Objective-Driven Data Sharing for Transit Agencies in Mobility Partnerships

White Paper

SHARED-USE MOBILITY CENTER

MOD
Mobility on Demand

Acknowledgments

July 2019,
Updated December 2024

Shared Use Mobility Center
Chicago, IL
312.448.8083
Los Angeles, CA
818.489.8651
www.sharedusemobilitycenter.org

SHARED-USE
MOBILITY CENTER

Contents

Overview

With advanced technologies enabling the growth of new mobility services, transit agencies and cities have looked toward partnerships with private mobility service providers for implementing Mobility on Demand (MOD) solutions to enhance their existing public transit systems. A primary challenge in implementing MOD solutions has been reaching an agreement between the public and private partners over data sharing. Drawing on lessons learned from the Federal Transit Administration's (FTA) MOD Sandbox program and beyond, this paper aims to support the decision-making process of transit agencies that are considering deployment of MOD or similar integrated mobility solutions in partnership with private-sector mobility service providers.

The paper first outlines the type of information an agency might need, based on their project type and objectives. This discussion helps guide the data request, whether it involves historical data for MOD service planning, trip information for MOD service evaluation, and/or real-time data for multimodal trip planning and payment apps.

The paper then discusses the challenges that local and regional agencies have faced in attempting to obtain the data that they might need. Challenges include those from both the public and private partners:

- Traveler Privacy
- Competitiveness
- Public Records Laws
- Data Security
- Aggregation
- National Transit Database (NTD) and Performance-Based Funding
- Capability Limitations

Based on observations from both inside and outside the Sandbox, a variety of possible approaches are available to agencies to address these challenges, and subsequently obtain and analyze the data that are necessary to meet project goals. Those approaches include project-level, regulatory, and legislative means to overcome the specific barriers mentioned above. In certain circumstances, the type of project will dictate which approach an agency can take, but the report also engages ways that agencies might better lay the groundwork for better data collection and analysis, in general. Approaches include:

- Finding a mutually agreeable parameter set and aggregation level
- Using third-party repositories
- Working with legislatures to modernize public records laws
- Establishing open data requirements
- Reaching individual Application Programming Interface (API) agreements with mobility service providers

The paper includes a decision tree that assists agencies in sequential decision-making to determine the best approach, based on their project type, project objectives, and constraints a synthesis of the lessons learned from both the MOD Sandbox and beyond, as well as proposed next steps to complement the FTA's ongoing work to establish metrics for shared mobility.

1. Purpose

One of the primary challenges of implementing MOD projects has been reaching an agreement between public and private partners over data sharing. Specifically, transit agencies and mobility service providers have had difficulty agreeing on the type of data to be shared and the frequency of that data sharing (e.g., real-time, daily, weekly, monthly). The Shared-Use Mobility Center (SUMC) has observed that this challenge has resulted primarily from each stakeholder having particular reasons for requesting or withholding data. For example, agencies may feel entitled to data if they are subsidizing trips. Mobility service providers, on the other hand, may be reluctant to share granular data due to concerns about traveler privacy and releasing proprietary information.

This paper is developed to assist transit agencies in deploying MOD solutions with mobility service providers through objective-based data acquisition and data sharing strategies that are further informed by performance measurement methodologies that help agencies understand the impacts of their projects. While understanding that specific data requirements will vary according to the nature of a project, this document aims to provide high-level guidance about how to develop an approach to obtaining this information.

In its best practices section, the Transportation Research Board (TRB) publication *Legal Considerations in Relationships Between Transit Agencies and Ridesourcing Service Providers* proposes that agencies understand data needs, specify data needs, maintain flexibility for data usage, ensure data security, and consider leverage in bargaining for data rights.[1] The following discussion will expand on these concepts, identify ways that agencies might overcome barriers to a data agreement, and chart a workflow for interested stakeholders through synthesis and a proposed decision tree. The analyses and discussions in this document are conducted by SUMC to support FTA's MOD Sandbox program.[2]

2. Know What Information You Need and Why

In the typical project planning process, a transit agency sets goals and identifies the target population for the service. Traditionally, fulfilling the data collection and reporting requirements for the National Transit Database (NTD) as well as for local and regional planning processes dictate the data needs.[3]

---

[1] See pages 109-110 National Academies of Sciences, Engineering, and Medicine 2018. Legal Considerations in Relationships Between Transit Agencies and Ridesourcing Service Providers (Legal Research Digest 53). Washington, DC: The National Academies Press. https://doi.org/10.17226/25109

[2] SUMC Innovation and Knowledge Accelerator. https://sharedusemobilitycenter.org/ika/

[3] National Transit Database homepage: https://www.transit.dot.gov/ntd.

Transit agencies should think through a series of questions that link their data requests to their MOD project's objectives.

- What do you need to learn?
- What is the objective of your project?
- What type of project is it?
- What types of analysis are required? Why? What types of data would be needed?

What do you need to learn? What is the objective of your project?

The data needed will be determined largely by the objectives and model of their project. In order to follow a structured approach, the agency should first determine the specific objectives of their particular project, then determine the factors that contribute to meeting those objectives that require information. The agency should also set a marker to identify which populations the project intends to serve. Finally, the agency should ask itself if the project addresses operational deficiencies and ineffectiveness in its service area, and if so, how the potential improvements will be measured.[4]

Once the objectives are determined, they should guide the agency's approach to identifying the following desired elements:

- Service outcomes
- Impact of the service
- Information needed to determine whether the service outcomes are achieved
- Activities, such as surveys and analyses, that are needed to determine the necessary information
- Data types and elements that need to be collected or analyzed to carry out the identified analyses

What type of project is it?

The type of project influences the type of data an agency will need to plan, implement, and assess the effectiveness of a project. For example, a ridesourcing partnership might require data about trips taken, or a multimodal trip-planning app might require access to real-time vehicle availability. Typical MOD pilot projects, such as those in the MOD Sandbox program, include first/last-mile connections to transit, late-night service, on-demand service for people with disabilities, and multimodal trip-planning and payment integration apps.

---

[4] McCoy, K., Andrew, J., Glynn, R., & Lyons, W. (2018). Integrating Shared Mobility into Multimodal Transportation Planning: Improving Regional Performance to Meet Public Goals (No. DOT-VNTSC-FHWA-18-13; FHWA-HEP-18-033). United States. Federal Highway Administration. Office of Planning, Environment, and Realty. https://rosap.ntl.bts.gov/view/dot/34689

What types of analyses are required? Why? What type of data would be needed?

In addition to the types of data, the timing and frequency of the data varies by need. Project data needs typically include those described below.

Transportation Trends – Archived, Post-Processed Data

The agency will want both its own data and, if possible, some of the historical data of demand responsive services, such as taxi or ridesourcing services (e.g. Uber, Lyft). This historical and aggregated data ideally shows travel patterns and pick-up/drop-off locations.

Service Planning – Archived, Post-Processed Data

Here again, the agency will look at factors such as response time and service area to determine operational characteristics of the project, such as geographic area, times of operation, and population served.

Service Delivery, Effectiveness, Accountability – Trip Data

This is the type of data needed for both on-going and retrospective evaluation.

- Operations data: The purpose of these data is to understand the performance of the service and behavior of users. These are trip-level data, analyzed in real time or periodically during the pilot, depending on the type of partnership. Types of data can include:

  » Origin-destination (O/D) pairs

  » Pickup and drop-off times

  » Wait times

  » Travel times

  » Vehicle occupancy

- Accounting data: The purpose of these data is to monitor the actual cost compared to the budget, and comprises trip and cost data analyzed post-hoc. Trip-level data include items such as traveler fares and total cost. The program-level aggregated data can examine parameters such as surge pricing trends from transportation network companies (TNCs), average fares, and pooled versus non-pooled rides, i.e., passengers-per-vehicle when applicable.

- Auditing data: The purpose of these data is to determine if the partner is providing the service as agreed. These include the accounting data as well as compliance data, such as fulfillment of wheelchair requests.

Other models might require a modification of the specific data required. The items mentioned are only examples of information that would be useful for agencies. Each agency should determine the specific data items needed based on the objectives of the project.

Integrated Trip-Planning and Payment: Real-Time Data

For integrated multimodal trip-planning, operators can provide data through an Application Programming Interface (API), which allows access to a real-time feed of an operator's trip availability, vehicle location, cost, and other basic and pertinent data.

For transit, the General Transit Feed Specification (GTFS) is the current standard. For bikeshare, it is the General Bikeshare Feed Specification (GBFS). Both GTFS and GBFS enable a private operator to better integrate public transit trips into their users' trip planning. While a transit agency could access the information on a ride, vehicle, and cost (bikeshare, carshare, TNCs, etc.), some companies do not allow their data to be shared on the same app as a competitor. Furthermore, in the United States, there is currently no broad precedent for open APIs, where providers must make real-time parameters available for third-party use.

Improving prediction of arrivals can be an area of data sharing as well. GTFS with real-time inputs, such as vehicle location, exist for transit, but typically not in combination with more sophisticated analytics based on historical arrival and departure data to more accurately predict real-time arrivals. Private mobility ventures are developing these analytics and computing-intensive software to maximize utilization of their services. Although public agencies have both additional mandates for service provisions and a longer funding horizon, both sectors might learn from each other. When coordinated together, real-time prediction of transit services can be improved, leading to a higher-quality experience for the traveler.[5]

3. Concerns and Challenges to Receiving Quality Data and Information

Through conversations and interviews with various agencies and providers in the MOD Sandbox program and experts in the mobility ecosystem, along with supplemental research, several challenges around data agreements were identified. These challenges exist for both the public agency and private mobility service providers, but some are of more concern to one sector than the other. The following discussions summarize the identified challenges and some other key findings.

Privacy Concerns with Location Data

Both the public and private partners are concerned that origins and destinations for point-to-point data might compromise privacy. Paired origin-destination data, even if anonymized, is at significant risk of identifying individuals when combined with publicly available data sets, such as property tax records. A recent study found that even coarse aggregation of anonymized data may put privacy at risk.[6] This emphasizes the need for security, and leads to the consideration about the accessibility of records from public agencies.

---

[5] Sun, F., Pan, Y., White, J., & Dubey, A. (2016, May). Real-time and predictive analytics for smart public transportation decision support system. In Smart Computing (SMARTCOMP), 2016 IEEE International Conference on (pp. 1-8). IEEE.

[6] De Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. Scientific reports, 3, 1376

Competitiveness

Data generated from partnerships may also contain proprietary information of the private mobility service provider. For example, the provider would have an interest to not share data pertaining to algorithms, predictive features, or data analytics methods. Commercial information such as targeted marketing, marketing strategies, and market capture are also valuable information that may be of proprietary interest or considered to be business sensitive information. The possibility of broader release of this information leads to the following concern over public records laws.

Public Records Laws and Related Regulations

The risks associated with privacy and competitiveness issues are largely due to public records laws (also referred to as freedom of information laws), which allow the public to gain access to government records. These laws predate the modern information age, and as such, do not explicitly address the technological advancements that have allowed for gathering and storing large amounts of granular data. Transportation authorities and transit agencies collect a large amount of personally identifiable information (PII), and any agency should realize and understand the risk in their potential exposure.[7]

These laws often list types of records that are exempt from public release, but exemptions for data privacy and security vary by state and jurisdiction.[8] It should be noted that non-disclosure agreements or contract provisions might not supersede laws regarding disclosure.[9]

Also, the service type determines what data an agency might reasonably expect. Where partnership-based services such as subsidized first/last-mile (FLM) trips may require data about those trips, general TNC operations might fall under regular permitting regimes, where operators are often not required to submit data to local jurisdictions. A prominent example can be found in San Francisco, which under California regulations does not regularly receive data from the California Public Utilities Commission, the state-level regulating agency.[10] Before implementation, agencies should refer to their local and state laws to determine how shared mobility modes are regulated in their jurisdictions. Many states have preemptive laws on the books, which might limit a local jurisdiction's ability to regulate or require data from certain mobility service providers operating in its area.[11]

[7] Thomas, L. W. (2017). Legal Issues Concerning Transit Agency Use of Electronic Customer Data (No. Project J-05, Topic 16-02). http://nap.edu/24730

[8] "State Freedom of Information Laws," National Freedom of Information Coalition. https://www.nfoic.org/coalitions/state-foi-resources/state-freedom-of-information-laws. Also, see National Conference of State Legislatures. Data Security Database: https://www.ncsl.org/technology-and-communication/data-security-laws-private-sector

[9] RASIER-DC, LLC v. B & L SERVICE, 237 So. 3d 374 (Fla. Dist. Ct. App. 2018). See discussion at Chow & Mingrove "Upon Reflection, Some Trade Secrets May Be Less Secret Than They Appear." https://blogs.orrick.com/trade-secrets-watch/2018/01/25/upon-reflection-some-trade-secrets-may-be-less-secret-than-they-appear/#more-2410

[10] "TNCs Today: A Profile of San Francisco Transportation Network Company Activity." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/overview/tncs-today-a-profile-of-san-francisco-transportation-network-company-activity-san-francisco-california-2017/

[11] National League of Cities "City Rights in an Era of Preemption: A State-by-State Analysis." https://www.nlc.org/resource/city-rights-in-an-era-of-preemption-a-state-by-state-analysis

Data Security

Whether the source is a partner or a third party, the data must be handled and stored appropriately. The primary concern on all sides is data management and security. This could affect collection, formatting, storage, archiving, access/authorization protocols, cybersecurity, reporting (internal and external), and compatibility with existing systems. Data handling and security protocols must be discussed thoroughly before forming partnerships and clearly defined in agreements, while adhering to the regulations regarding data and privacy at the local, state, and federal levels.

Data Aggregation

Data could be aggregated spatially and/or temporally to address the concerns discussed above, but the resulting resolution may be too coarse to be useful for the agency to monitor the service's performance relative to its metrics and objectives. Similarly, data that is not aggregated could still be reported without a level of resolution needed to meet agency objectives.

National Transit Database (NTD) and Performance-Based Funding

Each year, NTD performance data are used to apportion FTA funds to transit agencies in urbanized areas.[12] Agencies desire accurate reporting of trips, including for example, subsidized first/last-mile trips, to be able maximize their FTA formula funds.

Capability Constraints

Transit agencies may not have the technical expertise or infrastructure to handle, store, and analyze large amounts of data. This may be true even for pilot projects, depending on the characteristics of the agency and the breadth of the pilot service.

4. Obtaining the Data or Information You Need

Based on observations from inside and outside the MOD Sandbox perspectives, a variety of possible approaches are available to agencies to obtain and analyze the data necessary to meet project goals. Those approaches include project-level, regulatory, and legislative means depending on the specific barriers that the agency faces. The type of project will also affect which approach it should take.

Project-Level Data

Subsidized Service Projects

Depending where the agency operates, the regulatory regimes discussed below might affect the flexibility afforded to the agency as it negotiates data and other requirements for the project. Many aspects need to be negotiated between the agency and the operator. The private and public partners should try to build a data sharing agreement based on mutually agreeable parameters and

---

[12] US Federal Transit Administration. What is the National Transit Database (NTD) Program? https://www.transit.dot.gov/ntd#:~:text=The% 20NTD%20records%20the%20financial,provide%20public%20information%20and%20statistics

aggregation levels, focused around what data the agency deems essential for planning and evaluation purposes. These data often include:

- Trip length/distance
- Origin-destination pairs: The grantees found this to be one of the most difficult negotiations. Negotiations usually settle on which level of aggregation to use, such as ZIP Code, census Tract, or number of decimal digits of latitude-longitude coordinates.
- Cost/fare
- Customer response: Most app-based vendors gather customer satisfaction information and may be willing to integrate specific questions for the transit agency on subsidized rides.

The frequency of the reports is another important element. While a trip planning app may require real-time information, service projects may only need information about trips or customer feedback weekly or monthly.

As discussed above, data anonymization requires some care, and local agencies might want to consider measures to assist in these efforts. The following are a few examples to demonstrate the different parameters and aggregation levels used by different partnerships within and outside of the MOD Sandbox projects.

Pierce Transit (Washington) - Once the regulatory environment is determined—again, public records laws might preempt agreements—the agency or operator will often consider some form of non-disclosure agreement.[13] The first/last-mile MOD Sandbox project agreement between Pierce Transit and their partner Lyft involves monthly reports from Lyft to Pierce Transit with data for each trip as follows:[14]

- Month/Year
- Day of Travel
- Dispatch Method (Concierge or App-based)
- Coupon Code (to indicate the zone)
- Origin Census Tract and Destination Census Tract
- Trip Time Period (AM Peak, Midday, PM Peak)
- Trip Length Range (2-4 miles, 4-6 miles, etc.)
- Trip Duration Range (5-10 minutes, 10-20 minutes, etc.)
- Trip Cost, Refund, and Subsidy

[13] Matthew W. Daus (2018) "Transportation Network Companies: Passenger Data Security and Privacy Issues." Westlaw, Sharing Economy 300:100. See also Cubic (2014) "Privacy and Data Collection in the Transportation Industry"

[14] Correspondence with Pierce Transit. For more on the project, see: https://www.transit.dot.gov/research-innovation/mobility-demand-sandbox-demonstration-pierce-transit-limited-access-connections

Lyft and Pierce Transit are working under the presumption that these data could be released in a public records request, and the parameters and aggregation levels are determined accordingly.

City of Arlington (Texas) - The partnership between the City of Arlington, Texas and Via yielded a trove of publicly available documents from the issuance of the request for proposal through implementation. The partners are making efforts to anonymize the data, as discussed below. SUMC's analysis of their project shows how the City was able to work towards mutually agreeable terms through an iterative procurement process.[15]

These data will include, on a periodic basis:

- *Individual ride data (anonymized): requested origin, requested destination, number of passengers, time and length of ride, fare paid*

- *Aggregated service data (for a given period): completed rides, active drivers, driver hours, utilization (rides per vehicle per hour), average trip duration (minutes)*

- *Performance standards (for a given period): average estimated time of arrival (ETA) to pick-up, % of on-time rides, % completed rides, rider satisfaction metrics*

- *Historical trends (over longer periods): overall ride volume/growth, top requested origins and destinations, demand 'heat maps'*

These data shall be made available in formatted numerical and graphical reports. For the avoidance of doubt, the information above constitutes proprietary trade secrets of Contractor, and shall be subject to the confidentiality obligations set forth in the Contract.

MBTA (Boston, Massachusetts) - The Massachusetts Bay Transportation Authority's RIDE On-Demand program for people with disabilities formed agreements with Uber and Lyft that are similar to City of Arlington's agreement with Via, although it requires the monthly data to be reported on a ZIP code-level resolution.[16]

LA Metro (California) - In Los Angeles County, the Los Angeles County Metropolitan Transportation Authority has signed an agreement with Via to provide first/last-mile service in three stations around the county, as part of the agency's MOD Sandbox project. The service provides rides to transit stations from "virtual" bus stops located near the trip requester as determined by Via's routing algorithm. The data sharing agreement between the two is comprehensive, with a detailed list of parameters for individual trips and weekly aggregations of app, call center, vehicle, and ridership information.

---

[15] "The City of Arlington, TX and Via MOD P3 RFP & Contract Process." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/casestudy/new-the-city-of-arlington-tx-and-via-mod-p3-rfp-contract-process/

[16] "Lyft and Uber Agreements, MBTA, Boston, Massachusetts, 2016." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/overview/lyft-and-uber-agreements-mbta-boston-2016/

Among the many trip-level parameters listed in the agreement between LA Metro and Via are:[17]

- Requested pick-up and drop-off locations – latitude and longitude to three decimal places
- Actual wait times before pick-up, and actual pick-up and drop-off date and time
- Whether an accessible vehicle ride was requested and provided
- Whether there was more than one passenger in the car
- Trip cost charged to paying passenger
- Ride rating awarded by passenger (1-5 starts)

These examples highlight that different agencies are willing to agree to different parameters and aggregation levels of the data produced by their partnership. These depend on the project type, the public disclosure laws in the agency's jurisdiction, the agency's capacity to manage the data, the urgency to introduce service in the target areas and populations, and the partner's business interests. Thus, a single best practice has yet to emerge and may not even be useful. Agencies should understand what potential providers' approaches to data sharing are and aim to select a partner with whom they can find a mutually agreeable parameter set and aggregation level for data sharing.

Agencies should understand what potential providers' approaches to data sharing are and aim to select a partner with whom they can find a mutually agreeable parameter set and aggregation level for data sharing.

Data Repositories and Collaborations

Another approach to sharing data from mobility service providers while avoiding some of the aforementioned challenges is for the public agency to never hold the mobility service provider's disaggregated data. Instead, such data would be held by the provider or a trusted third-party repository under strong legal, cybersecurity, privacy, and technical protections such as encryption, anonymization, data lakes, etc. Those types of repositories would also require a clear access agreement, through which the public agency would be guaranteed access to "views" of the data at the levels needed for accomplishing its work such as service planning, performance measurement, annual reporting, etc.

One third-party model is currently being developed at the University of Washington Transportation Data Collaborative (TDC).[18] This model combines an understanding of how to balance the needs of public and private stakeholders with the institutional heft and technical capability of a major research institution. Capability includes both the technical know-how and the sheer storage and computing power needed to handle massive and ever-growing data flows.

---

[17] Los Angeles County Metropolitan Transit authority. "Contract for Federal First and Last Mile Mobility On-Demand Services," FAIN No. CA-2017-018, Exhibit A – Statement of Work. December 18, 2018

[18] University of Washington Transportation Data Collaborative: https://www.uwtdc.org.

TDC is working to develop the policy environment, governance model, and security protocols so that they can serve as a public/private data management turnkey solution for agencies facing constraints related to data disclosure or capability to perform large-scale analysis. In the case of Washington State, the TDC aims to use research exemptions in the State's public records laws to protect sensitive data from disclosure.

However, the TDC is experiencing challenges. First, the legal constraints of Washington's public records laws require that the setup of this type of institution be vetted through the state attorney general's office, which has added significant time to launch due to availability of resources. The TDC is also required to have an agreement with the private mobility service provider prior to generation of the data.[19]

Business interests are also a factor. Unless there is a requirement to report granular data, there will not be much incentive for a private mobility service provider to participate in a third-party repository. Submitting data requires resources from the business end, but without financial or other incentives, there is a limited business case to prioritize it. Further still, the University of Washington also requires that businesses pay to use the TDC, not just the public-sector clients, creating an additional barrier for its use. These fees may be marginal for a global company, such as a major TNC, but can be a significant burden for smaller providers.

A similar effort is being conducted by Shared Streets, supported by National Association of County Transportation Officials (NACTO) and the Open Transport Partnership (a Bloomberg Philanthropies-funded initiative). Shared Streets aggregates general operational data from TNCs and other anonymized transportation data to include in a repository for analyzing and understanding transportation behavior in cities. Shared Streets is a partner on pilot partnerships including the Department of For-Hire Vehicles and the Department of Transportation in Washington, DC to anonymize and aggregate taxi and TNC trip activity, and with the Seattle Department of Transportation and Ford to map and regulate their curb use. Other partnerships are planned around traffic speeds and incidents, construction, and closure data.[20] However, as of this writing, Shared Streets has not been used to share data in a partnership service, such as a subsidized first/last-mile service, but this kind of tool presents an opportunity for transit agencies to better understand travel patterns for planning and implementing new projects. It remains to be seen whether the tool can or will be used effectively for operational data of a specific service partnership.

The National Renewable Energy Laboratory's Transportation Secure Data Center is a partnership between the U.S. Department of Transportation and the U.S. Department of Energy that contains archived data that may be useful for planning.[21]

While the formation of these third-party repositories could possibly be a longer-term solution, the incentives between public and private sectors are not aligned enough to offer a reliable solution. While a research exemption to public record laws is available in the state of Washington, it

---

[19] Interview with University of Washington

[20] Shared Streets. http://sharedstreets.io/

[21] Transportation Secure Data Center home: https://www.nrel.gov/transportation/secure-transportation-data/tsdc-about.html

is not entirely resolved that a publicly funded university would not be subject to public disclosure laws, and this could vary by jurisdiction. Furthermore, agencies should consider whether the appropriate level of detail would be available, and if the resources dedicated by the mobility service provider through an arrangement with such an entity would meet their objectives. If constraints related to public records disclosures or agency capability are impeding progress to obtaining the data that they need, agencies should explore using a third-party repository, provided that the information that will be made available is supportive of agency objectives. The examples highlighted, along with the emergence of mobility data management companies, show that a third-party solution may be more useful in the project planning stage, where general travel patterns are of interest, as opposed to the operations stage, where more granular trip information is of interest.

While there are benefits to agencies not directly holding and managing the data, the data themselves might be of significant value to generate revenue or exchange for private sector data. The Transit Cooperative Research Program has initiated a research effort to assist public transit agencies on how to better understand the potential uses and value of data generated by public transportation and to be prepared to enter into agreements with entities to trade, share, and sell data.[23]

Integrated Trip-Planning Apps

Transit agencies are increasingly interested in becoming mobility managers for their regions, and some have embarked on developing multimodal trip-planning apps, where travelers can theoretically plan, book, and pay for multimodal trips that consider all participating mobility service providers operating in their area. Agencies have taken different approaches to achieve this goal.

TriMet (Portland, Oregon) - MOD Sandbox grantee TriMet is developing an open-source trip-planning app that aims to integrate bike sharing and TNCs into its public transit trip-planning options. The agency needed to balance the concerns of Uber and their proprietary, real-time feed, with the needs of the TriMet app to provide accurate information. The terms of use for major TNCs' APIs preclude the appearance of that data on the same screen with a competing service, even though these same data are available freely on Uber's own app. This presents a barrier to true mobility integration. However, for the MOD Sandbox project, TriMet and Uber agreed that TriMet could use Uber's fare and arrival time estimates in their trip-planning app.[23] However, if a competing TNC, such as Lyft, were to be included on the same screen within the app, say as a side-by-side comparison, then only the arrival time information could be used. TriMet also agreed to consider these data confidential information and protect it from public disclosure, including notifying Uber of public records requests involving the data. As of March 2019, TriMet had not reached a similar agreement with Lyft, and Lyft was not included in the public demo release of TriMet's trip planner.[24]

[22] Transit Cooperative Research Program J-11/Task 31. https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4445

[23] Uber-TriMet API agreement for MOD Sandbox project, February 2019.

[24] TriMet MOD Open Trip Planner Demo. https://betaplanner.trimet.org

Other Agencies (Valley Metro, DART, VTrans) - Other agencies in the MOD Sandbox who are creating integrated trip-planning apps have been in similar conversations with the major TNCs about use of their API data. Valley Metro (Phoenix, Arizona) is working to integrate Lyft rides into its integrated trip planning and payment app. It has delegated agreements to use TNC API data to its app developer, Routematch, in part to avoid possessing the data, which might then be disclosed in a public records request.

Dallas Area Rapid Transit (DART) in Texas has been negotiating with the major TNCs and other providers for first/last-mile discoverability in their trip-planning app, similar to TriMet. In DART's microtransit service zones, DART and Uber have negotiated an agreement to mutually discover on-demand shared rides through both DART's microtransit service (GoLink) and UberPool on each other's apps, with the UberPool rides being subject to the GoLink fare structure. As a first step, UberPool trips are discoverable through the DART GoPass app and can be booked through a deep link out to the Uber app.[25]

While individual agencies that are considering integrated trip-planning apps can attempt one-off agreements with TNCs, this appears to be an inefficient approach for the industry in the long term. As these efforts unfold, the contrasting approaches that TriMet, Valley Metro, and DART are taking—relying on a commercial agreement or a direct contract/agreement with a TNC—will provide insight toward which approach may be more viable and produce a more successful service and which areas of policy need addressing if communities indeed see integrated trip-planning as a worthwhile goal. The difficulties arise naturally from a lack of a requirements or incentives for providers to make these data available, either to the government or the public. The MOD Sandbox program presents an opportunity to evaluate these approaches.

However, bike share operators and smaller mobility service providers have been more open with others' use of their API data. These services can help accelerate integrated multimodal trip planning in areas where they play a bigger role. For example, the Vermont Agency of Transportation (VTrans) is developing an open-source trip planner, Go! Vermont (https://plan.govermont. org/), and using the GTFS-Flex data standard to integrate flexible transportation providers across the state. Private mobility partners, such as demand-response transportation and carpool providers, have agreed to integrate their data using this standard. This creates opportunities in rural areas, where TNCs may not be a significant provider of trips.

Regulatory Data Environments

Local governments are typically constrained by state regulations, especially public records laws, that factor in many of the approaches discussed earlier. Also, while most states (41 as of this writing) have some form of preemption law related to ridesourcing services, interpretation of those laws for data might vary.[26] Through their legislative process, the local agency could work with the state legislature

---

[25] DART hires Uber to give Dallas-area customers free, discounted rides. https://www.dallasnews.com/business/technology/2019/03/12/dart-hires-uber-to-give-dallas-area-customers-free-discounted-rides/#:~:text=Starting%20today%2C%20Uber%20is%20offering,will%20be%20subsidized%20by%20DART.&text=Uber%20will%20provide%20free%20and,areas%20with%20limited%20transit%20options

[26] For FOIA/FOIL laws see National Freedom of Information. See National League of Cities (2018), "City Rights in an Era of Preemption: A State-by-State Analysis" at https://www.nlc.org/resource/city-rights-in-an-era-of-preemption-a-state-by-state-analysis

to strengthen public records laws to prevent disclosure of travel data. This typically happens through exempting such data from disclosure by defining certain parameters as personally identifiable information (PII), to protect privacy, or as trade secrets, to protect proprietary information. What this means in practice is that agencies have to choose between seeking exemptions—working with the states to change the law—or seeking innovative options such as finding a third-party data repository.

Exemptions in public records laws are being used a data protection tool by the joint first/last-mile MOD Sandbox projects undertaken by LA Metro in Los Angeles County and Sound Transit/King County Metro in the Seattle area.[27]

LA Metro (California) - LA Metro's agreement with Via, mentioned earlier, has specific provisions around which data constitute trade secrets, and that as a result, those data will not be disclosed to the public under the California Public Records Act. Furthermore, the agreement reinforces the California Streets and Highways Code provision that any information that identifies or describes a person, including travel pattern data, is PII and exempt from public disclosure.

Sound Transit and King County Metro (Washington) - The State of Washington statute, however, does not include an exemption for location data. It is interpreted to exempt social security numbers, driver's license numbers, and other "highly sensitive" information as personally identifiable information, potentially forcing the Seattle project to disclose location data generated by their MOD service. However, integration of the service with ORCA, the regional transit fare card, has provided a path for exempting location data, due to provisions in the statute relating to the use of fare media.

Transit agencies have also played a proactive role in helping to modernize state laws to exempt sensitive data from public access.

TriMet (Portland, Oregon) - In 2014, TriMet was developing an electronic fare collection system, Hop FastPass (currently operating), that would generate large amounts of travel pattern data. Yet, there were no protections against disclosure of such data from public records requests. TriMet supported the Oregon legislature to update laws to exempt travel pattern data in possession of transit agencies, citing the potential of use by stalkers and domestic violence offenders.[28] The legislation listed PII "collected as part of an electronic fare collection system of a mass transit system" as exempt from disclosure, and classified location and other travel information from these systems as PII. The legislation permitted disclosure of "public records that have attributes of anonymity that are sufficient, or that are aggregated into groupings that are broad enough, to ensure that persons cannot be identified by disclosure of the public records."[29] As of a discussion with TriMet in December 2017, no requests had been made that would prompt TriMet to determine whether certain travel data would not be exempt.

DART (Dallas, Texas) - In 2015, during the Texas Legislative Session, the North Texas Toll Authority sought to protect toll customers' data, including location data gathered at toll gantries, from disclosure.

[27] Sourced from interviews for Sandbox: California Streets and Highway Code, Section 31490 and Revised Code of Washington 42.56.330
28 Oregon State Legislature, Legislative Information for House Bill 4086. Protect the Privacy of Public Transit Riders. https://olis.oregonlegislature.gov/liz/2014R1/Downloads/CommitteeMeetingDocument/35111
[29] Oregon Revised Statutes Chapter 192, Section 192.345(38). https://www.oregonlegislature.gov/bills_laws/ors/ors192.html

DART requested that the bill's authors modify the legislation to include the same protections for transit users. Upon DART's request, protections were added for transit customer's electronic data collected through DART's mobile application (GoPass), and it was further modified to protect customer data collected by all applicable transit systems in Texas. Location is protected since it is part of the transaction activity.[30]

> **Agencies can positively influence the modernization of public records laws to protect sensitive travel data... and still retain data internally for evaluation, planning, and monitoring the performance of their services.**

In both the Oregon and Texas examples, the updates to the public records legislation were uncontroversial and were easily passed in their legislatures.

The key lesson is that agencies can positively influence the modernization of public records laws to protect sensitive travel data from public disclosure and still retain data internally for evaluation, planning, and monitoring the performance of their services. The original laws were written before advancements in technology enabled the rapid collection of real-time, granular, and voluminous personal travel data. The advancement in fare collection technology for traditional fixed-route services is enough of a reason to better protect personal data, and considerations for point-to-point travel data should also be incorporated. In the interest of transparency about how public funds are spent on transportation, the public deserves to understand the data that serve as the basis of an agency's policy and service decisions. However, the public's access to individuals' travel patterns has not been shown to be in the public interest, and protecting the privacy of individuals has drawn broad political support. Agencies can work with legislatures to include location data as well as other necessary elements to be classified as PII.

To overcome the lack of availability of API data for integrated trip-planning, transit agencies, together with states or cities, can establish permitting systems to require providers to open up basic data parameters needed for developing trip-planning apps. Many of these parameters are already available through providers' own apps. These requirements could enable side-by-side comparison of trip itineraries in coordination with transit, or otherwise make providers' APIs publicly accessible. The Finland Transport Codes, effective in 2018, present an example of how a country the size of many US regions has required open data and interoperability from public and private transportation providers with the goal of seamless multimodal travel and payment.[31] The City of Arlington, Virginia requires micromobility operators in that city to provide publicly accessible APIs regarding the

---

[30] Texas Transportation Code Chapter 451, Section 451.061. https://statutes.capitol.texas.gov/Docs/TN/htm/TN.451.htm#451.061. Also sourced from correspondence with T. Plesko, DART.

[31] "National Mobility as a Service (MaaS) Policy, Finland, 2018." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/overview/national-mobility-as-a-service-maas-policy-finland-2018/

location of micromobility devices (bicycles and scooters available for rental. As the MOD Sandbox trip-planning projects and the regulatory environments like Finland and Arlington, Virginia mature, understanding the effects of these approaches will help to further refine this recommendation. Some cities require data sharing from all mobility service providers operating in their jurisdictions. For example, Chicago requires data from TNCs regarding vehicles, drivers, trip origins and destinations, and crashes.[32]

Such requirements can be important tools for cities to evaluate the impacts of TNCs and develop policies accordingly regarding congestion, equity, safety, or other goals. Similarly, some cities, including Los Angeles, require micromobility operators to use the Mobility Data Specification (MDS), a common standard for reporting trip, cost, and other vehicle and system data to the city to help the city actively manage the providers and the public right-of-way. However, these industry-wide requirements are not necessarily suitable for the project-level data sharing needs discussed in this paper, which emerge from the particular objectives of specific MOD partnerships.[33]

Federal Data Requirements

The lack of clarity regarding data requirements for shared mobility partnerships—as acknowledged by a recent US General Accountability Office (GAO) report to the US Senate Committee on Banking, Housing, and Urban Affairs—results in some confusion over reporting requirements to the National Transit Database (NTD).

Recipients and beneficiaries (i.e. transit agencies) of FTA grants are required to report certain trip data (depending on mode) to the NTD. These data, in tandem with those needed for planning, shapes how local agencies view their core data-gathering responsibilities. While no data requirements exist for shared mobility trips, the GAO report shows that the FTA is moving towards expanding NTD reporting requirements.[34]

The federal government could also consider requiring partnerships have to clearly defined data management strategies to be eligible for federal funding. This could place a mutual responsibility between the public and private sector participants to develop a strategy and infrastructure for data sharing that could then benefit other partnerships.

> **The federal government could consider requiring partnerships to have clearly defined data management strategies to be eligible for federal funding.**

---

[32] National Mobility as a Service (MaaS) Policy, Finland, 2018." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/overview/national-mobility-as-a-service-maas-policy-finland-2018/ "Shared Mobility Devices – Transportation." And, Arlington, VA city website: https://www.arlingtonva.us/Government/Programs/Transportation/SMD-API

[33] Institute for Transportation and Development Policy. "Ride Fair: A Policy Framework for Managing Transportation Network Companies" https://www.itdp.org/wp-content/uploads/2019/03/2019.03.13.TNC-Policy.V9.pdf. "Mobility Data Specification: A Data Standard for Shared Mobility Providers, Los Angeles, California, 2018." SUMC MOD Learning Center: https://learn.sharedusemobilitycenter.org/overview/data-standard-for-shared-mobility-providers-los-angeles-california-2018/

[34] See pages 24; 25, footnote 44, re: reporting; and 37; US Government Accountability Office, 2018, "Public Transit Partnerships: Additional Information Needed to Clarify Data Reporting and Share Best Practices." https://www.gao.gov/assets/700/693518.pdf.

5. Synthesis and Next Steps

Based on their project objectives, project type, and metrics for performance, agencies can create a strategy for establishing a data agreement with their partner. This strategy first requires:

- Defining project objectives clearly
- Developing the performance indicators to measure the performance of the project and its outcomes
- Creating a list of data elements truly needed for measuring the outcomes of the project, considering potential constraints around aggregation levels and reporting frequency

Decision Tree

Then, agencies will need to decide from various pathways to form the structure and details of their data agreement. The decision tree shown on the next page is intended to guide agencies that are considering mobility partnerships in a systematic way.
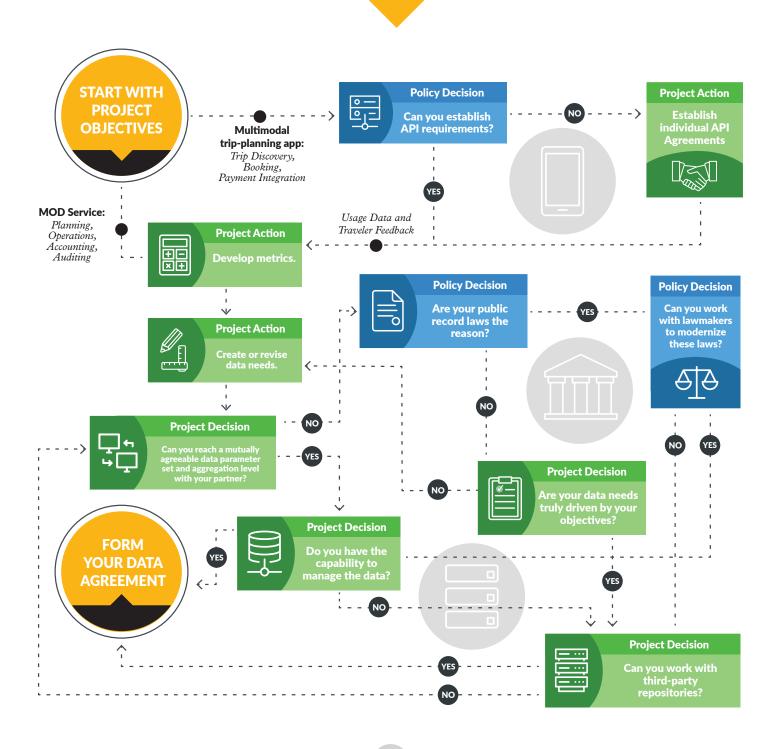
The decision tree sequentially directs the agency through a set of issues, with consideration of benefits and tradeoffs to inform important choices, such as:

- How to determine potential partners among institutions and private mobility service providers
- What is the best strategy for data handling and management, balancing considerations such as storage, sanitization, authorization, and security, including protocols for using third-party data repositories
- Identifying which policies might serve as barriers to collecting, using, and safeguarding data. Once identified, the agency can participate, as one of many stakeholders, in the legislative process to modernize those policies.

For example, if the agreement parameter and aggregation level are not initially achievable, the decision tree prompts the agency to examine whether its public records laws are the reason for this, and if so, encourages the agency to work with lawmakers to modernize those laws. If this is not feasible, then a third-party repository is suggested. Following this process will support the establishment of data agreements that respect individual privacy and proprietary information while forming integrated and multimodal service partnerships that serve the public's best interest.

Decision Tree

The orange circles indicate the starting and ending points for the decision process, beginning with the objectives agencies must determine before a data agreement can be reached. The green boxes represent actions and decisions on a project level, and the blue boxes represent decisions on a policy level.

START WITH PROJECT OBJECTIVES

Multimodal trip-planning app:
Trip Discovery, Booking, Payment Integration

MOD Service:
Planning, Operations, Accounting, Auditing

Policy Decision

Can you establish API requirements?

NO

Project Action

Establish individual API Agreements

YES

Usage Data and Traveler Feedback

Project Action

Develop metrics.

Project Action

Create or revise data needs.

Policy Decision

Are your public record laws the reason?

YES

Policy Decision

Can you work with lawmakers to modernize these laws?

NO

Project Decision

Can you reach a mutually agreeable data parameter set and aggregation level with your partner?

NO

YES

NO

Project Decision

Are your data needs truly driven by your objectives?

NO   **YES**

FORM YOUR DATA AGREEMENT

Project Decision

Do you have the capability to manage the data?

YES

NO

YES

Project Decision

Can you work with third-party repositories?

YES

NO

19

For each major decision presented in the decision tree, a set of advantages and disadvantages are listed in the table below to help an agency weigh their options. It is important to understand that there is no right answer, and the agency must manage tradeoffs based on their specific and desired values, capabilities, outcomes, and project objectives.

Advantages and Disadvantages of Decisions

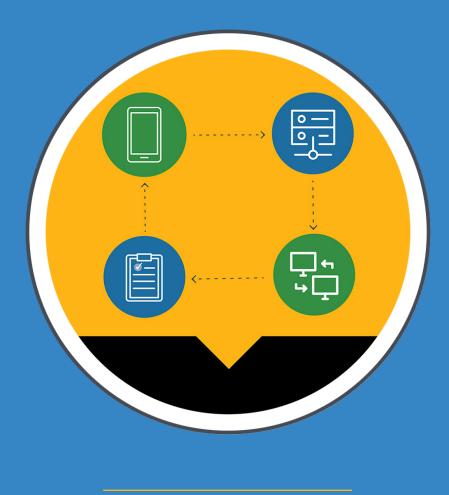| Project Type	Decision	Advantages	Disadvantages
MOD Service	**Mutually agreeable parameter set and aggregation level**	• Direct negotiation with provider to address concerns and data delivery format • Addresses specific needs for project and methods for data ownership and handling	• Data resolution may be too coarse • Lack of standard requires one-off agreements for every project, including extensions • Differences may not be reconcilable in the end
	Modernize public records laws	• Clarity on what data are protected • Establishes bounds for future projects/agreements	• May be a longer-term solution • Requires time and resources for legislative outreach and cross-departmental coordination
	Manage data in-house	• Data ownership sits with agency; can query as needed • Additional benefits for agency; can use for planning • Possible monetization in the future	• Data handling expertise required • IT infrastructure may be needed • More burdensome as services grow beyond pilot stage
	Use third-party repositories	• Possible protection for personally identifiable information from public disclosure • Warehousing, management, and/or analysis of raw data, as needed	• Data resolution may be too coarse • Requires additional work to check against public records laws • Data ownership sits outside agency • Legal and governance framework still not mature
Multimodal Trip-Planning App	**Establish API requirements for mobility service providers**	• All providers will be seen by travelers • Avoids separate or differing agreements for individual providers	• Lack of standards for most trip discovery and payment APIs • Requires legislative or regulatory action
	Individual API agreements	• Work with providers who are ready to integrate • Avoids legislative process	• Lack of uniform standard for API connections • Might not include providers with large market share

Conclusion and Next Steps

One of the primary challenges of implementing MOD projects has been reaching an agreement between public and private partners over data sharing. SUMC has observed that this challenge has resulted primarily from each stakeholder having particular reasons for requesting or withholding data—privacy concerns, competitiveness, proprietary information, public records laws, and data security being among the most prominent reasons. Through the analyses and discussions in this document, the following noteworthy conclusions were reached:

- Agencies should understand what potential providers' approaches to data sharing are and aim to select a partner with whom they can find a mutually agreeable parameter set and aggregation level for data sharing.

- If constraints related to public records disclosures or agency capability are impeding progress to obtaining the data that they need, agencies should explore using a third-party repository, provided that the information that will be made available is supportive of agency objective.

- Transit agencies and supporting organizations can proactively influence the modernization of public records laws to protect sensitive travel data from public disclosure and still retain data internally for evaluation, planning, and monitoring the performance of their services.

- To overcome the lack of availability of API data for multimodal trip planning, transit agencies, together with states or cities, can establish requirements to open up basic data parameters needed for developing trip-planning apps.

- The federal government could consider requiring partnerships to have clearly defined data management strategies to be eligible for federal funding.

- Regardless of their geographic location, size, coverage area, customers, or culture, transit agencies and their partners would benefit from following a structured approach while forming any partnerships to integrate their services. This structured approach would include clearly defining the objectives and associated performance indicators, identifying the required data and data constraints, identifying the partners and establishing agreements where the required data will be available, strategizing on data management, managing associated tradeoffs, and working on regulatory barriers that could provide opportunities for better service for the public.

The MOD Sandbox projects are in their early states of implementation, which will be followed by complete performance evaluation. The Sandbox initiative is a platform for testing different data arrangements and a convening platform to share lessons on this issue and work toward best practices moving forward.

This paper complements the FTA's development of the Mobility Performance Metrics (MPM), which aim to identify new supplemental performance indicators for integrated mobility systems and services. The MPM efforts introduce supplemental performance indicators and assess their feasibility for measuring various metrics that are typically not covered within the traditional performance measurement approaches. Furthermore, the MPM effort attempts to identify gaps and redundancies between the proposed metrics and goals of MOD and other mobility integration projects, and evaluate broader policy implications of operationalizing those metrics to measure the performance of integrated mobility services. This effort will shape the longer-term approach for the industry to frame the data sharing issue in the appropriate broader context of agency goals, project objectives, and emerging measurement methods.

A stand-alone Executive Summary of this paper can be found at
www.sharedusemobilitycenter.org/publications

**SHARED-USE
MOBILITY CENTER**

312.448.8083
www.sharedusemobilitycenter.org