

Twin Cities Shared Mobility Data Seminar and Workshop

Event Summary

Whether it is shared cars, bikes, scooters, or their relationship with transit, shared mobility transportation services are creating huge amounts of data, both about the trips that individuals are taking and the services themselves. These data sets are critical for planning and operating transportation services and understanding how these services are affecting our communities. They can show how people get around, who and where these services benefit, and where services may improve access or add to congestion and emissions.

While the opportunities are clear, accessing and using these data sets is not easy. The volume of data is immense, not everyone has the tools and expertise needed to turn it into useful information, and there are valid concerns that the data becoming available could reveal too much information about individual travelers, and about the transportation providers.

How can we balance the desire to create and use immense amounts of data and the need to protect individual privacy and intellectual property? To help answer these questions, the Twin Cities Shared Mobility Collaborative held a Data Sharing Seminar and Workshop on July 11, 2019. Local and national experts gave presentations on the many complex and interrelated details of working with shared mobility data.

Key takeaways from the day include:

- **Cities have leverage to manage mobility data.** The City of Los Angeles created an API to help the city manage shared micromobility vehicles. Los Angeles has also made equally critical strides to adapt its staffing to new needs and skills around data.
- **Minneapolis is also a leader**, using similar processes as Los Angeles the city is using to regulate micromobility, including creating a mobility management center.
- **Transit agencies and others need to balance privacy needs** and find new ways to tap valuable information while ensuring individual trust and anonymity.
- **Historical data and real-time data pose different challenges** and opportunities. Historical data is useful for planning and real-time data crucial to move towards mobility as a service, among other uses for each. In each case, it's critical that the data be right.
- **Privacy laws vary from state to state**, and are essential to understand in order to navigate them. In Minnesota, government agencies must give individuals notice when collecting private or confidential information from them.
- **Common ground between agencies and organizations facilitates data exchange**, including agreement about the level at which data will be aggregated.
- **Organizations need to lay the foundation for data collection.** Data management, data sharing, privacy policies, and interorganizational agreements are all critical.
- **There are tools and resources for decision-making.** Among them:
 - Cecilia Viggiano presented a [Data Sharing Decision Framework](#)
 - Prashanth Gururaja presented a decision tree for [Data Sharing in Mobility Partnerships](#)

Following the morning seminar, individuals working closely with mobility data were invited to an afternoon workshop. Participants engaged in a series of facilitated discussions aimed at generating real solutions to the challenges with shared mobility data in the Twin Cities. Conversation was focused on a variety of issues including both legal/policy/ethical issues as well as technical/logistical issues with creating, storing, accessing, sharing and using data associated with the growing range of transportation options. The workshop culminated in a set of action items, which are outlined at the end of this summary document.

Discussions included a group focused on technology and logistics, and one oriented towards governance and policy. Working groups have been created to develop solutions for the top priorities identified.

- **Top priorities expressed for technology and logistics work in data sharing:**
 - Connecting practitioners
 - Agreeing on data sharing procedures
 - Building staff capacity and skills
 - Identifying data uses and goals.
- **Top priorities expressed for governance and policy work in data sharing:**
 - Understanding the current laws to identify any gaps
 - Building public trust
 - Overcoming barriers for equity challenges

Details from individual presentations and discussion sessions follow in the full summary.

Keynote Presentation:

Mobility Data Specification

Marcel Porras, Chief Sustainability Officer, Los Angeles Department of Transportation

In preparation for the 1984 Olympics, the Los Angeles DOT developed an Automated Traffic Surveillance and Control (ATSAC) system. Over time, the ATSAC has grown and allowed the LA DOT to develop a robust support staff for data and technology systems. Across the country, the advent of smartphones and expansion of Transportation Network Companies (TNC) has increased the amount of mobility data drastically, but private company data has not always been accessible.

In Los Angeles, the role of the Department is that of regulator, and as such, permitted dockless on-demand mobility providers must share real time data with the LA DOT so that it may effectively regulate and manage the dockless mobility program. This data is managed with MDS, Mobility Data Specification. Data includes real time location, start trip data, end trip data, vehicle utilization, among other things. MDS is focused on vehicle specific data and purposely does not collect personally identifiable information. MDS utilization ensures that data is being shared in a common language, and allows Los Angeles to actively manage shared vehicles using the public right-of-way. MDS data is an open source project.

With real-time data, traffic control staffing and skill needs changed and grew. Questions arose surrounding use of the data – should cities manage and have access to this data? To combat some of these concerns, LA DOT published data protection principles. LADOT developed a master license agreement that governs how data is accessed and managed by third party companies with a focus on ensuring privacy.

As Los Angeles looks toward the future, it is developing capacity and infrastructure around being able to provide real-time data in a two way direction between the city and an end user, such as providing real-time curbside rules or parking rates to driver. Further development of drones and other new transportation technology may need to be incorporated. LA DOT is hoping to set the base standard for data sharing now to realize the potential of the future.

Key themes:

- Cities/agencies need to be proactive; create a framework for data sharing when entering a shared mobility environment
- Data standards and privacy agreements should be established early
- Collaboration is key

Local Implementation Examples

Josh Johnson, Advanced Mobility Manager, City of Minneapolis

When Minneapolis launched a scooter pilot program in 2018, the City focused on ways to enable consistent, reliable, and trustworthy data. Collaboration with the City's Information Technology and Analytics team were crucial to the development of these tools.

Data privacy is a primary concern, with several steps taken to anonymize the data. Trip data is processed, anonymized, and stored in a secure database. Raw, identifiable data is never saved.

Additionally, trip times are rounded to the nearest half hour and trip points are sorted to the closest of three pre-determined segments on the street centerline.

Providers are required to utilize MDS to provide the City with data. This data can then be used to track vendor compliance with the pilot program's requirements and analyzed for planning purposes. To handle and process this data, the City developed a data methodology. Anonymized data from the pilot project is also held in the City's open source portal. On the planning side, the aggregated data is being used to determine scooter parking locations, inform infrastructure improvements and system enhancements, and understand user mode choices.

Key themes:

- Data privacy and management must be planned and executed properly
- Possible uses for accurate mobility data are wide-reaching for cities

Eric Lind, Manager of Research and Analytics, Metro Transit

Metro Transit has already been collecting shared mobility data through its own vehicles, including vehicle tracking and passenger counting. Vehicle data allows Metro Transit to collect locations, schedule adherence, and analyze whether or not the lift ramp is employed. Door readers on the vehicles allow for passenger counting, including segment loads. These two data sets can be used to calculate throughput and measure and predict delays. There are also challenges wrangling this data. The logged messages must move through an operational database, be extracted and loaded into the agency data store, cleaned, and manipulated into reports that become useful to the agency.

A second set of data is produced from individual travel records. Smartcards, the Metro Transit Mobile App, and on-board interviews promise to help the agency understand individual behavior. In addition to understanding the base ridership, data can be analyzed to review the impacts of new routes or route changes. Challenges occur with this data set as well. The Smartcard system does not require tap-offs, resulting in unobserved behaviors. Neither the app nor smartcard have a way to track GPS or vehicle exit, which results in lower data utility.

Agencies need to balance utility versus privacy, especially in regards to individual ridership data. Public trust is critical for any City or agency service. According to NACTO, the four principles for managing mobility data include: public good, protection, purposeful, and portable. These principles will continue to be a challenging balance as increased technology allows for smarter, more thorough data collection.

Key themes:

- Properly anonymized data comes with limitations
- Privacy vs. utility in public data collection

[Real-time Information Stack](#)

John Levin, Director of Strategic Initiatives, Metro Transit

When collecting shared mobility data, it is important to differentiate between historical and real time data. Historical data includes who, when, where, and how of each trip. It produces an immense amount of data and is often accompanied by significant privacy concerns. These two elements cause a technical

complexity to data analysis. Real time data, on the other hand, consists of services and vehicles that are available at the present time. There is also a large amount of data in this group, but there are fewer privacy concerns. Again, there is a technical complexity to the data flow and the ability to aggregate it.

To successfully integrate shared mobility solutions, real-time customer information will be key. Real-time information would allow a traveler to access all available options to travel from point A to point B, select a mode, pay for the trip, and travel seamlessly. For this to work, service operators also need access to real-time data. Vendors and agencies must make sure their vehicles and services are positioned to meet needs. This real-time data flow involves six steps:

- Operational systems generate data
- Data is converted into standard formats
- Data “publishers” make the data available via the internet
- Data “consumers” download data files and interpret them
- The data is aggregated from multiple sources
- Aggregated data is made available in small amounts for mobile apps, kiosks, web services, etc.

To function, real-time data must be accurate. Wrong information is worse than no information. The data must be complete, accurate, timely, and multimodal. Service operators must have the technical capacity to generate the required data and then be willing to provide data in a standard format. End user applications will ultimately drive this data market, including data presentation and privacy.

Key themes:

- Real-time data will be required to integrate shared mobility solutions
- Challenges to successfully and safely aggregating this data are present for the public service operators, TNCs, and end users

[Panel: National and State Data Privacy](#)

Moderator: Danielle Elkins, Fuse Executive Fellow, City of Minneapolis

Stacie Christensen, Director of the Data Practices Office, MN Department of Administration

Matthew Daus, Chair of the Transportation Practice Group, Windels Marx Lane & Mittendorf, LLP

Frank Douma, Director of State and Local Policy Program, Humphrey School of Public Affairs

Gregory Rodriguez, Counsel, Best Best & Krieger LLP

The panelists discussed the status of regulations surrounding data privacy in the United States from their perspectives, including risks to data collection and sharing broadly. A brief recap of the questions and panelist responses are included below.

Question: How is the conversation in the U.S. changing based on these new data concerns and laws? How is the international conversation impacting actions in the U.S.?

Rodriguez began with an observation that movement has shifted attention to legal parameters of data use, but informed consent in terms of collecting personal data is still very muddled in the United States. With adoption of GDPR in Europe, seeing some movement recently at the state level, such as the California Consumer privacy Act. Conversations around national standards are starting, but there is a challenge in finding the right balance of privacy and functionality. Daus added that the future is multimodal integration, and the ownership, accessibility, and interaction of the data is crucial. There is

currently no legal standard or technology framework to allow this platform to happen. For both consistency and privacy, a third party repository may be a place to start. Douma cautioned that one complicating factor to consider is that many of these large companies are not acting in one state alone, or even one country alone. We need to consider that California's legislators, for example, could be impacting Minnesota operators. Christensen brought the local perspective. In Minnesota specifically, a Tennesen warning is required before collecting private data. Limited types of data are legally classified as private data. For example, most state data is considered public data.

Question: In addition to LA, as noted in our keynote, what cities around the country should the Twin Cities be looking to for examples of how to move forward, processes to consider in the next phase?

Rodriguez first noted the city of Seattle, which has developed a robust model that considers privacy issues. He also pointed to universities around the country, as academic institutions doing research for decades, which know how to collect and handle data safely. Christensen pointed to Rochester, Minnesota, and their temporary data classification from the commissioner of the administration. It is important for cities and agencies to be thoughtful and proactive prior to collecting the data. Daus also pointed to the international sphere. While the laws are different, countries such as Abu Dhabi have data analysis far beyond our own country. It can be helpful to look to these international examples for possible uses of data.

Question: For cities that do not have resources to dedicate toward data analysis and proper anonymization, what is the risk?

Douma pointed to loss of public confidence as a major risk if constituents realize their data is being collected and are not reassured of its use and handling. Once public confidence is lost, it can be very difficult for a city to move forward and innovate moving forward. Christensen agreed; agencies need to be thoughtful about data collection and be transparent with the public on how and why data is being collected and used. Rodriguez and Daus agreed, both suggesting that city attorneys can be a very valuable resource for agencies that are looking for guidance and resources. Contracts and agreements need to be clearly outlined before the data work begins.

Key themes:

- Understanding data privacy laws in Minnesota is critical
 - Be proactive with the legal community
- Foundations should be established before collecting data – a data management policy, privacy policy, allocate staff and resources
 - If a third party vendor is chosen to host regional data sharing, clear agreements need to be outlined for data management and ownership
- Demonstrate the need for data to build public trust

National Solutions

Cecilia Viggiano, Research Associate, Economic Development Research Group

For TRB's Transit Cooperative Research Program, Cecilia Viggiano and her research team have worked to evaluate data sharing from a public transit agency perspective by interviewing transit agencies, private mobility and data providers and other organizations across the country. Benefits of data sharing for

transit agencies include innovation and research, transparency and increased awareness of public transit, cost savings, data reciprocity, monetary gain, and customer benefits. Alongside these benefits are risks and potential costs: privacy, security, data misuse, staff time and resources.

In preparing for data sharing, it is important for agencies to understand the differences between public and private data sharing models and the risks of each. Public data sharing to consumers can take multiple forms, such as static reports, data repositories, interactive dashboards, and developer APIs.

The research team developed a data sharing decision framework that includes identifying the data sharing model, understanding applicable laws and risks, outlining the benefits, and weighing against costs. The research team also outlined the internal needs for data management structure and protocol:

- Establish data-focused staff or division
- Identify transit agency goals and objectives that can be accomplished through data analysis and data sharing
- Define data and analysis needs
- Define data sharing models and develop protocols, including risk assessment, privacy policies, and standard licenses

Prashanth Gururaja, Program Director, Shared Use Mobility Center

Prasanth Gururaja provided an overview of mobility partnerships and the most critical elements for success. Gururaja shared a decision-tree for addressing shared mobility data from a larger resource on [Data Sharing in Mobility Partnerships](#). Gururaja's organization, the Shared Use Mobility Center, fosters collaboration between public and private sectors to scale the benefits of shared mobility, conducts research, provides tools for the public to share policies and best practices, and provides technical assistance for shared-use pilot projects.

To successfully partner on shared data projects, agencies must identify the level of aggregation and partner agreeable to both parties. The level of aggregation may vary from partner to partner and project to project. What is most important is understanding the level of data required to produce the desired outcomes.

Equally as important as the data collection are the laws governing that data. Privacy laws vary state to state and agencies must be familiar with the regulations that may restrict or benefit their practices. Shared Use Mobility Center suggests that federal involvement could help encourage data management strategies, but for now agencies must rely on exemption requests or working to change the state level laws. These interventions can take time, so agencies should rely on a structured and systematic approach to developing their data sharing projects.

Key themes:

- Outcomes from data sharing should be identified in order to determine the best type of data model
- A structured decision making framework should be utilized when setting up data sharing agreements

Afternoon Workshop

Jim Rettew from CivicMakers kicked off the afternoon workshop session by asking participants to identify discussion points in two categories: Technology/Logistics and Governance/Policy. The discussion points identified are listed below by topic area.

Logistics Topics

- Staff capacity and capability in data engineering/analysis
- Data methodology and use identification
- Trust with private sector
- Third party repository
- Who can store granular data?
- Are data sharing agreements needed between public agencies?
- Ensuring agreements with the private sector provide the public sector the data requested.

Technology Topics

- What data do we have?
- Barriers to sharing data between agencies
- Ease of ability to use the data
- Missing connecting data
- Accuracy of data

Governance/Policy Topics

- Innovate contracting and procurement codes to be more nimble
- How to build public trust
- How to build vendor/service provider trust
- Who owns what process?
- What is the appropriate level of government?
- To whom should the legislature delegate authority?
- Should we create a third party data repository?
- Understand what the laws really say
 - Privacy
 - Data practices
 - Freedom of Information
- Avoiding pre-emption by the state; who decides?
- Local leverage for state laws
 - Funding
- Overcoming barriers to equity challenges when collecting data

In addition to identifying these topics, the group identified unanswered questions:

- Is it possible for a region to put together a protocol for data sharing, use, etc.?
- Can we create a work plan today and if not, what's missing or what do we need to determine? Who else needs to be involved?
- Should this just be about shared mobility? What about transportation data as a whole? Does this limit anyone?

Using the discussion items identified in the first exercise, participants brainstormed action items and mapped them on a scale of impact and effort. Action items were also sorted based on whether they were immediate or future items. Participants split into two working sessions to detail next steps, including barriers and pain points, for the most immediate action items.

Technology/Logistics Group

From these discussion topics, four proposed actions were identified as priorities to discuss as a small group:

1. Connecting practitioners
2. Agree on data sharing procedures
3. Build staff capacity and skills
4. Identify data uses and overall goal

Detailed next steps, questions, and barriers for each of these actions are bulleted below.

1. Connecting Practitioners

- Define – who is a practitioner and what are their roles
- Identify existing capacity and future capacity needs
- Mutual Benefits
- What does connecting mean?
- Can we solve today?
 - Need to hold a meeting of key stakeholders to answer these questions
- What do we need to create work plan?
 - Audience
 - Who and what
 - Understand capacity and skills
 - People at frontlines with data
 - Mobility hackathons
 - Org/regional focus
 - Understand broader goals/outcomes for transportation modes

2. Agree on Data Sharing Procedures

- Make data regularly available across jurisdictions
 - What's the model? Automatic or scheduled data deliveries?
 - Open data or APIs?
 - How to make data available for a meaningful purpose?
 - Consistency needs to be created – local regional standards, tools, techniques

- Needs to be a stand-alone priority, not part of a one-time project
- Identify existing methods
 - Tools and platforms
 - Storage
 - Capacity
 - Procedures
- Data catalog – what information are we missing? Do we even know?
- Can we solve today?
 - Develop recommendations
- Who else needs to be involved?
 - Policy makers => Open sharing gateway
 - MN data task force
- What do we need to create work plan/decision?
 - Data catalogue (internal & external)
 - Common language for use
 - Open source
 - Good metadata
 - Policies
 - Security measures to ensure proper use and privacy
 - Identify risks to a breach
 - Examine existing models (e.g. hierarchy sharing)
- Risks or obstacles
 - Barriers to entry are high
 - Misinterpretation
 - Staff time
 - Sustainability
 - How to share regularly and efficiently?
 - Lack of common language in org
 - Centerline segments, projections, etc.

3. Build Staff Capacity and Skills

- Skill needs are multidisciplinary
 - End to end skill set: data engineering to data visualization
- How to balance required skills vs. capacity
 - Share skills and funds to pool resources
 - Agree on procedures and scope
 - Connecting practitioners
- Data management methods need to be established (city or regional)
- Can we solve today?
 - Can create recommendations for capacity/skills
- What do we need to create work plan/decision?
 - Example funding models
 - University
 - Other regions, cities
 - Examples for data management
 - Identify education needs

4. Identify data uses and overall goal

- Questions to answer:
 - What are we missing/not considering?
 - What harms do we want to prevent?
 - What are other best practices?
 - How to define valuable data?
 - What are reusable privacy controls?
- Can we solve today?
 - No
- What do we need to create work plan/decision?
 - Tie data sharing practices to organizational and regional goals/plans
 - Establish a data management plan
 - Understand limitations of the data
 - Establish a sharing context – make it clear what the data can and cannot be used for
 - Build public trust
 - Safety
 - Cost savings
 - Positive benefits
 - Transparency
- Risks or obstacles
 - Lack of trust
 - Misuse
 - Inappropriate access/use or unintended
 - Lack of coordination/requirements for data between agencies/providers

Governance/Policy Group

From these discussion topics, three proposed actions were identified as priorities to discuss as a small group:

1. Understand the current laws to identify any gaps
2. Build public trust
3. Overcoming barriers for equity challenges

Detailed next steps, questions, and barriers for each of these actions are bulleted below.

1. Understand current laws to identify any gaps

- Definitions to understand
 - Constitutional Law
 - Tennessee Law
 - Incidental data
 - Data Practices Act - All data collected is public unless legislature notes otherwise
- Pain Points
 - What laws apply and what do they say? (confusion)
 - Situational differences: incidental vs. user specific

- Collection methods: direct or indirect
- Educating public on the laws and their application
- Solutions
 - Audit of current laws to provide context (first priority)
 - Research how Minnesota fits with other locations
 - Lean on legislative staff, University experts to conduct audits
 - Use City of Minneapolis staff to provide guidance on practicality
 - Explore including a chapter in Data Practices Act for mobility data (second priority)
 - Establishing public trust will be required

2. Build Public Trust

- Definition: Acquire data freely
- Pain Points
 - Backlash or public overreaction
 - Users might want more data
 - Skepticism
 - Middle man collection and handling
- Solutions
 - Share success stories and best practices
 - Verify data limits through regulation and education
 - Define benefits to demonstrate weight over costs
 - Utilize informed consent
 - Transparency
 - Improve security to eliminate skepticism
 - Collect the least amount of data needed (minimal viable data)

3. Overcoming Barriers for Equity Challenges

- Definition: Use data to demand equitable delivery of services
 - Selectively serving populations will skew data
 - Does privacy apply differently to different users?
 - Do providers need to serve entire city?
 - Data might give government leverage to demand more equitable treatment
- Pain points
 - Data is not representative of underserved areas – there is a danger to using misrepresentative data
 - Privacy goals may prevent us from being in a position to treat the service as a utility
 - Profit based decisions may not align with equity goals
 - Need to monitor equitable delivery
 - Not all areas may want the same services
 - How to measure demand and performance
- Solutions
 - If treated as a utility, require service everywhere and provide lower cost options
 - Help companies build demand for services if they align with city/regional goals
 - Improve demographic data collection to understand user service areas
 - Get more data of non-fixed uses
 - Regulation
 - Balance between regulation and incentives

- Start with areas of high density and high profitability
- Does data allow us to measure equity
- Measure equity of delivery and coordinating other municipalities
- Promote equity goals