning
- Universities and national laboratories, supported by the DOE and other agencies, providing the research to support the mobility understanding necessary to make good decisions during this transformation
- Private industry engaged in delivering vehicle, device, and service solutions to consumers

Each organization layer has the opportunity to generate data that will give us insight into mobility demands. Examples are shown in each entity layer.

Private Industry

- Consumer Service Providers
  - Functional role in the ecosystem
    - Traffic and routing services
    - HD map services for automated vehicles

        - Pooling services, car sharing services
        - Fleet management services
        - Management of logistics operations, multi-modal routing services, including bike share
        - Traffic sensing and control services
        - Personal data aggregation
- OEMs and Tier 1 Vehicle and Device Providers
    - Functional role in the ecosystem
        - In-vehicle sensors
        - Automation services (Level 1–5)
        - Vehicle context data
        - In-vehicle standards, unmanned aerial vehicle management
    - Vehicle-level data collection and requirements opportunities
        - Vehicle movements data via GPS sensors
        - Large labeled datasets from on-board cameras and other sensors
        - Urban mobility demands from smartphone app usage
        - Bicycle movement data from shared bike services
        - Geospatially tagged social media data
        - Unmanned aerial vehicle movements from GPS sensors or flight plans
    - Infrastructure data collection and requirements opportunities
        - Statuses of changeable message signs on major freeways
        - Ramp-metering signal profiles for major corridors
        - HOV lane occupancy profiles
        - Inductive loop sensor data
        - Signal timing plans for major signals
        - Video-based car counts at key locations
        - Road network status changes over time
        - Archived weather

Government

- Transportation Infrastructure Planning Organizations
    - Functional role in the ecosystem
        - Automated mobility districts
        - Fueling and infrastructure planning
        - Infrastructure standards
        - Government fleet data collection
        - Transit operations
        - Operational systems for traffic management control
        - Infrastructure sensor management
        - Grid integration—impact of CAV fleet if EVs
        - Innovation models
        - Evolution of public transit in context of social and environmental goals
        - Protocols for aggregations and access to data
        - Communication standards for automation

    - Data requirements and generation opportunities
        - Vehicle movement data from government fleets
        - GPS traces from government employees
        - Labeled datasets
        - Logistics demand—highway monitoring
        - Performance metrics
- Urban Planning Organizations
    - Functional role in the ecosystem
        - Land-usage plans
        - AV policy and usage limitations
        - Traffic management regulation
        - TNC regulation
        - Heterogeneous fleet policies
        - Data metrics for urban operations
        - Testbeds
        - Fair revenue sharing models
        - Open data policies in cities
        - Evaluation of longitudinal data
        - National transportation census
- National Laboratories and Universities
    - Functional role in the ecosystem
        - Early-stage research
        - CAV heterogeneous fleet control and planning
        - Modeling and simulation of scenario impacts
        - Fleet fuel consumption and optimization
        - Massive data handling—methods and facilities
        - Geospatial, temporal data analytics
        - HPC for large-scale simulation
        - Visualization
        - Impact of 3D printing, drones, courier network services, and alternative delivery models
    - Data requirements and generation potential
        - Mobility demand profiles
        - Logistics demand
        - Private transit operations
        - TNC operations
        - Transit operations

Moving Forward: Policy and the Information Economy

Data is forging a new economy, fueling growth and change in all sectors. It is inspiring innovation, driving commercialization, and influencing the choices we make. In this digital information economy, data is an asset, mined from a multitude of devices and the people who use them—and much of it is privatized. Our transportation system is a source and a conduit to all kinds of data, but who owns it and who benefits? Data can be a great democratizer or it can be siloed and benefit only certain segments of society. We are at the cusp of a transforming transportation ecosystem, and at this juncture, it's critical that commercial solutions don't fragment our transportation system and undermine our future.

To intelligently design infrastructure and systems for a cleaner, safer transportation network in the midst of this transformation, the workshop participants pinpointed these observations and action.

- Broaden the metrics that fuel transportation models

  Revisit metrics involving congestion, travel time, and throughput and ensure that metrics that encompass equity, access, and public health are incorporated. Ensure that a creative, multimodal transportation system is accounted for.

- Build models for acquiring data

  Create models that are more dynamic and adaptive and continuously update as new information is learned. Go directly to consumers to acquire data, and enable participants to volunteer their data. Collect and blend small, deep data and big, shallow data. Encourage collaboration between social sciences and AI.

- Engage the ecosystem to evaluate how planning and incentives and investment can support the future

  To address such a complex interconnected system, it's important to involve all stakeholders to understand their needs and motivations. Work directly with the companies that acquire the data, and develop partnerships to achieve win-win agreements. Ease the ability of people and companies to share data through advancements in privacy protection.

- Define a software and data system for speed

  Promote data sharing by standardizing the interface, and build system that could potentially be chained together. Can we learn from the global science communities regarding managing and sharing large data systems with metadata frameworks to accelerate our ability to impact the transformation underway?

- Speed up analytics

  Find ways to get access to data to do more scenario development. Many of the models available to planners take too long to run, limiting policy- and decision-making. We need to leverage DOE expertise in large-scale analytics and visualization and focus it on the transportation ecosystem.

- Scope and accelerate the city role

  Cities have an increasingly important role in our solutions and are a tremendous source for transportation data, which businesses are eager to have access to. Because they govern access to right of way and are stewards of the infrastructure, they need the authority to institute policies.

- Improve the sensing

The transportation system has been reliant on sensors to collect data, but the state of sensing is poor. Implement new technologies to understand mobility.

Appendix

Government Agencies

U.S. Department of Energy

Federal Highway Administration

National Labs

Argonne National Laboratory

Oak Ridge National Laboratory

Idaho National Laboratory

Pacific Northwest National Laboratory

National Renewable Energy Laboratory

Los Alamos National Laboratory

SLAC National Accelerator Laboratory

City and State Government Agencies

City of Chicago

City of Oakland

Colorado Department of Transportation

Contra Costa Transportation Authority

San Francisco Municipal Transportation Agency

City and County of San Francisco

San Diego Autonomous Vehicle Proving Ground

Industry

Iteris

eCalCharge

Volvo

TomTom

Uber

car2go

StreetLight Data

INRIX

HERE Technologies

Mercedes-Benz R&D North America

Voyomotive

Terbine

Metropia

Cambridge Systematics

Strategic Vision

Swiftly

Arup

Academia

University of California, Berkeley

University of Maryland

University of Michigan, Mcity

University of California, Davis

Clemson University

Texas A&M Transportation Institute

Portland State University

Purdue University

Silicon Valley Leadership Group