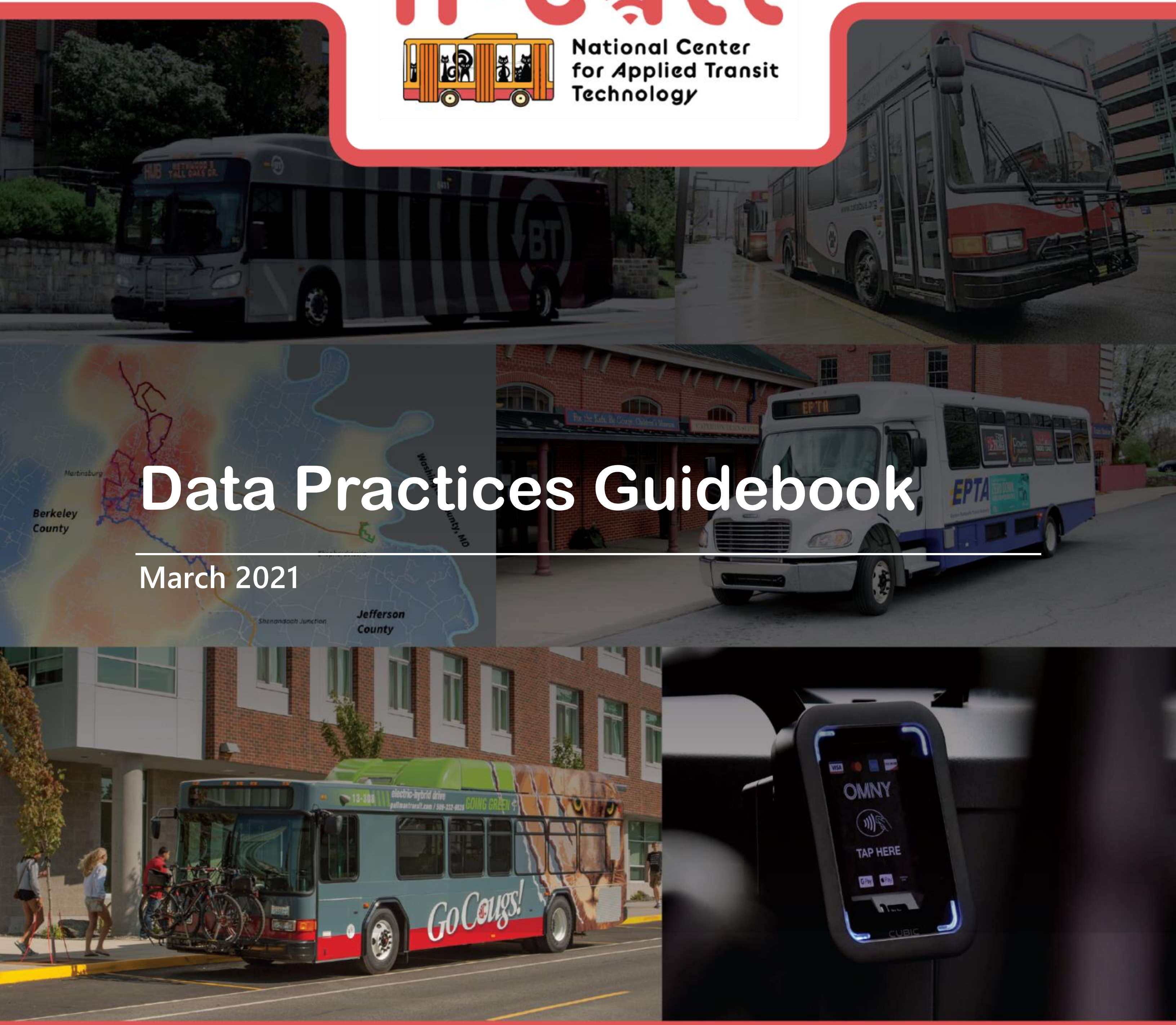
National Center for Applied Transit Technology



Data Practices Guidebook

March 2021

Prepared by:

FOURSQUARE ITP
INTEGRATED TRANSPORTATION PLANNING

CONTENTS

FIGURES

TABLES

EXECUTIVE SUMMARY

> Data Practices collect, structure, or analyze data to improve agency operations and how transit riders are served

The National Center for Applied Transit Technology (N-CATT) is a technical assistance center with a mission to provide small urban, rural, and tribal transit agencies with practical resources that help them apply technological solutions and innovations. To this end, N-CATT presents the Data Practices Guidebook, a resource for agencies seeking actionable insights from data. Beginning with an overview of how agencies can obtain more and better data, the Guidebook addresses the state of the practice and emerging trends in data collection, management, and the use of open data. On this foundation, the Data Practices Guidebook then guides agencies on how to use data through the lenses of planning and performance monitoring, partnerships, and access and equity.

The data practices featured in this Guidebook are tailored to small urban, rural, and tribal agencies. While the discussion of these practices is drawn from experiences at such agencies, some emerging data practices are found primarily at transit providers in major metropolitan areas. This Guidebook seeks to provide smaller agencies with both an understanding of the range of current practices, as well as others that may become more widespread in the coming years. Transit managers and planners using this Guidebook can use the prompts and checklists provided to consider current and emerging data practices that may be well-suited to your agency.

Types of Data Practices

Practices for Obtaining Data

For transit agencies, making good use of data begins with identifying, collecting, and storing the data sources your agency may already have—and in some cases, accessing useful data for planning from external sources. This data may be manually collected or represent digitized paper records, but increasingly agencies rely on passive, sensor-based systems to automatically harvest large quantities of data. With raw data in hand, a best practice is to use common data standards to format and manage data. Doing so paves the way for quicker analyses or allows for broader dissemination of data to customers and stakeholders. Whether publishing data or making use of open sources like the US Census, transit agencies stand to benefit from participating in the broader open data ecosystem.

Manual Data Sources

For some agencies, digitizing existing records or simply collecting more data manually will be the first priority when working with data. Such data includes ridechecks, trip requests, on-board customer surveys, and customer service complaints, as well as internal records on asset condition, safety incident reports, and maintenance data.

Passive Data Sources

Passive, or automatic, data sources collect information using electronic sensors and are capable of generating large volumes of highly granular data. Although these technologies carry a higher capital cost than manual data sources, passive sources have lower marginal costs as well as increased data coverage, availability, and reliability.

Emerging Advances in Data Collection

While sensor-based, passively generated data is increasingly part of the state of the practice for transit agencies, emerging technological advances are changing how such data is generated: from internet-connected parts and systems to customer cellphone location data, to the data generated as agencies provide more customized services.

Existing Data Standards

Standards can be either formal data standards that are adhered to across the industry or informal data practices that are defined at each agency. Regardless, standards are a key determinant of how much agencies—as well as customers and stakeholders—can get out of data. The General Transit Feed Specification (GTFS) for both static transit schedules and real-time updates is the most prevalent standard for data used in the transit industry today.

Emerging Data Standards

Standards are emerging to provide demand response and other flexible transit services with the same level of visibility as GTFS provides to fixed-route services. Systematically recording and managing trip history and ridership data are also set to be improved by emerging standards.

Open Data

As data producers and data users, transit agencies can benefit from accessing external open datasets as well as opening their transit data. Open data practices lead agencies to promote service awareness and transparency, improve efficiencies, spur innovations, improve customer satisfaction, become more engaged with customers, and improve service and data quality.

Practices for Using Data

Data-driven planning decisions are the key to maximizing the use of limited resources for the greatest benefit of transit riders. While the use of data to inform planning and performance monitoring decisions is not new, the richness of new data sources provides agencies with the ability to conduct more precise evaluations of performance and to plan services with greater clarity. However, many small agencies do not have the resources to collect or analyze all the data they need to. Partnerships with other organizations—whether they be another transit agency, a non-profit, a state agency, a business, or other entity—can help provide insights, technology, or expertise. Data also provides a means to assess who transit serves and how well it serves them in richer detail than ever before.

Planning

For transit service planning, data provides key insights for decision-makers when designing and evaluating service changes. Though some data like ridership and revenue hours will be tabulated by the agency itself, external datasets such as population and employment estimates will play a role in many planning decisions. As services change and fleets age, decisions on future year capital investment needs can also be made on the basis of asset data.

Performance Monitoring

As communities grow, travel behaviors change, and agency conditions fluctuate, so too do metrics like ridership, on-time performance, service efficiency, revenues, and costs. Transit performance monitoring is the process of reporting a set of performance measures repeatedly over time. However, rural, tribal, and small urban agencies structure their performance monitoring efforts, new data sources provide new opportunities to gauge performance.

Partnerships Providing Access to Data

Not all data is "open," and not all open data is readily used. As a result, agencies must rely on other organizations to obtain datasets vital to transportation planning. Both public and private sector entities can be potential partners, including local school districts, colleges and universities, major employers, metropolitan planning organizations (MPOs), regional planning agencies (RPAs), municipal departments, state DOTs, and other transit agencies.

Partnerships for Data Analysis and Technology

ITS devices and thorough data analysis come at the cost of both capital outlays and staff expertise. Transit agencies need not conduct all procurement or data analysis themselves: often, there are partners that can help agencies procure, develop, or maintain technology or who can process datasets that new technologies produce. In some cases, agencies are sharing their own proprietary technology with other agency partners.

Using Data to Build Partnerships

It can be difficult for transit agencies to prove their importance to the agencies and organizations that fund, oversee, or partner with them. When building a new relationship with a partner organization, providing proof of the effectiveness of an agency can be even more difficult. Certain types of data analyses can help illustrate a transit agency’s importance to a community and make the case for more funding support from current and existing partners.

Using Data for Access and Equity

Evaluating accessibility to key destinations and jobs is an emerging trend in the industry. Transit agencies have begun to pay more attention to accessibility metrics, which examine how well transit provides connections to key destinations. Although equity analysis is required through Title VI, there are opportunities for agencies to address these questions more thoroughly. Equity considerations include service offered, fare payment types, trip purpose analysis, and trip planning accessibility.

Profiled Data Practices

This Guidebook surveys both the current state of data practices at small transit agencies around the country, as well as recent advances in this domain. Following an initial desk review of the transit literature, notable implementations of data practices by small agencies were selected for more detailed research. For these profiled data practices, interviews with transit staff were used to document the context and methods of the practice, significant barriers faced, and the outcomes and lessons learned following implementation.

Technology Adoption at Frederick TransIT

Over the past several years, TransIT, a small urban agency in Frederick, Maryland, has implemented a number of technology upgrades to improve the passenger experience, data collection, and data reporting. These upgrades included installing automated passenger counters (APC) and automatic vehicle location (AVL) devices on its fixed-route fleet, launching a new mobile ticketing app, and purchasing nine fully electric vehicles.

Customer Mobile App at Pullman Transit

Pullman Transit, located in eastern Washington, deployed a new mobile app in 2019 when it updated its AVL system. The mobile app has many capabilities, including real-time vehicle tracking, passenger load tracking, a trip planner, automated announcements, and the ability to receive feedback. The feedback Pullman receives from riders through the app is also important as it helps identify real-time issues.

"Where's My Bus" at EPTA

After conversations with Bluefield Transit, another small transit agency in West Virginia, the Eastern Panhandle Transit Authority (EPTA) learned how they could create a real-time bus location web interface with their Automatic Vehicle Locator (AVL) data. EPTA enlisted the Berkeley County Information Technology (IT) department's help to develop a similar interface. The site can be viewed easily on a mobile phone in addition to EPTA's website.

GTFS Creation at CATA

Crawford County and Venango County, Pennsylvania consolidated their transit services in 2016, with Venango County Transit merging into CATA. The consolidation of services required a merging of both agencies' datasets, which were in very different formats. To accurately merge data, create a better inventory of stops, and improve trip planning for customers, the agency created a GTFS feed using the Rural Transit Assistance Program (RTAP) GTFS Builder and Google Earth.

Blacksburg Transit Real-Time Bus Location Text Service

Blacksburg Transit (BT) provides service in Blacksburg and Christiansburg, VA, with much of its service centered on Virginia Tech University. Driven by requests from students for more real-time information, BT partnered with a computer science program at Virginia Tech to launch a real-time bus location texting service. This service was later turned into a mobile app that enabled users to see bus locations and passenger loads in real-time, use a trip planner, and receive alerts from BT, among other features.

LIST OF ABBREVIATIONS

AASHTO	American Association of State Highway and Transportation Officials
ACS	American Community Survey
AFC	Automatic Fare Collection
APC	Automatic Passenger Counters
API	Application Programming Interface
AVL	Automatic Vehicle Location
BI	Business Intelligence
BRATS	Baldwin Regional Area Transit System
BT	Blacksburg Transit
CAD	Computer-Aided Dispatch
CASD	Computer-Aided Scheduling and Dispatch
CATA	Crawford Area Transportation Authority
CSV	Comma-separated values
CTAA	Community Transportation Association of America
CTPP	Census Transportation Planning Products
CUTR	Center for Urban Transportation Research
DI	Disparate Impact
DB	Disproportionate Burden
DST	Decision Support Tools
DOT	Department of Transportation
EPTA	Eastern Panhandle Transit Authority
EPC	Electronic Passenger Counters
FTA	Federal Transit Administration
GEOID	Geographic entity code
GTFS	General Transit Feed Specification
GTFS-Flex	General Transit Feed Specification for Flexible Transit
GTFS-RT	General Transits Feed Specification–Real Time
GIS	Geographic Information Systems
GPS	Global Positioning System
GPX	Global Positioning System Exchange Format
HVAC	Heating, Ventilation, and Air Conditioning
ID	Identification
iNTD	Integrated National Transit Database
IT	Information Technology
ITS	Intelligent Transportation Systems
IoT	Internet of Things
KML	Keyhole Markup Language (from Google)
LBS	Location-Based Services
LEHD	Longitudinal Employer-Household Dynamics
LODES	LEHD Origin-Destination Employment Statistics

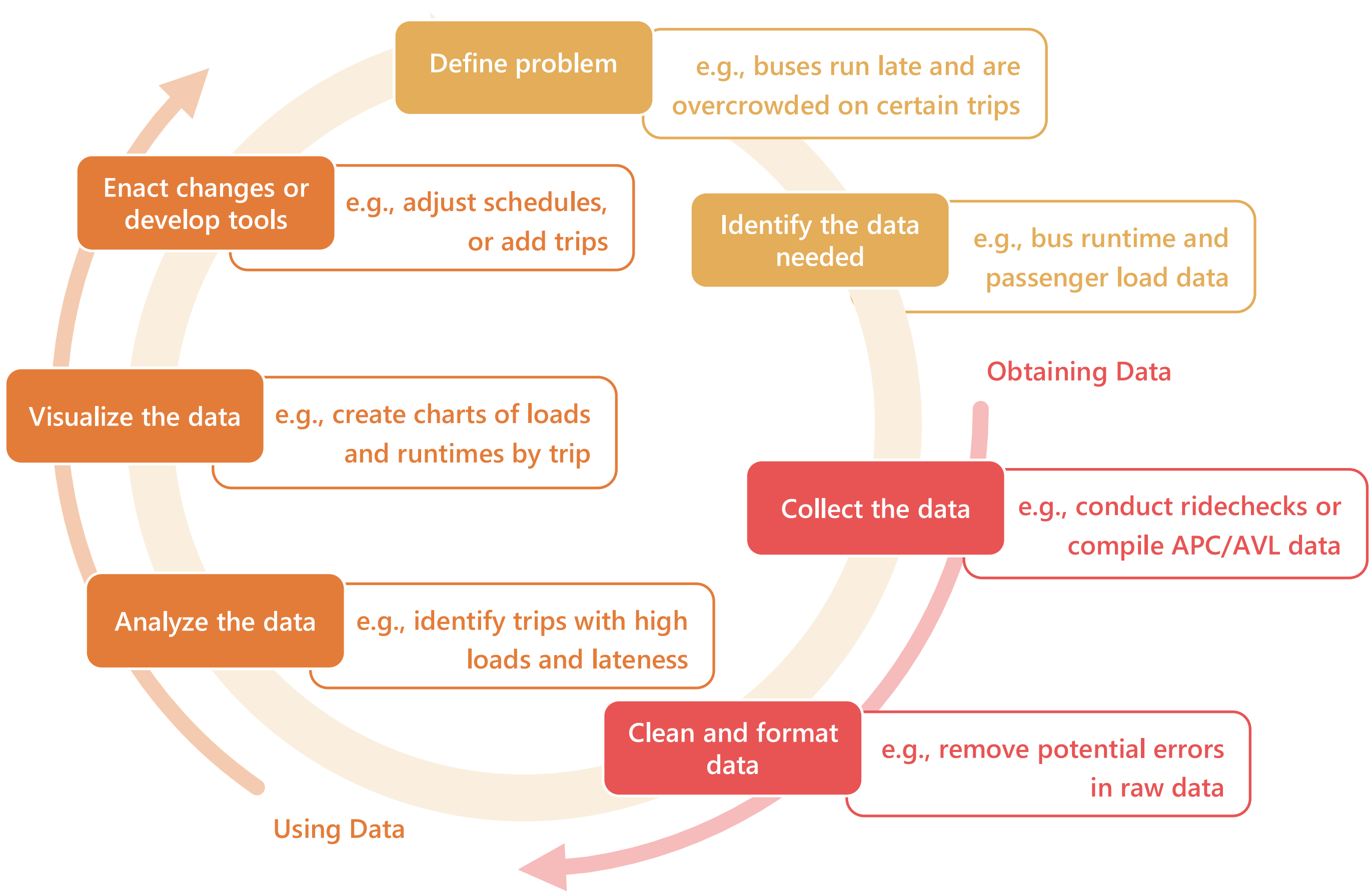
MaaS	Mobility-as-a-service
MARTA	Metropolitan Atlanta Rapid Transit Authority
MAP-21	Moving Ahead for Progress in the 21st Century Act
MATS	Muskegon Area Transit System
MDBF	Mean Distance Between Failure
MOD	Mobility On-Demand
MPO	Metropolitan Planning Organization
MTD	Champaign-Urbana Mass Transit District
N-CATT	National Center for Applied Transit Technology
NEMT	Non-Emergency Medical Transportation
NTD	National Transit Database
OD	Origin-Destination
OSM	OpenStreetMap
OTP	On-Time Performance
PAC-12	Pacific Coast Conference School
PAT	Pulaski Area Transit
PDF	Portable Document Format
PHP	Hypertext Preprocessor (programming language)
PMT	Passenger Miles Traveled
RFP	Request for Proposals
RPA	Regional Planning Agency
RTAP	Rural Transit Assistance Program
R.Y.D.E.	Reach Your Destination Easily
SEA	Service Equity Analysis
SGR	State of Good Repair
SMMPO	Southeastern Massachusetts MPO
SRTA	Southeastern Regional Transit Authority
STOPS	Simplified Trips-on-Project Software
TAM	Transit Asset Management
TAPT	Transit Asset Prioritization Tool
TBEST	Transit Boardings Estimation and Simulation Tool
TCaT	Tulare County Area Transit
TCRP	Transit Cooperative Research Program
TERM-Lite	Transit Economic Requirements Model Lite
TIGER	Topologically Integrated Geographic Encoding and Referencing
TMA	Transportation Management Association
TNC	Transportation Network Company
ULB	Useful Life Benchmark
VVTA	Victor Valley Transit Authority
WSU	Washington State University

INTRODUCTION

The National Center for Applied Transit Technology (N-CATT) is a technical assistance center with a mission to provide small urban, rural, and tribal transit agencies with practical resources that help them apply technological solutions and innovations. With better data, a transit agency can more clearly predict future needs, streamline its operations, and better serve the diverse needs within its community. But for small urban, rural, and tribal transit agencies, making good use of data—let alone generating or storing it—requires navigating an ever-growing set of data sources, standards, and analytical methods. To this end, N-CATT presents the Data Practices Guidebook, a resource for agencies seeking actionable insights from data.

At transit agencies, data addresses subjects from customer demographics to agency performance to financial conditions. Though data is often a byproduct of of fare collection, transit operations, or other agency functions, data can also be collected through open data portals, partnerships, or other third-parites. Through more deliberate internal data collection and management efforts, agencies can maximize the value of data in contributing to problem-solving and service improvement. **Figure 1** illustrates the cycle of using transit data to solve problems and improve service, while providing an implementation example.

Figure 1: Cycle of Data Use to Solve Problems and Improve Services



Navigating This Guidebook

For transit agencies, making good use of data begins with identifying, collecting, and storing the data sources your agency may already have. The three first chapters of this Guidebook examine these topics, including obtaining useful data for planning from external sources. The first chapter covers data sources ranging from manual data recorded with pen and paper to sensor-based systems that automatically harvest large quantities of data. With raw data in hand, Chapter Two discusses ways to use common standards to format and manage data. Chapter Three dives into how transit agencies, as data producers and data users, can benefit from accessing external open datasets, as well as opening their transit data.

Data-driven planning decisions are the key to maximizing the use of limited resources for the greatest benefit of transit riders. In this context, the three latter chapters of this Guidebook demonstrate how data can be used for planning and performance monitoring, for shaping partnerships, and for promoting access and equity. While the use of data to inform planning and performance monitoring decisions is not new, the richness of new data sources provides agencies with the ability to conduct more precise evaluations of performance and to plan services with greater clarity. Partnerships with other organizations—whether they be another transit agency, a non-profit, a state agency, a business, or other entity—can help provide insights, technology, or expertise. Lastly, data also provides a means to assess who transit serves and how well it serves them in richer detail than ever before.

Making good use of data is made easier with thoughtful guidance on best practices and lessons learned from peers. For this reason, best practices around the country are featured throughout this Guidebook in call-out boxes. Notable implementations of data practices by small urban, rural, and tribal transit agencies were selected for more detailed research following a desk review of the transit literature. Interviews with transit staff were used to document the context and methods of the practice, significant barriers faced, and the outcomes and lessons learned following the implementation. These data practices are:

- Technology adoption at Frederick TransIT in Chapter One
- General Transit Feed Specification (GTFS) file creation at Crawford Area Transportation Authority (CATA) in Chapter Two
- Real-time bus location web interface "Where's My Bus" implementation at Eastern Panhandle Transit Authority in partnership with Berkeley County in Chapter Five
- Mobile application deployment at Pullman Transit in Chapter Six
- Real-time bus location text service implementation at Blacksburg Transit in Chapter Six.

About N-CATT

The National Center for Applied Transit Technology (N-CATT) is a technical assistance center funded through a cooperative agreement with the United States Department of Transportation's Federal Transit Administration (FTA). Operated by the Community Transportation Association of America (CTAA), the mission of N-CATT is to provide small urban, rural, and tribal transit agencies with practical, replicable resources that help them apply technological solutions and innovations. Among its activities, N-CATT produces a series of white papers, technical reports such as this document, and other resources, all of which can be accessed online at https://n-catt.org.

About This Document

This document was prepared for CTAA by Foursquare Integrated Transportation Planning in March 2021 as part of the N-CATT cooperative agreement between CTAA and FTA. Primary authors were Adam Recchia, Reinaldo Germano, Wylie Timmerman, Jack McDowell, and Rebecca Martin of Foursquare ITP. Opinions expressed or implied in this document are those of the authors. Nothing in this document is to be interpreted as position, policy, or guidance from the United States Government. Incidental use of companies' names or the names of their products is made solely to facilitate discussion and should not be regarded as recommendations or endorsements.

WHAT EXACTLY DO WE MEAN BY "DATA?"

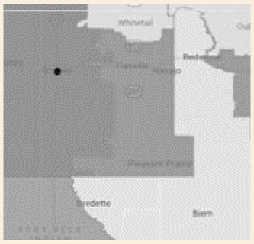
Data takes many forms, and the word itself has many meanings. For the purposes of this report, data is digital information that is collected, structured, and can be further analyzed. The first section of this report will speak to critical issues regarding collecting and structuring data for transit agencies, while the second section of the report will address how agencies are analyzing data.

Data collection is more than just a matter of clipboards, paper, and pens. Rural, tribal, and small urban agencies are increasingly making use of sensors to passively collect large amounts of data that can be mined for transportation insights. In some cases, agencies may be able to gain access to novel datasets based on aggregate cellphone location data from third-party providers.

Manual Data Sources

Passive Data Sources

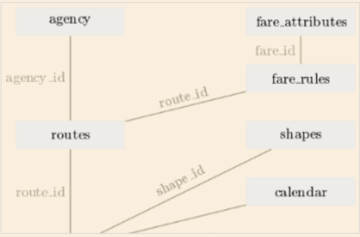
External Data



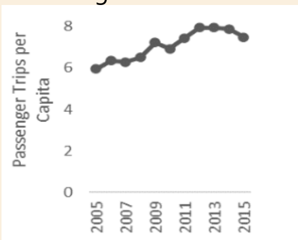
Data structure and management is the next challenge agencies must face. Most data maintained by transit agencies are structured into a tabular format: tables with rows representing an individual record and columns with the characteristics, often numeric values, text, or dates. The kinds of data that exist are, in fact, much broader, but, in practice, many of the forms of data described in this report will fit this description. For effective analysis of data to take place, agencies must be mindful of data cleaning, data structures, common data standards, storage considerations, and consistent key fields—like route names and dates—so that disparate data sources can be related together as needed.

A growing expectation of public agencies is that they will publish "open data" for inspection and analysis by the broader public. For some transit agencies, this will simply be a matter of submitting National Transit Database (NTD) reports, while others may create open data websites that provide route geometry files, transit schedule data, and ridership data. Open data, in particular, is more likely to have longitudinal characteristics (tracking changes to records over time, such as an agency's ridership) or spatial characteristics (recording the location of features, such as where stops are located). Even if your agency does not directly publish open data, this report will give you guidance on how to make use of open data sources that are available to you.

Data Standards



Longitudinal Data



Spatial Data

1. DATA SOURCES

Good data allows for good decision-making, and understanding your agency's data sources is the first step to collecting, storing, and analyzing data. Data sources can range from manual sources recorded with pen and paper to electronic sources recorded with mobile devices to passive sources that are automatically collected using installed hardware and connected technologies. In the digital age, new data sources including location-based services data are increasingly available to agencies. This fact sheet provides an overview of transit data sources that could be useful to small agencies and allow you to unlock insights into your customers' behavior, your system's performance, and market demand in your service area.

Transit data sources can be manual or passive, meaning they can be collected manually or automatically using on-board technologies

Manual Data Sources

Ridechecks

Ridership can be estimated through methods such as a ridecheck. Ridechecks are performed by staff that count activity at each stop on sampled trips. These sampled trips can then be used to estimate overall ridership for the system or, with a large enough sample, specific routes, months, or times of the day.

Trip Requests and Trip Logs

For demand response services, trip requests and trip logs are a valuable source of data. Trip requests are documented by dispatch, including the origin, destination, desired timing for the trip, and the number of passengers. Trip logs can then provide agencies with important information about the level of demand for their service.

Asset Management and Maintenance Data

For small agencies, transit asset management (TAM) data collection is necessary for state reporting. The foundation for a TAM plan is an asset inventory, which catalogs the rolling stock, service vehicles, and facilities operated by an agency. The condition of each of these assets must be assessed to determine whether the asset is in a State of Good Repair (SGR).

Safety Incident Reports

In the case of a safety incident, many agencies file a report or obtain copies of police reports. As defined by NTD, safety incidents include injuries, fatalities, collisions, significant property damage, and a transit vehicle or facility's evacuation. As a data source, detailed safety incident reports allow agencies to identify risks or patterns that can be addressed to improve their transit service's safety.

Customer Service Complaints

Customer feedback via phone calls, texts, emails, mobile applications, or online forms can help agencies bridge the gap between the rider and decision-making staff. To observe trends and track customer satisfaction over time, agencies may record the complaints they receive in a spreadsheet or database to enable comparison and aggregation.

On-Board Customer Surveys

In a customer survey, agencies engage with their customers on board vehicles or at transit stops to learn more about their age, income, trip purpose, or even their origins and destinations. Getting an adequate sample size so statistics are relevant to the entire system is key to an informative survey.

Passive Data Sources

Automatic Passenger Counting (APC)

APC devices measure boardings and alightings at the door of a transit vehicle. As a passive data source, APCs collect data at each stop a vehicle makes, providing agencies with more complete ridership data.

Automatic Fare Collection (AFC)

AFC systems are electronic fare systems that often include fareboxes, fare machines, and customer-facing websites and mobile applications that allow for fare transactions. With an AFC fare system, passengers can pay using a smart card, a mobile app, cash, or other smart fare media.

Automatic Vehicle Location (AVL)

AVL devices use global positioning system (GPS) technology to record the location of a bus, enabling real-time location monitoring as well as subsequent location-related analyses. Together with schedule data, AVL data can be used to evaluate on-time performance.

These findings are from the Data Practices Guidebook. The Guidebook is a resource to assist small urban, rural, and tribal transit agencies in understanding and applying good data practices.

Emerging Advancements

Internet of Things (IoT) Technologies

IoT devices are internet-connected, physical equipment that can relay information about their status or be controlled remotely. Devices are equipped with sensors that relay information about vibrations, pressure, heat, power usage, and other conditions to central servers. These features allow agencies to continuously monitor asset conditions.

Aggregate Cellphone Location Data

Aggregate cellphone location data purchased from vendors that obtain, aggregate, and anonymize Location - Based Services (LBS) data. The resulting datasets can provide agencies with information on origins and destinations in their service areas.

1.1. Introduction

For transit agencies, making good use of data begins with identifying, collecting, and storing the data sources your agency may already have—and in some cases, obtaining useful data for planning from external sources. For some agencies, digitizing existing paper records for further analysis and improving digital recordkeeping will be the first objective. Other agencies may be ready to begin installing automatic passenger counters (APC) or making use of emerging data sources like cellphone location data. These new sources provide richer insight for planning and performance management but demand greater resources and expertise from agencies.

This chapter provides a broad overview of the data sources small urban, rural, and tribal agencies can leverage for data practices ranging from data-informed planning to partnerships to equity analyses. As defined here, data sources are those records generated within an agency or proprietary information about travelers purchased from a vendor (open data sources like US Census data will be discussed in Chapter 3). The data sources discussed here may represent safety incident reports, ridecheck data, or other digitized agency records created manually by an individual. For some small urban, rural, and tribal agencies, data sources will also include data created by sensors like automatic passenger counters (APCs) that generate data passively.

The current trend among transit agencies is to make increased use of passive data sources that collect data automatically but that require more effort to analyze. For instance, many agencies seek to collect more ridership data by using automatic passenger counter (APC) devices on vehicles. Because this ridership data is not input by a human, these sources may be more accurate, but defects with sensor accuracy may create other problems when working with this data. Understanding your data sources and their shortcomings are essential elements of strong data practices.

The current trend among transit agencies is to make increased use of passive data sources that collect data automatically but that require more effort to analyze.

Tapping into new data sources—or sticking with familiar ones—is not without tradeoffs. Agencies will need to consider the costs of data storage, purchasing, and especially the staff expertise required to work effectively with more complex data sources. Data sets with customer information, such as on-board surveys and demand-response service records, also raise implications for privacy. Key characteristics of data sources are described below. Differences between manually and passively-collected data according to these characteristics are shown in **Table 1**.[1]

- **Capital cost:** The installation of sensors to collect data passively may require retrofitting your bus fleet, while collecting data manually may require little more than clipboards and a spreadsheet to store the collected data. Because of the frequency and scale with which passive data is collected, it often also requires more sophisticated databases and data feeds that are maintained by a vendor or an IT department staff.
- **Marginal cost:** Manually collected data requires the effort of staff to collect and input data. Each additional data observation requires a marginal effort that is relatively high and does not decrease with volume. Automatically-generated passive data is cheap—the cost of collecting and storing additional records of data might only be measured in pennies.
- **Sample sizes:** Drawing conclusions from data requires large enough sample sizes to draw statistically significant conclusions. Because of the high marginal cost of manually collected data,

[1] Wilson et al. "Public Transportation Systems: Data Collection Techniques and Program Design." (Lecture Notes, MIT, Spring 2017.) https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-258j-public-transportation-systems-spring-2017/lecture-notes/MIT1_258JS17_lec02.pdf. Accessed January 11, 2021.

generally only small sample sizes can be obtained. With the volume of data generated passively by sensors, you can obtain larger sample sizes and have more confidence in the conclusions drawn from your data.

- **Aggregation:** Data practices are put to use for different audiences: a board member looking at monthly trends or an analyst diving into distinctions by route or time of day. Both manual and automatic data collection provide opportunities to collect and report aggregate or disaggregate data.
- **Category of Data:** When data is collected manually, you have the opportunity to collect qualitative data about your system and passengers that a sensor alone could never detect. For instance, when conducting passenger counts manually, it may be more clear what types of customers—students, older adults, workers—get on and off at particular stops, and field observations may reveal how the system is used in practice. However, qualitative impressions are more subject to human biases and misperceptions.
- **Reliability:** The human eye can be fallible, such as when passengers are accidentally left uncounted on a ridecheck. It is also difficult to know how the data collected by two different persons may differ. Sensors are not without challenges, such as when automatic passenger counters accidentally count one person as two. Unlike manual data, issues with automatic data sources are more consistent over time and can be corrected prior to analysis. However, it is often difficult to outfit an entire fleet with APCs at once, and therefore getting an adequate sample size on every route can be challenging when only certain vehicles are equipped.
- **Coverage:** When relying on manually collected data, agencies may find gaps in datasets for periods where data was not collected or aggregated, especially if the demands of staff time to collect data are high. Especially on holidays, early in the day, or late in the evening, additional staff may not be available to conduct data collection. With passively generated data, such constraints on data collection apply less if sensors can be widely installed.
- **Availability:** In certain cases, automatically-generated passive data can be fed directly into analytical tools for real-time or near real-time answers to questions involving data. With manual data, more intentional efforts to collect and then aggregate data are required, limiting the frequency with which it is available.

This chapter will illustrate these key considerations for both emerging data sources and familiar data sources.

Table 1: Manual and Passive Data Collection Characteristics

Characteristic	Manual	Passive
Capital Cost	Low capital cost	High capital cost
Marginal Cost	High marginal cost	Low marginal cost
Sample Sizes	Small sample sizes	Large sample sizes
Aggregation	Aggregate or disaggregate	Aggregate or disaggregate
Category of Data	Qualitative or quantitative	Quantitative
Reliability	Unreliable	Errors and biases can be estimated
Coverage	Limited spatially and temporally	Ubiquitous
Availability	Not immediately available	Available in real-time or quasi-real-time

1.2. State of the Practice

1.2.1. Manual Data Sources

At small transit agencies, many sources of data may be collected manually. For these data sources, the volume of data is typically small, requiring a limited investment of time and personnel to gather data and enter it into a spreadsheet or database.

Manual Data Collection Methods: Analog vs. Digital

Tools for manual data collection fall into two categories: analog and digital. Analog methods, such as paper forms or phone calls with customers, often require data collection and data entry as separate steps, each requiring staff resources. Digital methods, such as online forms, mobile applications, or specialized electronic devices, merge the two steps into one by entering data into a spreadsheet or database as it is collected. Advantages of each type of method are listed below:

Analog Methods

- No new technology required (smaller up-front cost)
- Limited technical skills required.

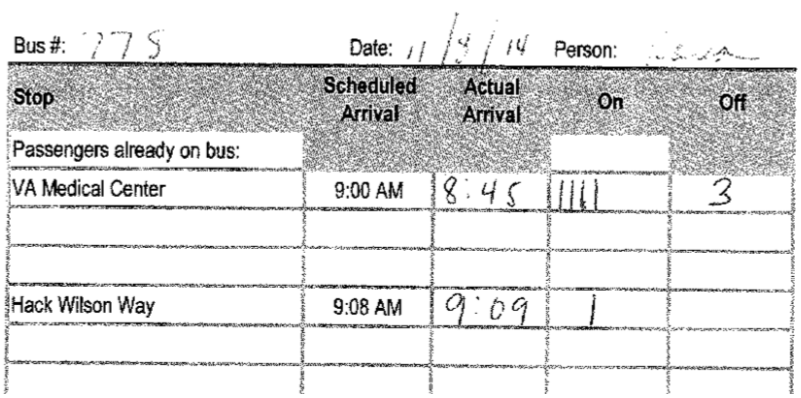
Digital Methods

- More efficient data entry
- Consistent formatting of data
- Reduced staff needs (smaller long-term cost)
- Partial passive data collection (e.g., timestamp).

Ridechecks

If passengers are the lifeblood of a transit system, ridership counts are the pulse. For smaller transit agencies, it may be practical to count 100 percent of passengers that board and alight each vehicle. For agencies where a complete count of boardings and alightings at each transit stop cannot practically be obtained, ridership can be estimated through methods such as a ridecheck. Ridechecks are performed by staff that count activity at each stop on sampled trips.[2] These sampled trips can then be used to estimate overall ridership for the system or, with a large enough sample, specific routes, months, or times of the day.

Figure 2: Manual Ridecheck Form



Bus #: 275 Date: 11/3/14 Person:

Stop	Scheduled Arrival	Actual Arrival	On	Off
Passengers already on bus:				
VA Medical Center	9:00 AM	8:45	IIII	3
Hack Wilson Way	9:08 AM	9:09	I	

Reliability, the ability to disaggregate data, and, most of all, the cost of a ridecheck depend greatly on sample size. Ridechecks conducted for National Transit Database (NTD) ridership reporting must follow a sampling strategy that randomly selects trips, be they an agency's busiest services, early morning trips, or

[2] For larger agencies that must report passenger miles traveled (PMT) to the National Transit Database, alightings must also be recorded. These requirements are specified in the NTD Policy Manual for Reduced Reporters.

weekend services.[3] When conducting ridechecks for other purposes, the sample size for a ridecheck will vary depending on available resources. Most agencies often aim to sample every trip on every route at least once, obtaining a sample across all days of the week and times of the day. Whatever the approach, known variations in ridership by the time of day, day of the week, and time of year must be considered when developing a ridecheck sampling strategy so that the ridership estimates are not inflated (or underestimated) by the choice of trips to sample.

Most agencies will perform ridechecks to satisfy NTD reporting requirements, which will vary according to the size of the fleet, urbanized area, and counting methods used by an agency. However, ridership data can also be used to perform other analyses, as will be discussed in Chapter Four: Planning and Performance Monitoring. Uses for ridership data may include:

- Measuring the maximum passenger load per trip
- Identifying ridership trends by time of day
- Comparing boardings and alightings along a route to identify stops with the greatest and least ridership.

Ridechecks may be performed using pen-and-paper-based methods or manual entry into a spreadsheet or online form. While riding on a bus route, the observer may be required to manually enter the route, trip, and stop information associated with each recorded observation. Alternatively, this information can be populated prior to sampling using the route's schedule. In some cases, data collected with an electronic device allows for route, trip, and stop information to be inferred during subsequent analysis based on timestamps and geographic location data associated with each record.

Figure 3: EPC Device Example

Electronic passenger counters (EPC) offer a manual form of ridership counting with many of the advantages of higher-cost passive data collection devices. With an EPC device, bus drivers press a button on a tablet for every boarding and alighting at each stop. The device can be programmed with the planned schedule, automatically recording the route, trip, and stop for each data point. As an electronic data source, data from EPC devices can be uploaded directly to a database. While the capital cost is greater than pen-and-paper or simpler electronic methods, the marginal cost is reduced by using bus drivers instead of separate staff to perform counts. In addition, the sample size is maximized for full coverage: EPCs enable agencies to measure ridership on every trip. As with all manually collected data, reliability is still low, particularly since drivers must multitask by counting riders while collecting fares and managing passengers. Because of the demands of both manual ridechecks and EPCs, larger agencies may conduct ridership estimates using Automatic Passenger Counters that collect far larger samples.

[3] Specific guidance for sample-based NTD ridership reporting is provided in the National Transit Database Sampling Manual. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/The_NTD_Sampling_Manual.pdf

Trip Requests and Trip Logs

For demand response services, trip request information and trip logs are valuable sources of data. Every trip that is requested and delivered is recorded by demand response dispatchers and drivers, with key data points that may be used to identify key trends in the operation and usage of the service. Common datapoints include the origin, destination, desired timing for the trip, and number of passengers. Dispatchers may also record the trip purpose and any special requests, such as wheelchair assistance or pickup details. Upon scheduling the trip, data recorded by dispatch may include the assigned vehicle and driver, the pickup window, and estimated trip duration. Independent of trip logs representing actual delivered service, trip requests provide agencies with important information about the level of demand for their service. Insights into who wishes to use the service, where riders live, when they make trips, and common destinations assist operators when scheduling drivers to meet capacity. Planners can also utilize this data when determining the service area, fares, or other service policies. Additional data, such as the reason for a declined trip request, may highlight the limitations of the demand response service.

Figure 4: Trip Request and Trip Log Process

Trip logs are usually documented by drivers, who record information such as the number of riders, fare payment, time, date, and trip information, including distance and duration. The time and distance traveled in non-revenue service are also an important part of a trip log. In contrast with trip request data, trip logs represent the actual service as delivered. For both internal and mandatory reporting applications, trip log data provides actual ridership, revenue and non-revenue hours and miles, and fare revenue information by trip that may be aggregated to calculate total ridership and total service delivered. In addition to aggregate metrics, agencies can use this data to analyze productivity metrics such as revenue miles per passenger and variable operating costs, including cost per passenger and cost per revenue hour. Beyond reporting requirements, agencies may record and analyze supplemental information on trip logs such as wait times, origin, destination, trip purpose, and other details such as wheelchair ramp usage.

Some of the richest insights from trip request data and trip logs can result from comparing the two datasets. In order to facilitate comparison, the datasets should use similar formats and a common identifier for trips. By comparing scheduled and actual pick up and drop off times, planners can measure on-time performance, the accuracy of trip duration estimates, average wait times. Such evaluations help agencies to identify opportunities for operational efficiency and to fine-tune policies to benefit riders and increase service quality.

Asset Management and Maintenance Data

"State of Good Repair" has become a well-worn phrase since the introduction of transit asset management (TAM) reporting requirements in 2012's Moving Ahead for Progress in the 21st Century Act (MAP-21) surface transportation bill. That legislation and subsequent federal law and regulations require transit agencies to produce a plan to maintain assets in a State of Good Repair. For small urban, rural, and tribal reporters, transit asset management reporting and planning is often facilitated by state departments of transportation, but data collection for these efforts still falls to local agencies. The foundation for a TAM plan is an asset inventory, which catalogs the rolling stock, service vehicles, and facilities operated by an agency. The condition of each of these assets must be assessed to determine whether the asset is in a State of Good Repair.

For revenue and service vehicles, an assessment of state of good repair is often based on vehicle age, but the types of data used to make this judgment may be increasing in the future. The most straightforward judgment of whether a vehicle is in a state of good repair is whether its age has exceeded a useful life benchmark (ULB) for its vehicle class. Of course, each agency (and each vehicle or asset) will experience wear and tear differently, so TAM plans may modify useful life benchmarks based on other factors. In the future, more state agencies creating TAM plans may take into consideration factors like mileage and maintenance history when determining whether an asset has exceeded its useful life. This will increase the amount of data tracked in asset inventories that transit operators must then collect and maintain.

Additional data on asset quality and maintenance come with the benefit of improved forecasting of capital needs and maintenance requirements. The asset data collected for a TAM plan will inform financial and capital planning models (often referred to as decision support tools) that guide how investments are prioritized. Models based on this data can also help to communicate the implications for service reliability and performance if investments are not made. For agencies themselves, increased data collection on assets–and the sub-assets within an asset, such as vehicle engines, wheels, fareboxes, and so forth—can support more pro-active, predictive maintenance work that minimizes disruptions to service from breakdowns. Opportunities to better track the condition of assets with new technology and analyze this data will be discussed in subsequent sections.

State of Good Repair assessments are often based on vehicle age, but the types of data used to make this judgment may be increasing in the future.

In addition to vehicle, technology, and facility assets, the state of bus stop assets is also important to track so that they can be repaired or replaced when necessary. Bus stop assets are particularly important to customer satisfaction as they are often the first piece of the transit system that passengers see. Bus stop assets can be tracked efficiently through Geographic Information Systems (GIS) databases.

Safety Incident Reports

Patterns of transit vehicle collisions and other safety incidents can be used by agencies to target risk reduction measures that improve the safety and health of staff and the public. In the case of a safety incident, many agencies file a report or obtain copies of police reports. As defined by NTD, safety incidents include injuries, fatalities, collisions, significant property damage, and evacuation of a transit vehicle or facility. As a data source, detailed safety incident reports allow agencies to identify risks or patterns that can be addressed to improve the safety of their transit service.

After the occurrence of a safety incident, involved staff typically file a report identifying the route, the agency’s vehicle, other vehicles, all persons involved, a count and description of injuries or fatalities, property damage estimates, and a narrative of what occurred, along with basic information such as the date, location, and time. Enforcement of agency policies regarding the completion of safety reports is important to ensure full coverage of safety data. Incidents that go unreported will reduce the reliability of safety statistics and will reduce opportunities for aggregate metrics such as the number of safety incidents per 10,000 passenger trips.

Greater detail in safety reports will expand the possible use cases for safety incident data. Filling out additional fields requires minimal marginal effort, with important insights made possible as a result. Consistency between reports is also key to enabling aggregation and comparison of quantitative safety metrics. However, many pieces of data will not be applicable to all types of safety events, limiting the usefulness of analyzing such data in aggregate. In addition, a large share of safety incident data is qualitative information that describes events in detail, which is difficult to compare or aggregate. Such qualitative fields, which are often critical for understanding individual events, may not be as useful for analysis.

Analysis of safety incident reports may reveal patterns, such as a vehicle's stairs with numerous falls or an individual driver with multiple incidents. By identifying patterns, agencies can take risk reduction measures, including additional training, vehicle maintenance, and procedural changes. In combination with other data sources, quantitative metrics that use safety incident data may include:

- Safety incidents per 10,000 passenger trips
- Safety incidents per 10,000 revenue hours
- Annual injuries and fatalities
- Frequency of safety incidents by type (e.g., fall, breakdown, collision, property damage, crime).

Customer Service Complaints

For most transit riders, the only agency staff they interact with is their driver. By soliciting customer feedback via phone call, text, email, mobile application, or an online form (see **Figure 5**), agencies bridge the gap between the rider and decision-making staff. In order to observe trends and track customer satisfaction over time, agencies may record the complaints—or kudos!—they receive in a spreadsheet or database to enable comparison and aggregation.

Customer feedback ranges from positive experiences to issues with on-time performance, personal comfort and safety, cleanliness, crowding, and behavior of staff. Typically, complaints are collected as a written comment with basic information regarding the route or service in question and the time. Some agencies may collect additional information with complaints, including the trip purpose, origin and destination, or the vehicle number. These supplemental details may help identify exactly how and why the rider experienced a specific issue. Qualitative responses are also important to make the rider feel heard and understood in their complaint.

However, the qualitative nature of customer service complaints and the different ways it is collected (e.g., through customer service representatives, dispatch, drivers, etc.) often limit the types of analysis possible with such data. Strategies to enable analysis include using categorical data (e.g., "drop-down" selection boxes in a form) instead of open-ended inputs, collecting quantitative data such as a satisfaction rating, and centralizing data from different sources into a master database. Another common challenge with analyzing customer service data in aggregate is the sample size. Relative to the total ridership on a service, customer service data may not be statistically significant. With a large enough sample of quantitative or categorical data, an agency can compare complaints and customer satisfaction between services and track changes over time.

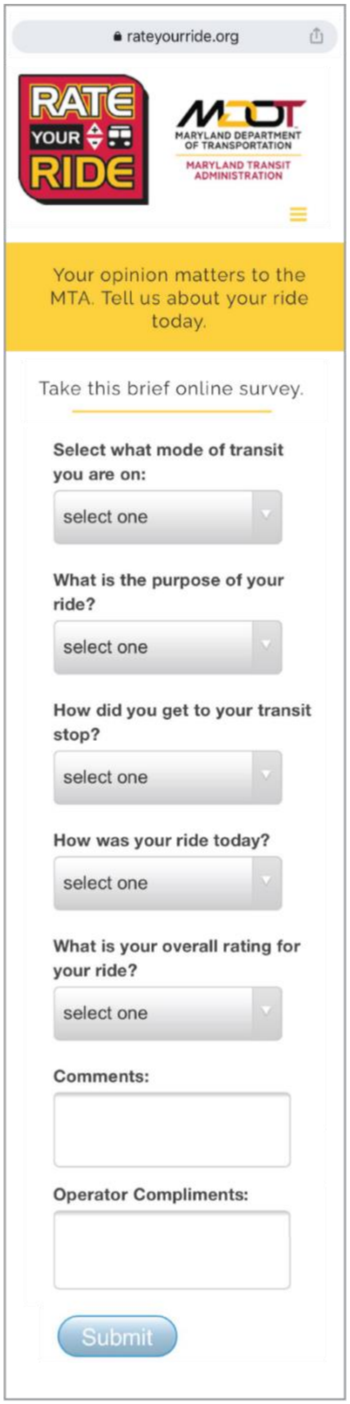


Figure 5: MTA Maryland Rate My Ride Customer Service Feedback Interface

On-Board Customer Surveys

A customer survey (sometimes referred to as an on-board survey or intercept survey) reveals who rides transit services, and how. During these survey efforts, an agency engages with its customers onboard its vehicles or at transit stops to learn more about them: their age, income, trip purpose, or other characteristics. The most sophisticated surveys are origin-destination (OD) surveys that capture more detailed information about how customers use transit to make their trips. While on-board surveys are a required element for larger, urban agencies that must conduct Title VI service equity analyses, they can nevertheless play a useful role for smaller agencies and regional planning bodies looking for data to support planning studies and conduct evaluations of service and fare equity.

When planning a customer survey, the most important consideration for an agency is determining how many responses are needed or the sample size. While an agency may have no expectation of a particular response rate for an online customer survey or a survey about changes to a transit system, the sample size is an important consideration for on-board customer surveys. In order to generalize from a few responses to the entire system, it is important to obtain a sample that reflects the agency's full spectrum of customers—from the first trip of the day to the busiest trip, and from the most frequent services to a quieter route. The calculation of sample size will also play a role in how many staff are required to conduct the survey. Resources provided at the end of this chapter provide straightforward ways for agencies to determine sample size, design a survey, and execute it successfully.

Increasingly, large agencies are relying on emerging data sources to supplement or replace the data gathered from origin-destination on-board surveys. These data sources include synthetic origin-destination data created from AVL and APC data to data about customer travel behavior purchased from private providers. These new sources are discussed in the following section, Passive Data Sources.

Figure 6: Cutout of the Pulaski Area Transit Customer Survey

Pulaski Area Transit

Transit Rider Survey

Please help Pulaski Area Transit improve its service by completing the survey below.

All results will be kept strictly confidential

Take this survey online or on your smartphone at: **pulaskitransit.org**

1. On what route did you receive this survey? ☐ Town Trip ☐ Demand Trip ☐ County/Dublin/Fairlawn Trip ☐ New River Express
2. Where did you begin this one-way trip?
   Name of location: ____
   Address or Intersection: ____
   City/Town: ____
3. Where is your final destination on this one-way trip?
   Name of location: ____
   Address or Intersection: ____
   City/Town: ____
4. How did you get to this bus today?
   ☐ Another bus route: Agency: ____ Route: ____
   ☐ Got a ride/carpooled
   ☐ Taxi/Uber/Lyft
   ☐ Walked
   ☐ Biked
   ☐ Other: ____
5. How will you get from this bus to your final destination today?
   ☐ Another bus route: Agency: ____ Route: ____
   ☐ Get a ride/carpool
   ☐ Taxi/Uber/Lyft
   ☐ Walk
   ☐ Bike
   ☐ Other: ____
6. How did you pay your bus fare today?
   ☐ Town Trip Cash
   ☐ Demand Trip Cash
   ☐ Pulaski County/Dublin/Fairlawn Cash
   ☐ Monthly Pass
   ☐ New River Express Cash
   ☐ New River Express Student Cash

13. What is your race/ethnicity?
   ☐ American Indian/Alaska Native ☐ Black or African-American ☐ Asian ☐ Native Hawaiian or Pacific Islander ☐ White ☐ 2 or more races ☐ Hispanic or Latino ☐ Other:
14. What is your approximate household income?
   ☐ Less than $10,000 ☐ $10,000-$19,999 ☐ $20,000-$34,999 ☐ $35,000-$54,999 ☐ $55,000-$74,999 ☐ $75,000 or more
15. Which of the following best describes your employment status?
   ☐ Employed Full-Time ☐ Retired
   ☐ Employed Part-Time ☐ Unemployed
   ☐ Student ☐ Other: ____
16. Based on your experience riding Pulaski Area Transit services, how strongly do you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Service is reliable	☐	☐	☐	☐	☐
Service gets me where I need to go	☐	☐	☐	☐	☐
Hours of service meet my needs	☐	☐	☐	☐	☐
Wait time for service is reasonable	☐	☐	☐	☐	☐
Fares are reasonable	☐	☐	☐	☐	☐
Buses are comfortable and well-kept	☐	☐	☐	☐	☐

1.2.2. Passive Data Sources

Passive, or automatic, data sources collect information using electronic sensors and cellular technology. These devices are capable of generating large volumes of highly granular data, creating a data point every second or less. Data from passive sources provide a wealth of information, typically including timestamps and enabling many levels of analysis by route, vehicle, stop and more. Although these technologies carry a higher capital cost than manual data sources, passive sources have lower marginal costs as well as increased data coverage, availability, and reliability.

Automatic Passenger Counting

Automatic passenger counting (APC) devices measure boardings and alightings at the door of a transit vehicle. As a passive data source, APC collects data whenever an equipped vehicle serves passengers. If not all vehicles are equipped, however, agencies must work to ensure that vehicles are rotated across routes so that trips are sampled adequately.

Most APC devices use a camera or infrared sensors to count how many passengers enter or exit the vehicle at an individual door. The resulting data is highly accurate, except under circumstances of miscalibration or extremely high bus load with many passengers boarding and alighting at the same time. For buses or trains with multiple passenger doors, it is important to have a device at each door to achieve a complete count. Although ridership data from APC devices is typically more reliable and comprehensive than a manual sample, agencies must periodically validate APC counts by performing manual ridechecks in order to use the data for NTD reporting. Many APC devices can also be linked to bike racks and wheelchair ramps to provide data on how often they are used.

APC data is often linked with stop location data to determine the boardings and alightings at each stop. Planners have the ability to analyze ridership levels for individual stops, trips, and routes and can identify trends over time. Due to the large volume of data, such analysis may require staff with more advanced data analysis skills.

Automatic Vehicle Location

Automatic vehicle location (AVL) devices use GPS technology to record the location of a bus at a specified time interval or its time of arrival at a particular location. While a bus is in use, an AVL device records location data nearly continuously. When used for real-time arrival information or service monitoring, these cellular-enabled devices will transmit the vehicle's location to a central server once every thirty to sixty seconds. In mountainous or remote terrain, either the GPS location of the vehicle may be inaccurate or the ability to transmit the location of t he vehicle may be impeded. Once the data is transmitted and processed, it can also be linked to customer-facing applications and webpages, giving customers the ability to see where their bus is currently located. After a vehicle ends its service for the day, more detailed AVL data can be downloaded from the vehicle for further processing and analysis. Given the high rate of data collection, AVL devices produce very large volumes of data for each bus. Typical AVL systems will produce metrics including on-time performance, runtime, and dwell time for each trip. However, more complex analyses may require staff with advanced data analysis skills. Further uses of AVL data are discussed in Chapter Four: Planning and Performance Monitoring.

Linked APC and AVL Systems

Many agencies use APC and AVL devices that are integrated by a single vendor. By linking these two data sources, agencies benefit from a streamlined data infrastructure and can access and cross-analyze the datasets together. Valuable insights arise from the association of ridership and location data. For example, passenger miles traveled can be calculated using the passenger load from APC together with the distance traveled between stops from AVL.

In combined systems, information that may only be collected by one of the devices, such as door opens and closes, wheelchair ramp deployments, and bus-kneel deployment is linked to both datasets. Additionally, staff can view the data from both devices and conduct analysis in a single location, resulting in time savings and, in some cases, reduced staff training.

Automatic Fare Collection

An automatic fare collection (AFC) system is an electronic fare system that often includes fareboxes, fare machines, and customer-facing websites and mobile applications that allow for fare transactions. With an AFC system, passengers can pay using a smart card, a mobile app, cash, or other smart fare media. These systems utilize complex datasets representing all financial transactions, fare payments, and transfers across an agency's services.

The parts of an AFC system vary greatly depending on the size of the system. A small system may consist of only electronic fareboxes and a mobile application, whereas a large system could also include hundreds of fare machines, thousands of fare gates, and a detailed website for registering, loading, and managing smart card accounts. In every case, data is automatically collected from all devices into a central database, from which queries and analysis can be conducted. Many systems provide a user-friendly interface for viewing AFC data without requiring serious data analysis skills.

Figure 7: AFC Systems

Compared to traditional fare collection methods, including cash, tokens, and paper tickets with simple fareboxes, AFC systems require lower marginal cost with high reliability and availability of data. Since the system is fully digital and mostly cashless, the majority of financial data is readily available without transporting and counting cash or tokens. Capital costs may be high for an AFC system since the entire fare system must be replaced at once. However, AFC also comes with significant savings in upkeep since most transactions are automatic and digital, requiring no exchange of physical fare media.

AFC data may include a wide array of data points to represent all fare types and all fare media, including every transaction and usage of fare with the corresponding fare amount. Such transactions include loading a smart card, purchasing a single fare or fare pass, using a pre-purchased fare, making a transfer, exiting a fare gate, or paying with cash. In addition to the transaction type and amount, AFC data captures the time, vehicle or fare gate, route, account number, and in some cases, the location of the transaction.

The uses for AFC data extend from financial analysis, such as revenue measurements and financial efficiency of services, to advanced insights on how passengers use the service, such as transfer analyses and ridership behaviors of anonymous individuals. As with manually collected farebox records, AFC data is also often used to validate ridership counts.

AGENCY PROFILE: FREDERICK TRANSIT TECHNOLOGY ADOPTION

FREDERICK TRANSIT HAS IMPLEMENTED APC, AVL, AND MOBILE TICKETING SYSTEMS, IN ADDITION TO MOVING THEIR FIXED-ROUTE BUS FLEET OVER TO ELECTRIC VEHICLES

Context

Over the past several years, TransIT, a small urban agency in Frederick, Maryland, has implemented a number of technology upgrades to improve the passenger experience, data collection, and data reporting. These upgrades included installing automated passenger counters (APC) and automatic vehicle location (AVL) devices on its fixed-route fleet, launching a new mobile ticketing app, and purchasing nine fully electric vehicles.

Resources Needed

TransIT's technology upgrades required financial and personnel resources. Federal Transit Administration (FTA) capital grants were used to purchase hardware for the APC and AVL systems and electric vehicles. The implementation process also required personnel resources on a number of dimensions. Operators needed training on how to log into new tablets installed in each vehicle for the APC/AVL system and how to record fares on tablets properly. Data analysts needed to learn how to use and download data from Route Match software, while other staff needed to learn how to properly program Route Match software to include the agency's schedule. As part of the purchase, APC/AVL vendors provided IT support, and county IT staff also helped with troubleshooting.

Results

TransIT's technology initiatives transformed the way the agency collected and analyzed data and has vastly improved the customer experience.
Vehicle tracking and schedule adherence have been automated with the AVL system. The entire fixed-route fleet was equipped with AVL, allowing passengers to see where their bus is located and its estimated arrival via the Route Shout mobile app.

Ridership reporting is also streamlined. The AVL system is used to track bus locations and for the drivers to record fares collected, from which overall ridership is derived and reported to the National Transit Database (NTD). The APC system is used for detailed data analysis at the stop level.

Mobile ticketing has improved the collection of fares and made purchasing fares easier for riders. It has also improved fare reconciliation, as reports the vendor, Token Transit, provides fare reports that can be compared to what drivers have entered into their tablets.

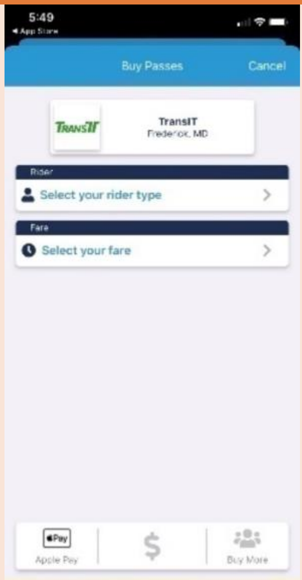


Figure 8: TransIT Mobile Fare Payment App

Results (cont.)

Data collection for the AVL and fare systems are linked through the drivers' tablets. Drivers enter fares into the tablets that were provided with the AVL system, and then all of this data is able to be accessed by agency staff.

While TransIT's electric vehicles have cut down on fuel costs, they do not currently have any dashboards for reporting metrics on the fleet. The manufacturers for their electric vehicles, Complete Coach Works and BYD, have partnerships with external vendors that provide dashboard reporting, but TransIT does not currently use them.

Lessons Learned

TransIT staff noted several lessons learned throughout the process of implementing these new technologies. They include:

- Ensure that vendors understand the data formatting required for NTD reporting
- Review platforms and user interfaces before purchase to ensure they are intuitive and easy to use for drivers
- Keep some system redundancy for the first year when installing new technologies to help with data validation and unexpected issues with the new systems
- Consider vehicle requirements when switching from diesel to electric vehicles.
    - Due to charging limitations, one-to-one replacements of vehicles often cannot be made. TransIT uses their electric vehicles during the peak periods and recharges them at their garage during the midday period. Careful attention needs to be paid to charging times and battery life, particularly during the times of the year when HVAC systems will be operating.

Key Takeaways

- Technology upgrades are essential to increasing the accuracy of an agency's data and increasing data collection efficiency.
- Different technologies can be integrated to improve data validation and ease of use.
- New technologies should be vetted before purchase. Redundant systems should be in place during implementation of new technologies to ensure data consistency and system reliability.

For More Information

National Center for Mobility Management: Frederick County's TransIT: Embracing Technology to Maximize Mobility Management

Frederick TransIT

1.3. Emerging Advancements

While sensor-based, passively generated data is increasingly part of the state of the practice for transit agencies, emerging technological advances are changing how such data is generated from internet-connected parts and systems to customer cellphone location data, to the data generated as agencies provide more customized services.

Internet of Things (IoT) technologies

A detailed asset inventory is the foundation of transit asset management, but maintaining this data can become more time-consuming as fleets grow in size or requirements for data collection increase. Internet of Things (IoT) devices are internet-connected, physical equipment that can relay information about their status or be controlled remotely. A handful of large urban transit agencies like MARTA in Atlanta are beginning to incorporate IoT-based data collection for transit asset management, and this trend may eventually spread to other agencies.[4]

IoT devices are equipped with sensors that relay information about vibrations, pressure, heat, power usage, and other conditions to central servers. For some IoT-equipped facilities, these devices can be used to better regulate building energy costs from heating, air conditioning, and lighting. For vehicles and other transit assets, IoT data can then be used to predict the health of the asset and inform capital investment decisions. In extreme cases, these signals can also be used to signal a need for preventative maintenance before the asset fails. The introduction of sensors is not without costs: sensors must be installed, calibrated, and maintained, and processing the data from these sensors will require software licenses that may be cost-prohibitive to agencies, especially for relatively small fleets or facilities. As applications of IoT continue to develop in the transit industry, the potential for small urban, rural, and tribal agencies is likely to continue to improve.

Many large and small agencies with battery-electric buses have already entered the realm of IoT-enabled assets. Electric buses automatically collect and transmit data such as battery levels, power consumption, and battery health in real-time to allow dispatchers to maintain a sufficient charge on each vehicle. Many buses also keep a full record of vehicle speed and braking, which provide insight into the most power-efficient driving styles. With real-time access to these metrics, agencies can move toward intelligent dispatch of electric vehicles based on their battery levels and the availability of charging infrastructure. Moreover, agencies are learning how to best procure internet-enabled assets and implement sophisticated data management practices. As the transit industry explores complex IoT technologies like self-driving vehicles and connected infrastructure, the systems and processes developed for battery-electric buses are opening doors for agencies of all sizes.

Aggregate Cellphone Location Data

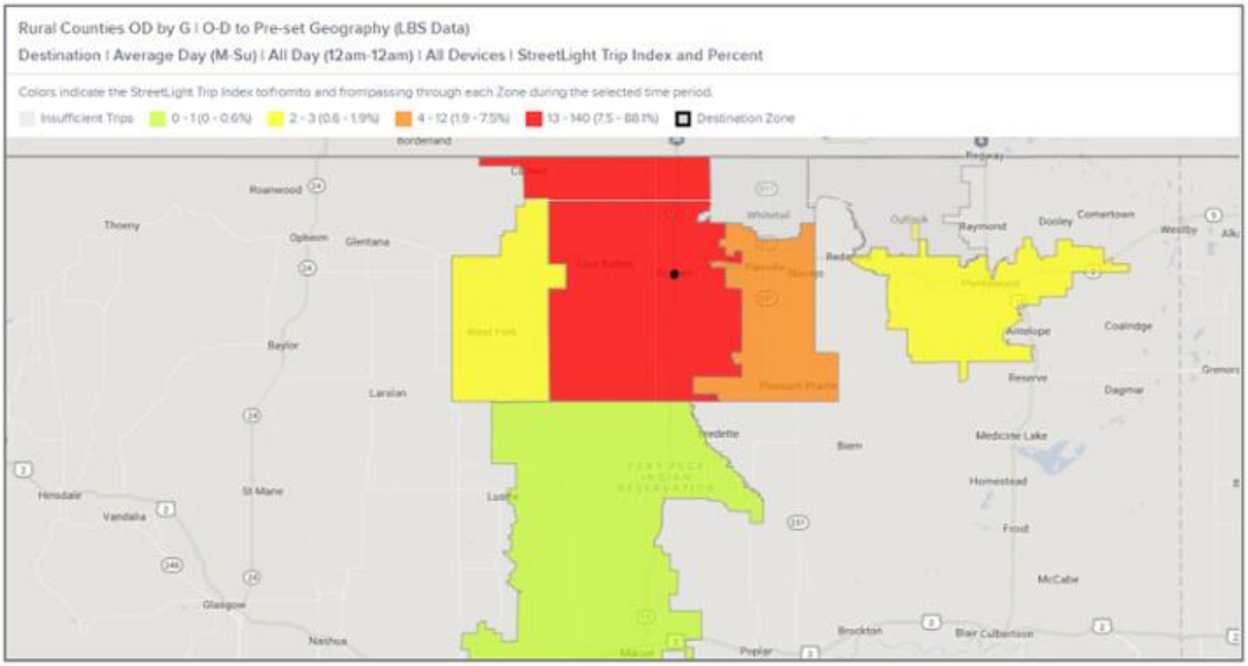
While on-board customer surveys can provide rich data on who riders are and how they use a transit system, the manual data collection effort they require mean that survey data is infrequently gathered, limited in the extent of routes or times that are covered, and expensive to procure. Especially as travel patterns have shifted drastically during the Covid-19 pandemic, with many rural areas seeing an influx of urbanites and normal work travel disrupted, having up-to-date and detailed travel data can be useful to transit agencies.

One of the solutions to this challenge pursued by some larger agencies is to use aggregate cellphone location data purchased from a vendor such as Streetlight or AirSage. These data providers obtain,

[4] Keenan, Sean Richard. "MARTA is turning to the Internet of Things to lower costs, reduce system hiccups." Curbed Atlanta. https://atlanta.curbed.com/2019/2/28/18243947/marta-ibm-transit-internet-things-transportation. Accessed January 13, 2021.

aggregate, and anonymize Location-Based Services (LBS) data. LBS data consists of the location 'pings' generated by cellphones when certain apps request location information from a phone, as well as other Global Positioning System (GPS) and cellphone triangulation data. Though rural areas may have incomplete cellular coverage, the wide array of location data sources that feed into LBS datasets can help to correct for gaps in coverage; an example of LBS data as processed by Streetlight is shown in **Figure 9**. In addition to LBS data, certain transit trip planning apps such as Moovit and Transit offer datasets on trip itineraries that users select in addition to their locations.

Figure 9. Example of Location-Based Services Data in Montana



For transportation providers, these platforms provide information on origins and destinations and travel speeds on roadways. In some cases, the mode of transport and demographic characteristics of travelers can also be inferred. Because of sample size and privacy concerns, this data is often not available for rural areas at the fine-grained level of detail of urban areas, but these datasets may be useful for identifying volumes of travel between rural towns for long-distance services. State departments of transportation and regional planning bodies may be able to obtain this data for your service area. Additional uses of origin-destination data are discussed in Chapter Four: Planning and Performance Monitoring.

1.4. Conclusion

Data sources are the fuel for data practices that inform planning decisions, foster strong partnerships, and answer critical questions about access and equity. These sources range from a variety of manually generated sources—ridechecks, safety records, and customer feedback—to automatically generated, often sensor-based datasets. Automatic or passive datasets like automatic passenger counters are commonly found among larger agencies and may steadily become more available for small urban, rural, and tribal agencies as well. On the horizon are new sources of data like aggregate cell phone location data that may eventually become part of the state of the practice. Use the checklist on the next page to assess what data sources your agency currently has and what it needs to consider obtaining.

Checklist: What are your agency's available data sources?

Manual Data Sources	Description	Currently Use This Data Source	Want to Obtain This Data
Manual ridechecks	Sampled boarding and alighting counts by trip and stop	☐	☐
Trip requests and trip logs	Demand response trip information recorded by drivers, including riders, distance, and fare payment	☐	☐
Asset management and maintenance data	Asset inventory and condition data	☐	☐
Safety incident reports	Information recorded after safety incidents	☐	☐
Customer service complaints	Customer feedback or ratings	☐	☐
On-board customer surveys	Sampled data on trip characteristics and passenger information	☐	☐

Passive Data Sources	Description	Currently Use This Data Source	Want to Obtain This Data
Automatic Passenger Counting	Boarding and alighting counts for every stop	☐	☐
Automatic Vehicle Location	Location data for a bus, associated with a route and trips	☐	☐
Automatic Fare Collection	Automatically generated fare transaction data	☐	☐
Transit smart card data	Transaction and travel data associated with individual accounts	☐	☐

Checklist: What are your agency's available data sources? (Cont.)

Emerging Data Sources	Description	Currently Use This Data Source	Want to Obtain This Data
Internet of Things (IoT) devices	Sensor-equipped physical devices that relay device condition data	☐	☐
Aggregate cellphone data	Origin-destination and travel speed data generated from cellphone location data	☐	☐

For automatic sources, data formats and standards are important for storing data consistently and translating data between formats to meet particular uses. In the next chapter, we'll examine common data sources and their importance to an agency's data practices.

Key Resources on Data Sources

- MIT Data Collection Techniques and Program Design
- NTD Sampling Manual
- TCRP 113: Using Archived AVL-APC Data to Improve Transit Performance and Management
- TCRP 177: Preliminary Strategic Analysis of Next Generation Fare Payment Systems for Public Transportation
- Public Transit Rider Origin Destination Survey Methods and Technologies
- Customer Surveying for Public Transit: A Design Manual for On-Board Surveys

2. DATA STANDARDS

Having good data is the first step to having better tools for planning, reporting, and providing information to customers, but how agencies handle and process data matters greatly. Data standards provide a way for organizations to store and use data in a logical and accessible way, helping agencies with internal reporting and facilitating dissemination of data to the public. This fact sheet provides an overview of established and emerging data standards that can be useful to small agencies and provides examples of how these standards are being implemented in small agencies around the country.

> Data standards can help agencies disseminate data to the public, plan for future needs, or report on performance metrics.

Existing Standards

GTFS Schedule Data

General Transit Feed Specification (GTFS) has become the de facto standard for tracking and disseminating schedule data. It can be used in Google Transit as well as many other transit route planning applications.

Real-Time Arrivals

In addition to schedule data, GTFS-RT provides real-time updates to riders, like stop closures, route changes, or bus delays.

Transit Asset Management

Transit Asset Management tools provide standardized procedures for tracking the condition of vehicles and other transit assets.

Ridership Data

Transit Asset Management tools provide standardized procedures for tracking the condition of vehicles and other transit assets.

Internal Standards

Creating standardized methods to track internal data can help agencies better use the data they already produce.

Emerging Advances

GTFS-Flex

GTFS-Flex is an extension to GTFS that provides information about flexible transit services, including demand-response and deviated fixed-route services.

GTFS Validators

Inaccuracies and errors in GTFS feeds impede their usefulness. GTFS validators test GTFS feeds for a variety of feed issues, and some validators also test for data quality.

Intelligent Transportation Systems (ITS)

ITS systems, including AVL, APC, and AFC, are becoming more common ways to collect ridership data and performance metrics. Standards can help to facilitate the use of this data.

Demand-Response Standards

TCRP has published data standards that, if adopted, could make collecting, sharing, and coordinating data across demand-response service providers more seamless.

2.1. Introduction

Having good data is the first step to having better tools for planning, reporting, and providing information to customers. Standards can be either formal data practices developed by industry leaders that allow for consistency among vendors and agencies, or standards can be less formal data management practices defined at each agency to ensure internal consistency. Unlike standard operating procedures or other process guidelines, the data standards and management practices here inform the structure and format of data. Because they allow users and applications to process important information quickly, standards are a key determinant of how much agencies–as well as customers and stakeholders–can learn from what they have collected.

General Transit Feed Specification (GTFS) feeds are the canonical example of data standards in transit. Because of the standardized format, trip-planning applications can help customers plan trips for any transit service that publishes a GTFS feed to consume agency schedule data regularly. Validator applications can also help agencies catch potential errors in their GTFS feed files. For other types of data, agencies may need to develop their own internal standards. Internal agency standards can help ensure that staff from different departments know about available datasets, communicate about what they need from the dataset, and contribute to their creation and upkeep. Formalized data structures, which remain consistent across time, can lead to faster, more accurate, and easier reporting, as the setup required to process the data can be reused in the future.

Trends in the industry include moving towards formats that allow for real-time data to be disseminated to customers and include flexible transit data in trip planning applications. Extensions to GTFS feeds can allow agencies to do both. GTFS-RT builds on the GTFS framework to provide real-time data like bus delays, temporary stop closures, or other pertinent information to riders. GTFS-Flex extends the GTFS framework to flexible transit, like demand-response or deviated fixed-route services. Also, ridership data collected from Automatic Passenger Counts (APC) or Automatic Fare Collection (AFC) data is beginning to be used more widely, including among small and rural transit agencies, and combined with Automatic Vehicle Location (AVL) data. Converging around standards for APC, AFC, and AVL data will allow agencies to process and disseminate this data much more easily. See Chapter 3: Open Data for more information about collecting data.

This chapter provides an overview of existing standards, emerging standards, and internal methods for improving data management that small urban, rural, and tribal agencies can leverage to use the data available to them better. As defined here, data standards are uniform methods for collecting, storing, and processing data. This chapter also provides several examples of how data standards and practices are being implemented in small and rural agencies around the country.

Collecting and managing data can take time and effort. Agencies can choose to do this in-house, which may require significant staff time and knowledge, or manage data using proprietary software, which may be more flexible but with a higher up-front cost.

2.2. State of the Practice

2.2.1. GTFS Schedule Data

General Transit Feed Specification, or GTFS, is the de facto standard for disseminating transit schedule data in the United States. Before GTFS was developed, each transit agency was responsible for publishing schedules in PDF format on their website or assisting riders who called a helpline, creating barriers for riders to find up-to-date information. Since multiple transit providers might exist in the same area, riders might also have to navigate multiple websites, schedules, or help lines. In addition, travelers or tourists might not be familiar with the local transit providers and might not know where to look for information.

To address these concerns, Google and TriMet, a transit agency in Portland, worked together to create the first GTFS feed in 2005.[5] They used TriMet's data and created a prototype of the GTFS feed, aiming to make the data format easily replicable so that it could be widely used. The ability to provide customers with journey planning through the popular Google Maps app and open source tools like OpenTripPlanner led to the steady adoption of the GTFS format.

GTFS feeds are produced in a standardized way using schedule and geographic data. The file structure is relatively simple: it is a collection of tables stored as comma-separated text files in a zipped folder. The tables are a decomposition of transit schedules into their constituent parts—stops, routes, trips, and so forth—with the relationships between these tables defined through common ID fields. This structure allows for efficient storage of large amounts of information. GTFS feeds require certain files and fields, while other files are optional, as shown in **Table 2**.[6] GTFS was designed to use tables to ensure that it could be edited and used with a variety of programs and coding languages.[7] Google provides resources to help agencies create GTFS feeds and submit them for use in Google Maps.[8]

Table 2: Required and Optional GTFS Files

Required GTFS Files	Optional GTFS Files
Agency.txt	Fare_attributes.txt
Stops.txt	Fare_rules.txt
Routes.txt	Shapes.txt
Trips.txt	Frequencies.txt
Stop_times.txt	Transfers.txt
Calendar.txt or Calendar_dates.txt	Pathways.txt
Feed_info.txt	Levels.txt
	Translations.txt
	Attributions.txt

Standardizing the data used to provide schedule information in Google Maps allowed other agencies to take advantage of this new technology by providing their data in the same format. Since GTFS feeds were designed with simplicity in mind, this allowed agencies to access to powerful trip planners like Google Maps or to deploy an agency-deployed trip planner at a relatively low cost. The open data format allows Google and other application developers to use this data to create journey-planning applications easily.

[5] Bibiana McHugh. 2013. "Chapter 10: Pioneering Open Data Standards: the GTFS Story" *Beyond Transparency*. https://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story/
[6] Google. "Static Transit: Reference" Google Transit APIs. https://developers.google.com/transit/gtfs/reference. Accessed January 20, 2021.
[7] McHugh, "Chapter 10"
[8] Google. "GTFS Static Overview." Google Transit APIs. https://developers.google.com/transit/gtfs#SubmitFeedToGoogle. Accessed January 20, 2021.

Today, thousands of agencies worldwide publish GTFS feeds.[9] GTFS data enables easy trip planning for transit riders, improving customer information.[10] It allows agencies of all sizes to take advantage of cost savings by relying on third-party application developers. In addition, GTFS allows for analyses of accessibility and other performance metrics, such as route frequency. The original GTFS feed format, which published schedule data for fixed-route services, is being expanded to other applications, like providing real-time arrival information and information about flexible transit services (see the next section, Emerging Advancements).

RTAP GTFS Builder: Producing GTFS Feeds for Rural and Tribal Agencies

Although GTFS feeds were designed to be simple to produce, generating a feed can still pose a burden to small agencies that might not have the specialized software to generate GTFS feeds or staff with technical abilities to operate such software. To aid rural and tribal agencies, the National Rural Transit Assistance Program (RTAP) created a GTFS Builder that consists of two Excel files. Users input schedule data (routes, stops, stop order, stop times, calendar, etc.) into one spreadsheet and then run a macro that generates the schedule into GTFS format. The schedule GTFS data is then copied into the second file called ImportExport and additional data such as agency information, fares and service dates are added. Finally, the entire ImportExport file is exported into a complete GTFS data feed. GTFS Builder works well with free applications such as Google's MyMaps and Google Earth to geocode bus stops and draw route shapes (see Agency Profile: Crawford Area Transportation Authority Use of the RTAP GTFS Builder below).

AGENCY PROFILE: CRAWFORD AREA TRANSPORTATION AUTHORITY USE OF THE RTAP GTFS BUILDER

THE CRAWFORD AREA TRANSPORTATION AUTHORITY (CATA) DEPLOYED A GTFS FEED IN OCTOBER 2019 USING THE RTAP GTFS BUILDER.

Context

Crawford County and Venango County, Pennsylvania consolidated their transit services in 2016, with Venango County Transit merging into the Crawford Area Transportation Authority (CATA). South of the city of Erie, Crawford County borders Ohio and is centered on the city of Meadville. Venango County is just to the southwest of Crawford County. The agency operates 13 fixed routes, including three in the Meadville area, one in the Titusville area, four in Venango County, and an additional five in rural areas of Crawford County (known as "lifeline" routes). Lifeline routes generally operate only two to three days per

[9] As of January 27, 2021, 1,327 providers are listed on OpenMobilityData and over 2,500 providers are listed on Transitland.

[10] Transportation Research Board. Pending publication. "TCRP G-18: Improving Access and Management of Transit ITS Data." https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4687

Context (cont.)

week, whereas other routes typically operate Monday through Saturday except for the Titusville route, which operates Monday through Friday.

The consolidation of services required a merging of both agencies' datasets, which were in very different formats. To accurately merge data, create a better inventory of stops, and improve trip planning for customers, the agency created a GTFS feed, launched publicly in October 2019. Staff utilized the National Rural Transit Assistance Program's (National RTAP) GTFS Builder to create the GTFS feed and created a new stop inventory and route alignments in Google Earth.

National RTAP's GTFS Builder consists of two Excel files. Users input schedule data (routes, stops, stop order, stop times, calendar, etc.) into one spreadsheet and then run a macro that generates the schedule into GTFS format. The schedule GTFS data is then copied into the second file called ImportExport, and additional data such as agency information, fares and service dates are added. Finally, the entire ImportExport file is exported into a complete GTFS data feed. GTFS Builder works well with free applications such as Google's MyMaps and Google Earth to geocode bus stops and draw route shapes.

Similar to other GTFS applications, once the GTFS feed is exported from GTFS Builder, agencies selecting Google Maps as a trip planner will need to sign up, sign a license agreement and create a partnership account with Google to upload the feed. National RTAP offers an in-depth review of a [sample license agreement](). The feed then undergoes two steps of validation, one to verify that schedules are accurate and match with what is on the agency's website, and another to make sure the tables internal to the feed are cross-validated and there are no missing data points or major errors.

Figure 10: The RTAP GTFS Builder in Use

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Routes	route_id	service_id	direction_id	trip_headsign	stop_id	stop_desc	stop_headsign	pickup_type	drop_off_type	timed_stop	wait_time	timepoint	wheelchair_accessible	bikes_allowed	T01	T02	T03	T04	T05
124	Bus Route: BlueA_Weekday													1	1	BLUEA	BLUEA	BLUEA	BLUEA	BLUEA
125		BlueA	mtwtf	1	DOWNTOWN MALL	1	Downtown Mall	CATA Office Meadville			1		1	1	1	6:30	7:00	8:00	9:00	10:00
126		BlueA	mtwtf	1		4	CATA Office Meadville	Active Aging			1		1	1	1	6:32	7:02	8:02	9:02	10:02
127		BlueA	mtwtf	1		5	Active Aging	Holland Towers			1		1	1	1	6:32	7:02	8:02	9:02	10:02
128		BlueA	mtwtf	1		6	Holland Towers	Fairview/Fairmont Apts			1		1	1	1	6:33	7:03	8:03	9:03	10:03
129		BlueA	mtwtf	1		34	Willow St & Park Ave	Fairview/Fairmont Apts						1	1	*	*	*	*	*
130		BlueA	mtwtf	1		105	Willow St & S Cottage Ave	Fairview/Fairmont Apts						1	1	*	*	*	*	*
131		BlueA	mtwtf	1		106	S Main St & Davis St	Fairview/Fairmont Apts						1	1	*	*	*	*	*
132		BlueA	mtwtf	1		107	S Main St & Poplar St	Fairview/Fairmont Apts						1	1	*	*	*	*	*
133		BlueA	mtwtf	1		108	Poplar St & Liberty St	Fairview/Fairmont Apts						1	1	*	*	*	*	*
134		BlueA	mtwtf	1		109	Poplar St & East Al	Fairview/Fairmont Apts						1	1	*	*	*	*	*
135		BlueA	mtwtf	1		110	Poplar St & Fairmont Ave	Fairview/Fairmont Apts						1	1	*	*	*	*	*
136		BlueA	mtwtf	1		111	Fairview/Fairmont Apts	Meadville Medical Center Grove Street			1		1	1	1	6:36	7:06	8:06	9:06	10:06
137		BlueA	mtwtf	1		112	Meadville Medical Center Grove Street	Forest Green Apts			1		1	1	1	6:38	7:08	8:08	9:08	10:08
138		BlueA	mtwtf	1		113	Pine St & Spencer Alley	Forest Green Apts						1	1	*	*	*	*	*
139		BlueA	mtwtf	1		114	Pine St & East St	Forest Green Apts						1	1	*	*	*	*	*
140		BlueA	mtwtf	1		115	Liberty St & Poplar St	Forest Green Apts						1	1	*	*	*	*	*
141		BlueA	mtwtf	1		116	Liberty St & Willow St	Forest Green Apts						1	1	*	*	*	*	*
142		BlueA	mtwtf	1		117	Liberty St & Lamont Dr	Forest Green Apts						1	1	*	*	*	*	*
143		BlueA	mtwtf	1		118	Liberty (Near Ranz)	Forest Green Apts						1	1	*	*	*	*	*
144		BlueA	mtwtf	1		119	Liberty St & Arthur St	Forest Green Apts						1	1	*	*	*	*	*
145		BlueA	mtwtf	1		120	Forest Green Apts	Morgan Village Dr			1		1	1	1	6:42	7:12	8:12	9:12	10:12
146		BlueA	mtwtf	1		121	Graff St & Talon St	Morgan Village Dr						1	1	*	*	*	*	*
147		BlueA	mtwtf	1		122	Meadowland Dr & Neason St	Morgan Village Dr						1	1	*	*	*	*	*
148		BlueA	mtwtf	1		123	Williamson Rd & St. Brigids Cemetery	Morgan Village Dr						1	1	*	*	*	*	*
149		BlueA	mtwtf	1		124	Williamson Rd & Nealson Hill Elementary	Morgan Village Dr						1	1	*	*	*	*	*
150		BlueA	mtwtf	1		125	Morgan St & Ravine St	Morgan Village Dr						1	1	*	*	*	*	*

Notes | Schedules V3 | stop_times | trips | local_routes | local_calendar | Local_stops | timed_stops

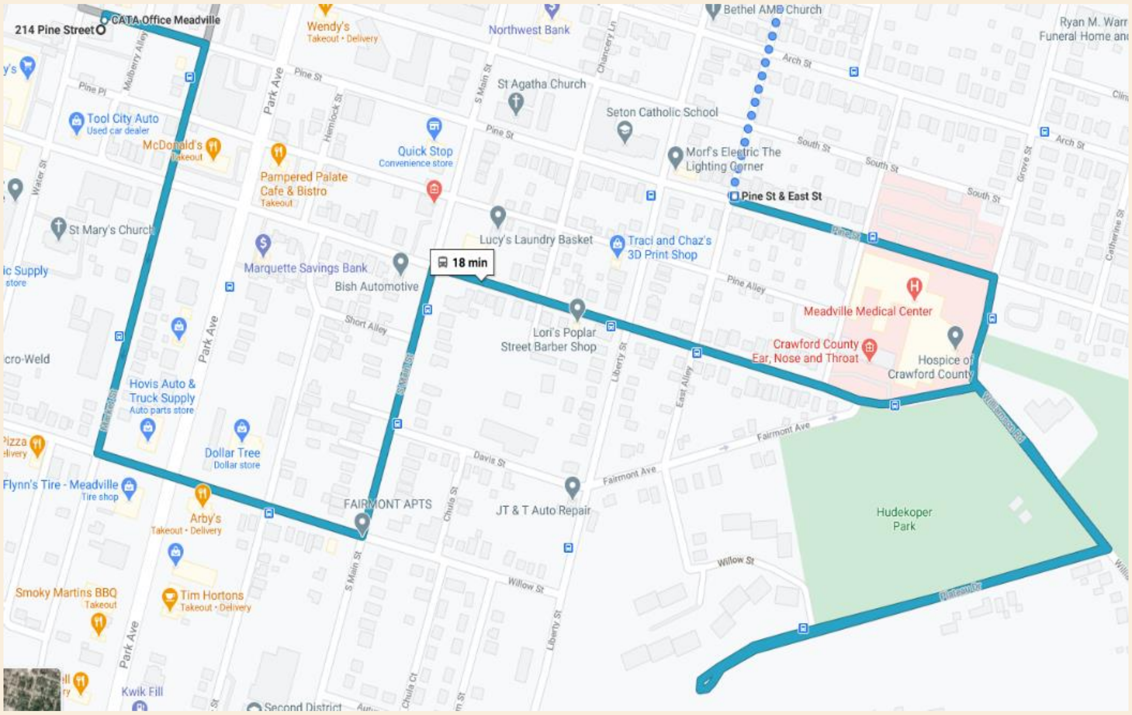
Resources Needed

Transit agency staff need to create a National RTAP cloud account and log in to download the Builder, and then begin using it. Before inputting agency data into the Builder, users need to create the route and stop spatial data in a GIS program or Google Earth. They also need to have the schedules by route and direction with stop times constructed. From there, the data can be entered into the Builder intuitively. An RTAP staff member is dedicated to helping agencies use the Builder.

CATA staff spent approximately four months creating their first GTFS feed using Google Earth and the Builder. The agency did not have an accurate bus stop inventory nor any spatial files for routes, so planning staff and operations staff had to first inventory stop locations and confirm route alignments. From there, a planning staff member was able to draw route alignments in Google Earth and set bus stop locations to get coordinates for each stop. With the help of RTAP staff answering questions about the Builder, this staff member was then able to build the bus stop inventory and GTFS feed, have it validated with Google, and update it whenever service changes are made.

There are no special skills required to use the Builder, other than a basic working knowledge of Excel and Google Earth. Users do not need to know how to build macros, and RTAP staff can help with troubleshooting issues that arise in the process. CATA staff mentioned that organizational skills are helpful when using the Builder, as data needs to be well organized even before entering it into the Builder's spreadsheets. In particular, stop ID's and stop names should be consistent across routes; before going through this process, CATA's stop ID's differed between routes, which would have led to issues when creating the feed with the RTAP GTFS Builder.

Figure 11: Routing Options



Results

The use of the Builder and the creation of the GTFS feed brought a number of positive results and changes to the agency, such as:

- The agency was able to build a trip planner into their website that passengers can now use to plan their trips. Passengers can also use Google Maps in a web browser or in a mobile app to plan trips.
- The feed and trip planner's creation drastically reduced the number of phone calls to the agency requesting trip planning assistance. In the first year of the GTFS feed being live, over 1,000 people used the trip planner on their website.
- The feed's creation also forced the agency to better organize its data, particularly stop locations and route alignments. Updates to the GTFS feed are now very simple to make.
- The creation of the feed also had several unforeseen positive outcomes, one of which was simplifying the creation of route brochures. Now that the agency has stops and route alignments in Google Earth, which they can send to their brochure maker instead of having to sketch out route alignments and stop locations manually.

Lessons Learned

CATA staff noted some lessons learned when creating their GTFS feed using the RTAP GTFS Builder. They include:

- Organizing stop locations, stop ID's, and stop names before using the Builder
- Making sure the times assigned to stops on your agency's website are accurate
- Using National RTAP staff as a resource as much as possible, particularly when running into error messages in the validation process with Google
- National RTAP staff can help clarify new terms, expedite best practices, and review data input into GTFS Builder.

Key Takeaways

- Creating a GTFS feed is a great way to organize your agency's schedule data to provide trip planning services for your customers.
- The RTAP GTFS Builder provides a free framework for organizing a GTFS feed that only requires basic Microsoft Excel skills and some beginner knowledge of Google Earth.
- Small agencies can create GTFS feeds with their existing staff and do not need to hire outside specialists. Once feeds are created, they are easy to maintain using the Builder and can lead to new efficiencies with creating schedule brochures and other tasks.

For More Information

CATA and Venango County Transit

RTAP GTFS Builder

Google Earth

GTFS Overview from Google

Demystifying Trip Planner Licensing Agreements

2.2.2. Real-Time Arrivals

A GTFS feed can help the customer plan their trip but will not help them know whether their bus will be on time–or even if it has been temporarily rerouted away from their stop. GTFS-Realtime (GTFS-RT) is an extension to a traditional GTFS feed that attempts to solve this problem. It includes: trip updates, like delays, cancellations, or route changes; service alerts like moved stops; and vehicle positions, including location and congestion. Since it provides real-time information, it must be updated frequently—up to once every five seconds—using Automatic Vehicle Location (AVL) data.[11] Google Maps currently supports GTFS-RT in Boston, San Francisco, San Diego, Portland, Madrid, and Torino and provides resources for other agencies to construct and deploy GTFS-RT.[12]

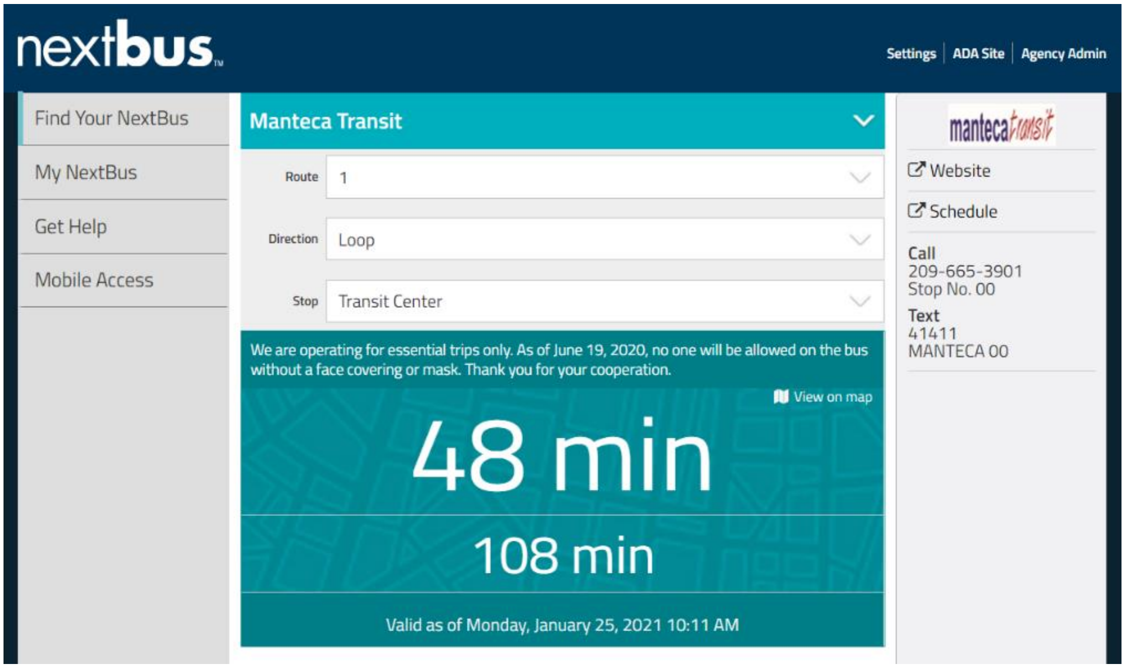
In addition to Google's format, other developers have created open-source or proprietary formats to provide real-time bus information that use GPS or AVL data to locate buses. OneBusAway is an open-source example that provides real-time information, available through the OneBusAway app or with the agency's own branding.[13] TheTransitClock is another open-source system that converts AVL data to real-time information, operating in the Twin Cities in Minnesota.[14] NextBus is a proprietary software owned by Cubic that provides similar services, offering a web and mobile application that provides real-time transit information.[15]

NextBus is the most widely used, providing real-time arrival information for approximately 75 systems, including small urban and rural transit systems like Manteca Transit in California, Winston-Salem Transit Authority and Greenville Area Transit in North Carolina, EZRide -Charles River Transportation Management Association (TMA) in Massachusetts, and AppalCART in North Carolina. NextBus is a logical choice if an agency already uses other Cubic services. OneBusAway and TheTransitClock are both open-source and can be deployed in-house with an IT team or with assistance from a nearby university or independent developer.

---

[11] TRB, "TCRP G-18."
[12] Google. n.d. "Create a GTFS-realtime feed." *Transit Partners Help*. Accessed January 20, 2021. https://support.google.com/transitpartners/answer/6368391
[13] OneBusAway. n.d. "Homepage." Accessed January 20, 2021. https://onebusaway.org/
[14] TheTransitClock. "Homepage." Accessed January 20, 2021. https://thetransitclock.github.io/
[15] Cubic. n.d. "NextBus: Real-Time Passenger Information. Accessed January 20, 2021. https://www.cubic.com/solutions/transportation/nextbus, https://www.nextbus.com/#!/ttc/5/5_1_5B/1405

Figure 12: Manteca Transit's NextBus App



2.2.3. Transit Asset Management Data

Safe and reliable service depends on agencies maintaining their transit vehicles and assets in a state of good repair, but the data to track asset condition is often difficult to manage. While there is no formalized standard for tracking this information, the Federal Transit Administration (FTA) has published Transit Asset Management (TAM) guidelines to help agencies organize and manage their data and produce TAM Plans that meet federal requirements.[16] The TAM model recommends tracking a variety of attributes related to assets, such as age, expected life, general attributes, model, available parts, utilization, condition, defects, inspection dates, investment, and life cycle costs.

For small urban, rural, and tribal reporters, reporting on asset condition is often facilitated by other agencies, but tracking this data nonetheless falls to individual agencies. Asset information can readily be tracked in a spreadsheet or in a database, but many agencies will use purpose-built Decision Support Tools (DST) maintained by state or federal governments. The FTA provides a number of resources for agencies to produce TAM plans. These include a TAM Plan Excel template for small providers and a Microsoft Access-based tool called Transit Economic Requirements Model Lite ("TERM-Lite") that is designed to help agencies track and forecast annual investments needed to maintain assets in a State of Good Repair (SGR).[17] In addition, FTA provides specific assistance for tribal agencies through TAM for Tribes.[18] This assistance outlines the requirements for reporting and the possibilities for sponsoring or joining a group for reporting requirements.

---

[16] Federal Transit Administration. 2020. "Transit Asset Management." Accessed January 20, 2021. https://www.transit.dot.gov/TAM, https://www.transit.dot.gov/sites/fta.dot.gov/files/2020-10/TAM-Systems-Handbook.pdf

[17] Federal Transit Administration. 2015. "TERM-Lite." https://www.transit.dot.gov/TAM/TAMPlans/SmallProviderV2_0, https://www.transit.dot.gov/TAM/TERMLite

[18] Federal Transit Administration. 2020. "TAM for Tribes." Accessed January 20, 2021. https://www.transit.dot.gov/TAM/Tribes

Tracking TAM data not only helps with maintenance and performance monitoring but can also help with both short-term operational needs and long-term strategic planning. This information can then be used by a variety of staff, including maintenance staff, accounting staff, or for budget and planning purposes. Further methods of analyzing transit asset data are discussed in Chapter 4: Planning and Performance Monitoring.

Alabama DOT Rural TAM System

The Alabama Department of Transportation uses a TAM system to track vehicles. They track 40 variables across five main criteria: engine issues, running condition, interior condition, air conditioning, wheelchair lift operation, exterior condition, and mileage. In addition to tracking assets, they also use a prediction model based on linear regression to help plan maintenance needs based on the projected budget. These efforts also help the agency with reporting.

Figure 13: Example TAM Spreadsheet, Akron METRO Bus[19]

Capital and Fleet Costs	Total Cost	METRO Cost	State Contribution	Federal Contribution	2013	2014	2015
Primary Transit Corridors	Total Cost	50%	10%	40%			
Line A - West Market	**$21,689,265**	$10,845,000	$2,169,000	$8,676,000	$0	$2,820,785	$2,877,200
Line B - Arlington	**$13,286,000**	$6,643,000	$1,329,000	$5,315,000	$0	$0	$0
Line C - Copley	**$16,770,000**	$8,385,000	$1,677,000	$6,708,000	$0	$0	$0
Line D - Kenmore	**$29,236,600**	$14,619,000	$2,924,000	$11,695,000	$0	$0	$0
Line E - Howard/State	**$8,554,000**	$4,277,000	$856,000	$3,422,000	$0	$0	$0
Line F - Twinsburg-Macedonia	**$10,133,000**	$5,067,000	$1,014,000	$4,054,000	$0	$0	$0
Copley Extension	**$12,164,500**	$6,083,000	$1,217,000	$4,866,000	$0	$0	$0
Norton Extension	**$9,616,500**	$4,809,000	$962,000	$3,847,000	$0	$0	$0
Stow Extension	**$11,982,500**	$5,992,000	$1,199,000	$4,793,000	$0	$0	$0
Unit cost CNG	**$460,000**				$3,284,400	$0	$0
Unit cost MCI	**$600,000**				$0	$0	$0
Unit cost Artic	**$850,000**				$0	$0	$0
Total Capital Cost (Infrastructure + Fleet)					$3,284,400	$2,820,785	$2,877,200

2.2.4. Ridership Data

For many small transit agencies, drivers wear many hats. In addition to safely operating the vehicle and providing excellent customer service to passengers, many must also keep track of ridership and fares. When ridership is tracked manually, it is important to have standards in how it is collected, formatted, and summarized. Ridership data collection on fixed-route services should use pre-populated, standard forms developed by the agency that include the scheduled trip start time, the vehicle used, and stop name and ID as relevant. Columns to record the number of passengers boarding and alighting (along with any special categories of passenger types to note) and the time each stop is departed should also be included. For systems with flag stops, popular boarding and alighting locations should be pre-listed with sufficient space before and after them for other boarding and alighting locations to be listed. Pre-populated forms reduce the likelihood that drivers make errors in recording stop locations or trip start times and allow them to focus on providing excellent customer service and safe trips for passengers.

When manually- collected data is entered electronically, each row should be a single stop location on a specific route and specific trip or vehicle to allow the data to easily be summarized at a later point at the stop, route, trip, or vehicle level. When using Microsoft Excel for ridership tracking, Pivot Tables are an

[19] Federal Transit Administration. March 2016. "Asset Management Guide for Small Providers." FTA Repot No. 0092. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/Asset%20Management%20Guide%20FTA_Report_No._0092.pdf

easy function to transform raw ridership data into summary data at various levels. They can also be used for data quality control, for example, to make sure that the number of boardings and alightings on a single trip is balanced.

2.2.5. Internal Standards

While there may not be an existing industry standard for every type of data, an agency can create its own data standards for internally collected and managed data. A variety of considerations may be useful when designing data standards. First, consider who will use the data. There may be multiple staff or departments who will use the same information for different purposes. Bringing these staff members together to discuss how they will use the data and what they need to be maintained in the dataset can ensure that the final product is usable for many purposes. Next, data can be formatted in a variety of ways. Decide on the unit of each record—what one row will represent—and assign identification variables to track each row. Decide how this dataset might relate to outside data sources and include relevant ID variables to allow for this. Since datasets are often used for reporting purposes, consider which attributes are necessary for reporting. Finally, determine whether geospatial information (e.g., latitude and longitude) will be needed and how it will be collected or related to a dataset. If the dataset is collected over time, include columns that track the year, month, or day—this can help agencies understand changes in metrics over time. This allows agencies to track trends in ridership, bus speeds, or other metrics that provide valuable information on what types of service modifications could be needed in the future.

Do you need a database?

Once an agency collects data, where to store it is the next question that must be answered. Most agency staff are familiar with the use of spreadsheet tools to store and analyze data. As your agency prepares to increase its use of data practices, a database may be a better home for data in many cases. In many ways, databases and spreadsheets have similar qualities: data can be stored in a table of rows and columns. But while most databases will not include the formatting or charting functions of a spreadsheet tool like Microsoft Excel, they will instead have better tools for validating data inputs, relating datasets to other sources, querying and reporting data, and making data more readily available to others in your agency. While spreadsheet tools like Microsoft Excel or Google Sheets are well known, database software (other than Microsoft Access) is generally less familiar to the average user and may require the assistance of a technical consultant or an IT department to set up. Below are rules of thumb to consider when deciding how to store your data.

Guidelines for Choosing How to Store Your Data

Consideration	Spreadsheet	Database
Number of individuals regularly entering or accessing data	One or two people	More than two people
Length of time data will need to be stored	One to two years	More than two years
Frequency with which data will be created	On an ad-hoc basis	On a regular basis
Volume of records created	Relatively few records (<1,000)	Relatively more records (more than about 1,000)

Do you need a database?

Consideration	Spreadsheet	Database
Consistency of data collected	Characteristics of data to be collected varies	Standard data variables are stored and collected
Technical ability of staff	Less experienced	More experienced
Relationships to other Datasets	Few or no relationships to other datasets	Many relationships to other datasets
Importance of data	Consequences for errors or data loss not severe	Data is critical to the agency

- **The number of individuals regularly entering or accessing data**: The flexibility of spreadsheets can lead to informal data entry practices that may vary across individuals. Databases by default include more strict data validation requirements, such that a particular variable must always be a number or a date. Microsoft Access allows for both data entry forms and standardized reports that also help to facilitate its use by multiple individuals. Other databases can also limit who accesses data, log who has changed a record, and track when data has changed.
- **Length of time data will need to be stored**. The regularity enforced by databases can be an asset to agencies when data needs to be stored and accessed over a long period of time. This can help to minimize differences in data quality as the staff responsible for a dataset change. Suppose you expect only to need to refer to data for a short period of time (such as for a project or analysis of limited duration). In that case, a database may not be necessary.
- **Frequency with which data will be created.** Spreadsheets can quickly be created to support a data collection and reporting effort. If you expect to enter data regularly over time, it may be helpful to use a database to ensure consistency between entries.
- **Volume of records.** While spreadsheet tables can accommodate hundreds of columns and many individual rows, it is often impractical to edit individual records as the spreadsheet grows in size. Databases can more readily store gigabytes of data and often come with reporting tools that allow users to access relevant data quickly.
- **Consistency of data collected.** A spreadsheet's flexibility is an advantage if you are beginning a data collection effort and are not yet sure what variables are important to collect or how they should be stored. Once the variables that will be tracked in a dataset are well understood, a database may be more useful. Some NoSQL or document databases like MongoDB are more flexible than typical relational databases when the types of information tracked about each record may vary.
- **Technical ability of staff.** Even user-friendly databases like Microsoft Access are more challenging to set up and maintain. If staff turnover is high, the cost to set up a database, only to see it abandoned, may be too high.
- **Relationships to other datasets.** Databases are especially well suited to relating tables through key fields. For instance, a table of route ridership by stop might be related to a separate table with information about those stops' locations and characteristics, like whether a bus shelter is present.
- **Importance of data.** If your agency depends on this dataset for reporting to external agencies or partners, the rigor imposed by a database is very useful. If the loss of data would only result in some inconvenience, tracking your data with a spreadsheet may be sufficient.

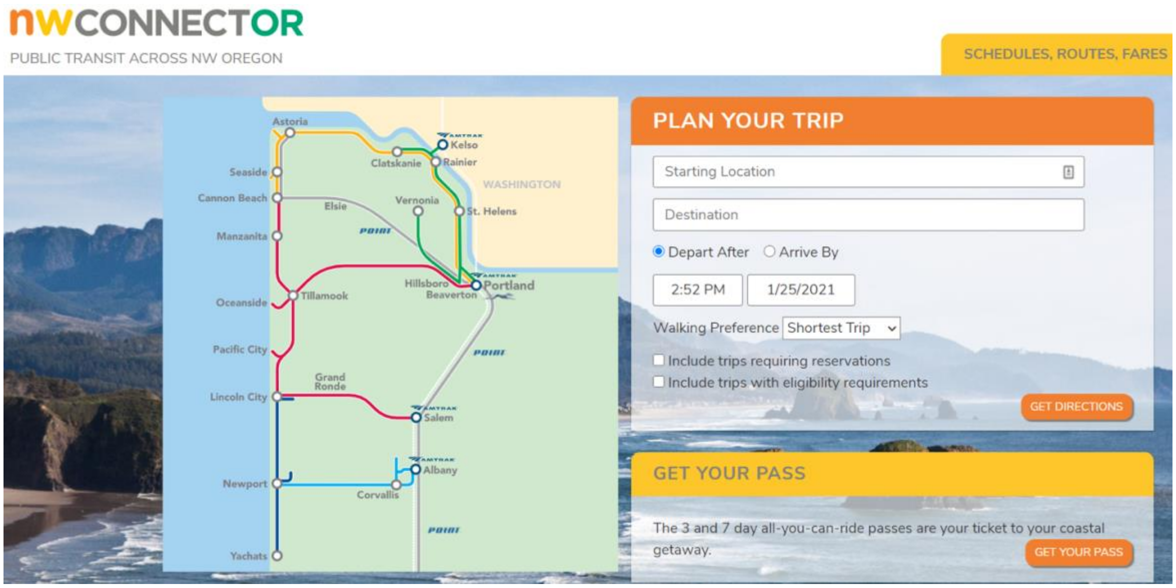
2.3. Emerging Advances

While some data standards have been around for years, the field continues to develop new standards to collect and process new types of data. While GTFS provides a standard for fixed-route services, this format is being extended to provide the same level of access and visibility to flexible transit services. Standards are also emerging to collect and share demand-response-generated data that can allow for greater coordination between demand-response providers. Greater use of GTFS leads to a need to check GTFS feeds' validity to ensure that they are usable and accurate. As intelligent transportation systems (ITS) are used increasingly to automatically collect data about vehicle location, fares, and passengers, data standards are emerging to standardize the collection and processing of this data.

2.3.1. GTFS-Flex

Existing GTFS provides schedule data for fixed-route transit but does not provide information for flexible transit services, like demand-response or deviated fixed routes services. Users for whom fixed-route services are not preferable or an option also deserve to access information about transit services easily; this could include passengers with disabilities who need specialized paratransit services or rural residents who may have access to a demand-response service but no fixed-route services.[20] GTFS-Flex is an extension of the GTFS format that addresses this need by adding booking rules, location groups, and locations and using an updated stop_times table.[21] The extension aims to provide the same visibility and ease of use for flexible transit that GTFS provided for fixed-route services. More resources for using GTFS-Flex, including a white paper by N-CATT, are provided in the resources section at the end of this chapter.[22]

Figure 14: Example Trip Planner Using GTFS-Flex, NWConnector



---

[20] Thomas Craig. 2020. "GTFS-flex." (Presentation, Trillium & Minnesota Public Transit Associatoin). https://www.mpta-transit.org/sites/mpta/files/uploads/documents/events/2020-10/TechnologyGTFS-flex%20for%20Minnesota.pdf

[21] Github. "MobilityData / gtfs-flex." Accessed January 20, 2021. https://github.com/MobilityData/gtfs-flex

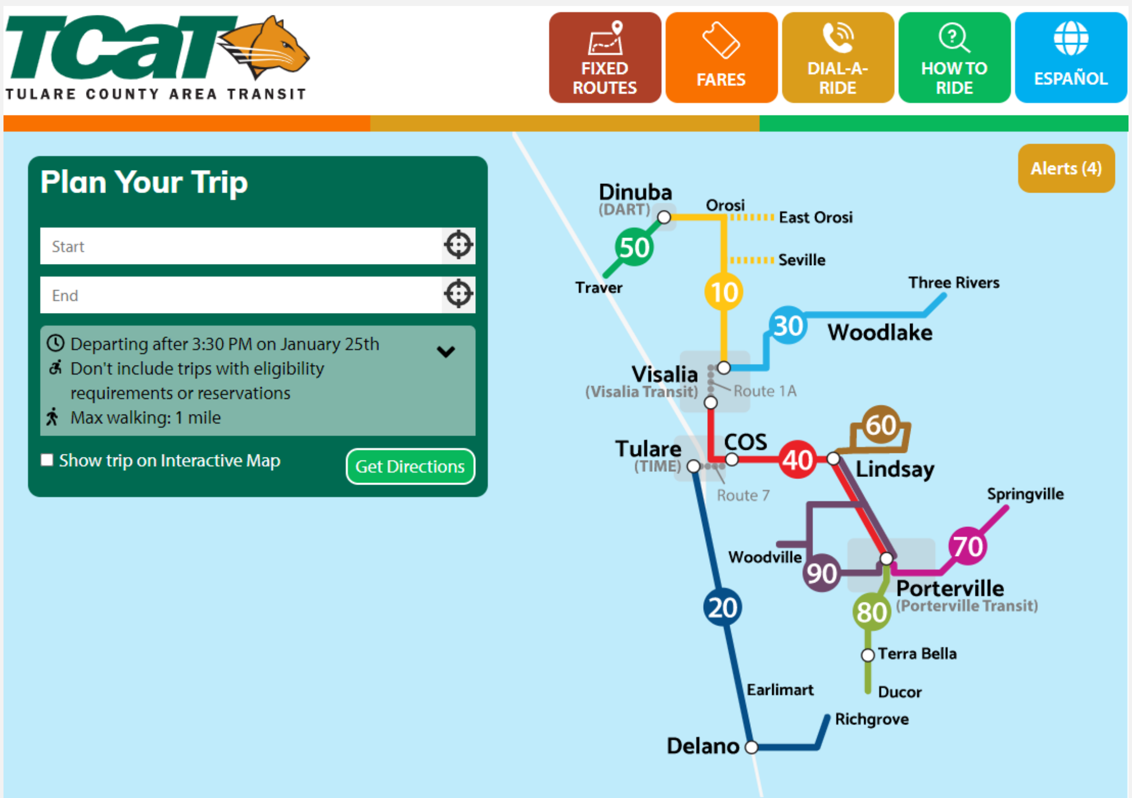
[22] N-CATT. 2020. "GTFS Flex: What is it and How is it Used?" December. https://n-catt.org/tech-university/gtfs-flex-what-is-it-and-how-is-it-used/

Over 100 transit services currently use GTFS-Flex, including: NW Connector, a partnership through five agencies in Oregon; Vamos Mobility, a partnership between 12 agencies in California, and Denver Regional Transportation District in Colorado. GTFS-Flex remains under development and is subject to change based on feedback from users. While Google Maps does not currently support GTFS-Flex, OpenTripPlanner supports several GTFS-Flex features, including: flag stops, which allows boarding and alighting between defined stops; deviated-route services, which allows a vehicle to deviate from its route within a defined radius; and call-and-ride services, which plans trips from origin to destination upon request.[23]

Tulare County GTFS-Flex

Tulare County Area Transit (TCaT), an agency providing transit services in Tulare County, California, offers a trip planner powered by GTFS-Flex. The trip planner enables riders to access information about both fixed-route and flexible, demand-response services. It also provides users with information about how to access demand-response services, including eligibility rules for demand-response trips, requirements for how to schedule a ride, estimated pickup and drop-off times, and fares.

Figure 15: TCaT's Trip Planner Powered by GTFS-Flex

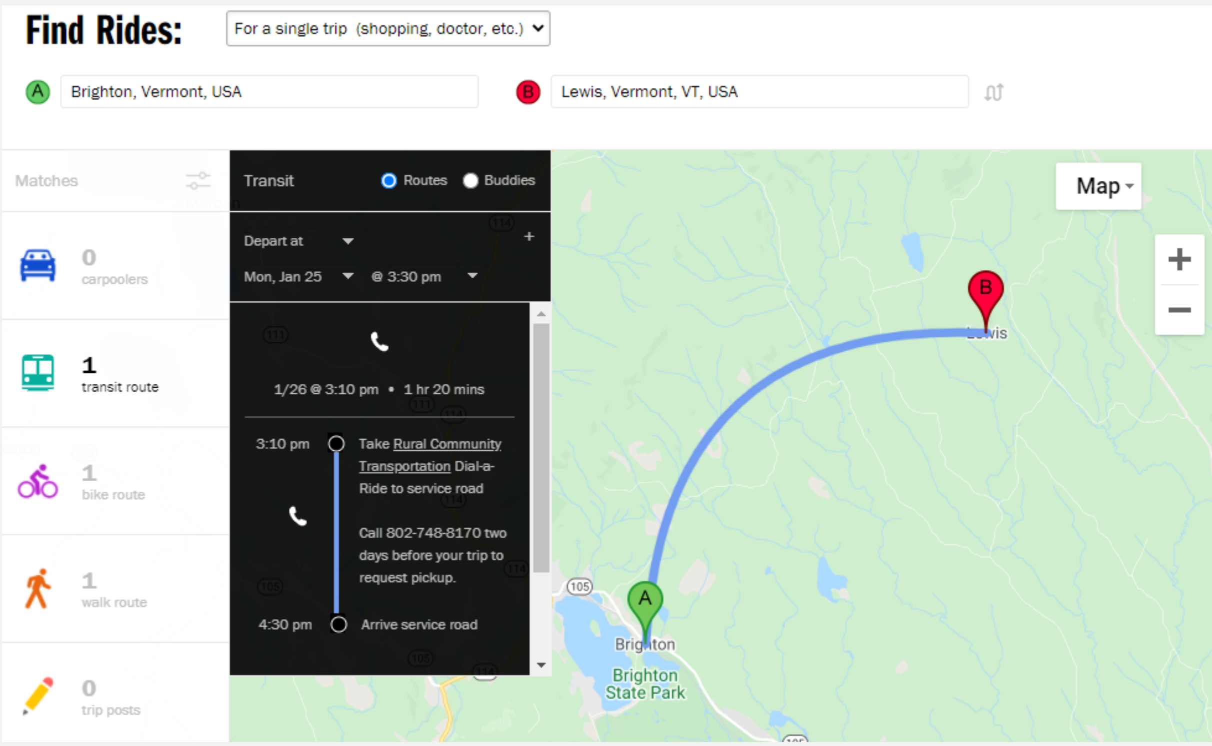


[23] OpenTripPlanner. n.d. "GTFS-Flex Routing." Accessed January 20, 2021. http://docs.opentripplanner.org/en/v1.5.0/Flex/

VTrans Statewide Trip Planner

The Vermont Agency of Transportation (VTrans) provides a statewide trip planner that uses GTFS-Flex to provide information about both fixed-route and flexible transit services. This trip planner was one of the first to use GTFS-Flex to combine multiple modes into a single trip planner. Combining modes allows the service to be used in any part of the state, which allowed nearly 100 percent of Vermont's population to plan transit trips, up from around 50 percent before using GTFS-Flex. To achieve this, VTrans worked with software consultancy Trillium to use GTFS-Flex within OpenTripPlanner, which has since been incorporated into OpenTripPlanner. A second and third iteration of the VTrans trip planner plans to incorporate carpooling, vanpooling, taxis, airport shuttles, and other transportation options. The VTrans trip planner not only helps rural residents find transit services but also helps to connect residents with disabilities to the transit services they can use.

Figure 16: VTrans Statewide Trip Planner



To learn more about Vermont's GTFS Flex in Statewide Trip Planner, visit:

- N-CATT's white paper "Open Source Software and Open Data – What They Are and How to Use Them."
- N-CATT's "Promising Practices Guidebook: Transit Technology Adoption"

2.3.2. GTFS Validators

While GTFS tables are simple to produce, there are few built-in checks to prevent errors from being introduced into the feed. GTFS Validator tools can check feeds for errors that result in the feed not conforming to GTFS specifications. In addition, some tools will also check for inconsistencies in the data, like stops that are far away from the service area or routes with no trips listed. While not formatting errors, these inconsistencies can affect the accuracy of information provided in trip planners. There are several tools that help agencies check or validate their GTFS feeds. Google provides a transit feed validator and provides a list of other resources to validate GTFS feeds.[24] Conveyal produces a GTFS validator that processes feeds faster than Google's validator and conducts a different set of tests.[25] Other feed validators are actively under development, including validators that assess GTFS-RT feeds' quality and other GTFS extensions.

2.3.3. Intelligent Transportation Systems (ITS) Standards

More and more small agencies are adopting intelligent transportation systems (ITS) that automatically track vehicle location (Automatic Vehicle Location, or AVL), count passengers (Automatic Passenger Count or APC), and collect fares (Automatic Fare Collection, or AFC). Of the three technologies, AVL is the most used by small agencies; in one survey, its use increased from six percent of survey respondents in 2010 to 51 percent in 2020.[26]

AVL is produced using sensor technology attached to vehicles. As discussed in Chapter One: Data Sources, combining it with other information like announcements and route patterns can provide useful secondary datasets. AVL can be used in concert with GTFS-RT to provide real-time information to customers. APC data is often collected through infrared screens on doors that can count both boardings and alightings and can often track location and time data. AFC generally accepts a variety of fare media and has the potential to replace APC data. In systems that use smart cards as fare payment, individuals' trips can be tracked over time. Compared with surveys done by staff throughout the year, automatic ITS data is collected at all times of the day throughout the year on equipped vehicles.

There are a variety of ways to process this data after collecting it, which include processing in-house, using software provided by equipment vendors, or contracting with third-party software providers. The Transit Cooperative Research Program (TCRP) is currently conducting research about how to create data standards around collecting and processing ITS data.[27] RouteMatch and Ridecheck Plus are two example products that can collect and parse ITS data. The key to using ITS data effectively is assigning IDs (for stops, trips, trip instances, block, service days, and vehicles) that match other datasets across the agency, including with GTFS. Matching IDs across datasets allows the agency to compare scheduled times to actual times and calculate on-time performance. In addition, reporting data with high granularity– at the route, stop, and trip instance level with trip IDs – allows an agency to examine the data at the route, stop, or trip level. It also allows for quality control checks to take place, such as balancing boardings and alightings on individual trips. Best practices for recording demand response data include recording coordinates for pickups and drop-offs, along with the trip start and end times and vehicle ID.

---

[24] Google. n.d. "Testing GTFS Feeds." *Transit APIs – Static Transit.* Accessed January 21, 2021. https://developers.google.com/transit/gtfs/guides/tools

[25] Github. n.d. "Conveyal / gtfs-lib." Accessed January 21, 2021. https://github.com/conveyal/gtfs-lib/

[26] Del Peterson, Jeremy Mattson, Kenechukwu Ezekwem. 2020. "ITS Technology Usage and Feasibility in Small Urban and Rural Transit." Fargo, ND: *Small Urban and Rural Center on Mobility (SURTCOM 20-01)*.

[27] TRB, "TCRP G-18"

Collecting this data with a high degree of granularity—i.e., at the smallest unit possible—allows agencies more flexibility in extracting a variety of tables or reports from the data. Data at the stop or trip level can help identify popular stops requiring more boarding time or identifying common transfer points. From AVL data, an agency can calculate bus speeds along a route to identify particularly slow or fast travel segments. Combining the three data types can help agencies infer origin/destination pairs as well as popular transfer locations.

Once collected and processed, agencies can use this data for a variety of purposes, including service planning, performance monitoring, scheduling and operations, financial planning, customer information, and reporting. It can even be combined with non-transit data like traffic congestion, demographics, or economic data to understand how service is being used or provided. Although these applications are possible, in most small agencies, ITS data is more likely to be used for day-to-day operations rather than for longer-term planning applications and performance monitoring.[28]

2.3.4. Demand-Response Transit Transactional Data Specifications

Demand-response services are deployed on an individual basis, with each demand-response service either developing their own data communication methods or purchasing software to do it for them. This means that demand-response services cannot communicate with other shared services or collect and disseminate data in a standardized way. TCRP has published data standards that, if adopted, could make collecting, sharing, and coordinating data across demand-response service providers more seamless.[29] They propose a transactional data specification in which a set of rules determine data collection and interactions between software systems. This method establishes a common language for tracking relevant demand response data and provides a recommended approach for how communication about data can occur between agencies or software systems. Implementing these standards could allow agencies to coordinate with one another to provide trips to customers, providing better services for customers, reducing wait times, and improving productivity. Standardized data formats would allow more flexibility in choosing a software provider and would also help produce reports.

Union County Demand Response Spreadsheet

Union County, North Carolina, tracks their demand-response trips in a way that is easy-to-use and analyze for agency staff. Each observation is one individual trip, which comes with: a trip ID; customer ID, which allows tracking of trips by the customer; and vehicle ID, which allows the agency to track vehicle load. Information about the trip is stored, such as whether the trip was canceled, completed, or the customer did not show for the trip. The cost and funding source of each trip is also recorded. Detailed pickup and drop-off times are easy to compare with promised pickup times. The address and latitude and longitude of origins and destinations are tracked, which allows for detailed spatial analysis later. From this spreadsheet, the County can easily generate the number of pickups per hour and see which areas of the County have high demand using only Microsoft Excel and Google Earth.

---

[28] TRB, “TCRP G-18”

[29] National Academies of Sciences, Engineering, and Medicine. 2020. *TCRP 210: Development of Transactional Data Specifications for Demand-Responsive Transportation*. The National Academies Press. http://doi.org/10.17226/25800.

Denver Trip Exchange Project

The Denver Trip Exchange Project is a unique collaboration of multiple agencies providing demand-response services. This project allows the agencies to share data about requested trips, acting as the bridge between different data systems between agencies. A participating agency (such as a social service agency or senior citizen transport agency) fills out a trip ticket for a customer who needs a ride. The exchange automatically identifies agencies who have the capacity to fulfill the trip, and that agency is notified and can choose to claim the ticket and fulfill the ride. Notably, the system is not a centralized reservation system, but trips are requested and claimed on an ad-hoc basis.

2.4. Conclusion

What are your agency's informal data standards for practices without a standard that could be made standard-operating procedures?

Whether formal or informal, data standards can help an agency collect, structure, and process data to be easily usable. Standards are ultimately a key determinant of what an agency, customer, or stakeholder can get out of the data. Formal data standards, such as GTFS, GTFS-RT, or GTFS-Flex, allow schedule data, real-time data, and demand-response data to be broadcast widely via Google Maps and other trip-planning applications. Emerging standards, such as transit asset management, allows an agency to track internal data like vehicle maintenance. In contrast, ITS systems and standards give agencies access to a large range of automatically generated data. Standards allow agencies to use open data sets published by other government agencies and private entities and allow for open data to be more readily published and used by others.

Checklist: What are your agency's internal standards?

Internal Standard	Currently Use This	Want to Prioritize This
ID variables used are consistent with other agency data sources	☐	☐
Staff from multiple departments contribute to the development and maintenance of the data	☐	☐
Dataset has all of the attributes necessary for reporting	☐	☐
Geospatial data is included	☐	☐
Year, month, or day data is included	☐	☐
Data conforms to external standards (as applicable)	☐	☐

Key Resources on Data Standards

- FTA's Transit Asset Management Systems Handbook
- FTA TAM Resources
- Google's GTFS Validators help page
- Google's GTFS-RT Overview
- GTFS Best Practices
- GTFS-Flex on GitHub
- TCRP 210: Development of Transaction Data Specifications for Demand-Responsive Transportation (2020)
- TCRP G-18: Improving Access and Management of ITS Data (pending publication)
- RTAP GTFS Builder

3. OPEN DATA

As both data producers and data users, transit agencies can benefit from accessing external open datasets and by opening their own transit data to the public. Among many benefits, open data practices promote awareness of transit services, increase transparency, improve efficiencies, and spur innovations. Not all open data is created equal: useful open data depends on strong data collection, processing, and documentation to ensure data can be used effectively once published. Challenges associated with providing open data often relate to resources and organizational limitations, data quality and timeliness, and technical issues.

> Open data is data that can be freely used, re-used and redistributed by anyone—subject only, at most, to the requirement to attribute and share-alike.[30]

A variety of open data sources are available to transit agencies to use themselves, often from regional or state agencies or other public data portals. Data from these sources can be used for a broad range of transit planning and geographic information systems (GIS) applications, including socioeconomic and travel demand analyses. While open data sources are easy to access, many of these datasets require collecting and analyzing large volumes of tabular or spatial data, which may lead agencies to seek technical assistance from cities, state DOTs, or private consultants.

Open Data and 'Free-to-Use' Data Sources

Benefits and Challenges

Decisions around open transit data must weigh benefits against the challenges in opening data at small transit agencies. However, findings show that the benefits of opening route and schedule data in standardized formats, and vehicle locations when possible, clearly outweigh any risks and efforts in overcoming challenges.

Open Data Sources

Open data sources are available to transit agencies directly from regional or state agencies or across the web on government websites and public data portals. These include socioeconomic data from the U.S. Census, Census's LEHD/LODES data, and spatial data such as TIGER/Line and OpenStreetMap.

Emerging Advances

Web Tools to Obtain and Visualize External Open Data

A number of websites and free-to-use applications allow users to download, view, or visualize open data created by the Census Bureau and other sources, including Census Reporter, Policy Map's Open Map, On the Map, CTPP Data Viewer, Remix, Google's MyMap, and RTAP applications.

New Data Sources

New data sources, including location data from smartphones, data from trip planning and fare payment apps or mobility-as-a-service (MaaS) apps, and data from new mobility options, such as shared scooters and bicycles and transportation network companies (TNCs), can provide additional insight into how people move around their communities.

---

[30] Open Data Handbook. n.d. "What is Open Data?" https://opendatahandbook.org/guide/en/what-is-open-data/

3.1. Introduction

As both data producers and data users, transit agencies can benefit from accessing external open datasets and by opening their own transit data to the public. Accessing external datasets can help agencies serve their community and operate efficiently, while sharing internal datasets can bring benefits ranging from improved customer information to improved performance. In this context, this chapter aims to help agencies make decisions about opening transit agency data and accessing external datasets.

N-CATT's white paper "Open Source Software and Open Data – What They Are and How to Use Them," published in October 2020, explores the characteristics, benefits, and challenges of open transit data, as well as what open-source software and open data each can and cannot deliver for transit systems in rural, tribal, and small urban areas. As a complement to the white paper, this chapter dives deeper into how agencies can also take advantage of other open data sources.

From Census tools and datasets to local government data portals, agencies already have a bounty of open data at their fingertips. Open data refers to openly accessible repositories of data collected by government agencies or made available by private entities with a permissive license. In addition to being freely accessible, data is open if it can be used, modified, and shared by anyone for any purpose—subject only, at most, to requirements to provide attribution and/or share-alike.[31] Several 'free-to-use' data sources may have restrictions on how the data is used or how much can be obtained and so are not open data but may still be available to use. Transit agencies may also access data through sharing agreements and partnerships, further discussed in Chapter Five: Partnerships.

Either manually-collected or passively-generated, and whether following a standard or not, the data generated by transit agencies is vast. An agency recipient or subrecipient of Federal Transit Administration (FTA) Chapter 53 funds collects and reports annual data on the types of services provided, operating data, and asset conditions, among other data elements. Demand-response providers keep a database of trips. By publishing a General Transit Feed Specification (GTFS) feed, an agency shares standardized data on the location of bus stops, schedules, and more. An agency with a fleet equipped with automatic passenger counting (APC), automatic vehicle location (AVL), or automatic fare collection (AFC) technologies generates data useful for operational activities. Though all these data examples fall under the umbrella of transit data, the most impactful open transit data is shared in a standardized, machine-readable format.

There are costs and a level of effort involved in opening and maintaining updated datasets. Effective open transit data depends on good internal data management and may require changes to data collection, data processing, and data documentation to ensure data accuracy. Many small transit agencies operate with limited personnel and lack dedicated staff or divisions responsible for data collection and management. Despite the challenges, those processes may yield benefits from simplifying agency workflows to producing better performance statistics.

3.2. State of the Practice

A range of potential benefits motivate agencies to open their transit data and consume open data, though agencies must consider the risks, challenges, and costs (including staff hours) of doing so. Issues around

[31] Open Data Handbook. n.d. "Glossary: Open Data." http://opendatahandbook.org/glossary/en/terms/open-data/

open data legislation and data privacy also underly these decisions. This section covers some observed benefits of open transit data and discusses aspects of opening transit data and consuming open data.

3.2.1. Benefits of Open Transit Data

There are many beneficiaries of open transit data practices, from individual transit agencies to customers to third-party developers to the transit industry as a whole. Open data practices lead agencies to promote service awareness and transparency, improve efficiencies, and foster improved relationships with data stakeholders. By delivering data through customer-facing applications, transit agencies can improve customer satisfaction, become more engaged with customers, and improve service and data quality.[32] Lastly, third-party developers, who typically rely on transit agency-provided data, can continue to spur innovations and establish themselves as links between transit services and customers. For these reasons, the transit industry has continued to push for the development of data standards, as discussed in Chapter Two: Data Standards.

Agency Benefits

Increased awareness of the available transit services is one of the most common benefits agencies experience as a result of providing open data.[33] Given trends in government transparency enabled by web technologies, the general public increasingly expects transit providers to publish data in free-to-use and open formats.[34] Beyond improving public perception and providing a sense of transparency, agencies that provide standardized, open transit data increase awareness of their transit services. Open transit data can serve to publicize available transit services to existing and potential new customers. The higher quality transit information results in improved customer service and experience and potentially increased ridership.

Central to many open data initiatives is the commitment to customers and their needs. To address these, agencies embraced a continuous process of improvement of the data shared and services offered. By releasing more customer-oriented transit data, many agencies face lower costs associated with customer requests, among other efficiencies. More often than not, open data initiatives also streamline data collection and management practices at transit agencies. These realizations illustrate the efficiencies and potential savings associated with open data.

Open transit data encourages innovation outside agency walls that can ultimately benefit the agency in turn. Private developers have responded to open streams of public transit route, schedule, and vehicle data by developing travel apps that provide trip planning and vehicle arrival information to customers. This allows agencies to extend their reach to existing and potential customers through applications without having to invest in software application development. Additionally, by opening up data, transit agencies can identify problems with datasets more easily through feedback from external users of the data.

---

[32] Federal Transit Administration. 2016. *FTA Open Data Policy Guidelines*. Washington, DC: FTA Report No. 0095. https://www.transit.dot.gov/sites/fta.dot.gov/files/FTA_Report_No._0095.pdf

[33] National Academies of Sciences, Engineering, and Medicine. 2015. *TCRP Synthesis 115 Open Data: Challenges and Opportunities for Transit Agencies.* Washington, DC: The National Academies Press. http://doi.org/10.17226/22195

[34] National Academies of Sciences, Engineering, and Medicine. 2020. *TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future*. Washington, DC: The National Academies Press. https://doi.org/10.17226.25696

Customer Benefits

One of the most significant benefits that transit agencies consider when publishing open transit data is its potential to impact customers positively. Almost 90 percent of transit agencies in a recent TCRP survey release route, schedule, and vehicle location data in order to provide customers with more information.[35] These open data feeds provide customers with greater certainty about their journeys and potentially save passengers time resulting in improved customer satisfaction and even inducing higher ridership.

Almost 90 percent of transit agencies in a recent TCRP survey release route, schedule, and vehicle location data in order to provide customers with more information.

Open data policies and practices can also create a conduit for agency and customer interaction, improving customer relations. For example, through open data activities, agencies can engage customers with satisfaction surveys, and by tracking application downloads and the number of times real-time data feeds are accessed. Application and customer satisfaction evaluations, application tracking, and continuous improvement processes associated with open data practices orient the agency toward engaging with its customers and recognizing their experiences.[36]

Third-Party Developers

In addition to transit agencies and transit customers, third-party developers are crucial players who use transit agencies' data to create innovative products and tools for the public. Private companies offering trip-planning apps increasingly serve as de facto front ends for transit agencies.[37] Serving customers in ways agencies may not be capable of doing, these companies commonly invest significant resources into improving transit data to conform to quality standards and add new features to the dataset to meet the rider's expectations. Beyond mobile applications, developers also use multiple operational and service data to create innovative practices in service planning, operations, and maintenance, among others.[38]

3.2.2. Publishing Open Transit Data

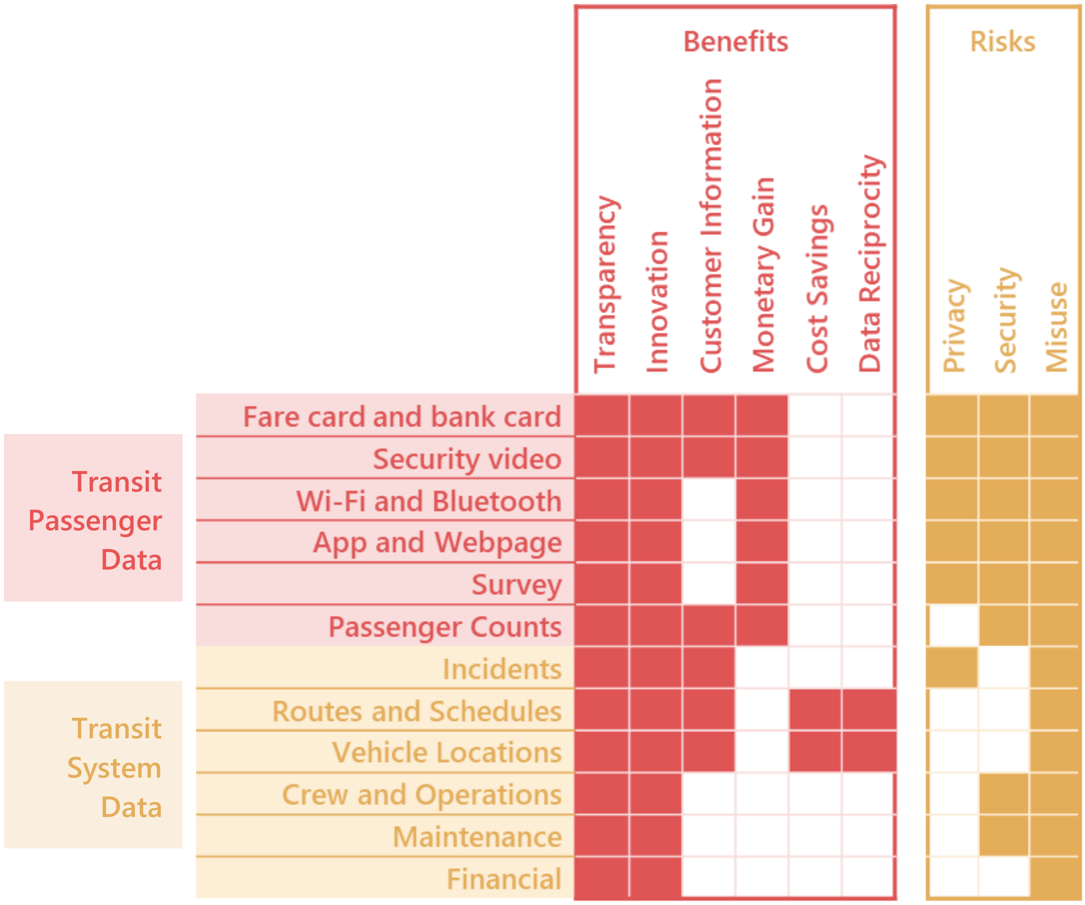
Challenges

Decisions around open transit data must weigh benefits against the risks in opening data while considering the challenges in doing so. **Figure 17** indicates the potential benefits and risks of sharing several transit passenger and system datasets.[39] Passenger data include passenger count data and survey data as well as newer data types, such as fare or bank card transactions, video, Wi-Fi, Bluetooth, and app and webpage usage data. Most types of passenger data contain records of individual passengers or records that have the potential to identify an individual, posing a privacy risk. Data about the transit system includes route and schedule data, vehicle location data, maintenance, staff and operations data, and financial data. Risks related to these datasets typically do not contain privacy risks, and for routes and schedules and vehicle locations, the benefits are generally considered to always outweigh the risks.

---

[35] National Academies of Sciences, Engineering, and Medicine. "TCRP Synthesis 115."
[36] FTA. "Open Data Policy Guidelines."
[37] Transit Center. 2018. "The Data Transit Riders Want: A Shared Agenda for Public Agencies and Transit Application Developers." https://transitcenter.org/publication/transit_data/
[38] FTA. "Open Data Policy Guidelines."
[39] Adapted from TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.

Figure 17: Transit Data Types and Sharing Characteristics



Source: Adapted from TCRP Report 213: Data Sharing Guidance for Public Transit Agencies—Now and in the Future (2020).

Effective open data depends on good internal data management, which may require changes to data collection, data processing, and data documentation to ensure data is used effectively and appropriately when published. Considering the diversity of transit datasets, publishing some types of transit data can be resource-intensive and both technically and organizationally challenging for some transit agencies. Challenges associated with providing open data often relate to resources and organizational limitations, data quality and timeliness, and technical issues.[40]

There are many steps required to prepare data for being published. These steps require staff time and may also require contracting with external vendors, in some cases. Often data is collected without analysis or sharing in mind, stored across different locations and systems, and not documented for external use. These data silos present challenges for both the internal and public use of the data. Good internal data management practices make data sharing easier. A well-documented, centralized internal data repository, for example, helps transit agency staff make use of data and reduces steps required to distribute data but may require dedicated staff time to develop and maintain it.[41]

[40] TCRP Synthesis 115 Open Data: Challenges and Opportunities for Transit Agencies.
[41] TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.

Personnel and organizational demands, particularly in small transit agencies, also have consequences for open data sharing. Developing technical capabilities among transit agency staff requires staff training, and the additional challenge of staff turnover can make it difficult to ensure that progress in data management is sustainable. Establishing a staff member or team that is dedicated to data management is an important step in addressing challenges in opening transit data.

Technical limitations related to data collection influence the types of data, the data quality, and coverage available to be shared. While large transit agencies generally have full coverage of AVL systems in their vehicle fleets, GPS polling rates and accuracy vary from one implementation to another. Additionally, vendors that install and maintain systems, such as AFC, APC, or AVL, may retain ownership of the data generated, offer limited data reports, or export data in standards particular to the vendor. Although there has been a shift in ownership of AVL and AFC systems to transit agencies and standardization in real-time service/vehicle location, this issue may still persist for other data types. For example, if maintenance is outsourced, some maintenance data may be owned by the maintenance company rather than the transit agency.[42]

How to Successfully Open Your Data

Route and schedule information, system alerts, and the real-time location of transit vehicles are often the first pieces of data that come to mind when discussing open transit data. In most cases, route and schedule information is published in the standard GTFS and GTFS-related formats, such as GTFS-Flex and GFTS-RT. However, agencies commonly publish information on performance indicators, including route or line-level ridership, passenger counts at bus stops and train stations, on-time performance, and reliability indicators. Transit agencies also provide summaries of survey data, including travel surveys and customer satisfaction surveys.

An Application Programming Interface (API) is a way computer programs can talk to one another. APIs used to publish data allow programmers to choose specific records they need and are authorized to view. The API is often linked to a database that is updated periodically.[43]

Transit data can be made public in many forms. Agencies have maps, data, and reports that can be downloaded from their websites. Interactive dashboards allow customers to interact with the data in a controlled way. GTFS feeds can be available online for download. Among other kinds of data, route, schedule, and vehicle location data can be shared using an Application Programming Interface (API). These formats have advantages and disadvantages, as shown in **Figure 18**.[44]

[42] TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.
[43] Open Data Handbook. n.d. How to Open up Data. http://opendatahandbook.org/guide/my/how-to-open-up-data/
[44] TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.

Figure 18: Methods for Sharing Data Online

Static Reports	Interactive Dashboards	Data Repositories	Developer APIs
✓ Accessible to all audiences ✓ Protect against misinterpretation of data	✓ All audiences can interact with the data	✓ Researchers and innovators can download and manipulate, generating new analysis and insights	✓ Developers can efficiently pull data into apps
× Data cannot be manipulated	× If underlying data is not downloadable, data manipulation is limited × Dashboards require significant effort to develop	× Can be difficult to use for non-technical audiences × Risk of data misuse × Not appropriate for disaggregated data on individuals	× Only appropriate for data sources that many developers want to access at frequent intervals × Difficult for non-technical users

Increasing interactivity

Source: Adapted from TCRP Report 213: Data Sharing Guidance for Public Transit Agencies—Now and in the Future (2020).

High-quality transit data begins with the public agencies that produce and publish the data. Agency leadership is responsible for ensuring that potential customers can access the information they need to make transit their preferred choice.[45] FTA's Open Data Policy Guidelines Report lists several guiding principles and techniques for transit agencies to open transit data based on agencies' insights that have successfully moved data into the open arena. These are:

- **Follow best practices:** Lessons learned from agencies regarding practices, procedures, and experiences are brought to light through workshops, webinars, interviews, and research.
- **Embrace an altered perspective:** View information and data systems as tools to improve decision-making and expand the influence over customer experience instead of overhead expenses.
- **Become data-centered:** Many agencies that embrace open data practices do so because it is part of an ongoing business enterprise that requires updating and curation.
- **Focus on the customer:** This customer focus is at the core of the mission of open data within the transit industry.
- **Implement successful strategies:** For agencies that have the support of staff and leadership, the implementation strategies guide them towards more improved open data programs.
- **Develop an open data policy:** A formal open data policy can guide the decision-making process, including procurement and technology considerations.
- **Adopt a phased approach:** When leaders are skeptical, and a formal open data policy is not politically palatable, agency staff may find success through pilot projects.
- **Overcome technology barriers:** Developing the skills within the organization to manage, create, and distribute the data, to outsource the responsibility for data management, or create a blended approach as is typical for most agencies.

[45] Transit Center. The Data Transit Riders Want: A Shared Agenda for Public Agencies and Transit Application Developers.

- **Engage with data users:** Agencies that successfully share their data often have a relationship with the main data stakeholders.
- **Overcome fear of hazards:** Several factors contribute to agencies maintaining a closed data position, including loss of control over data and data quality and fear of legal hazards of open data.

A step-by-step process for publishing open transit data is provided below.

I want to open transit data[46]

Step 1. Engage the community

- Engage the community to determine why and how to open your transit data.
- Find out what the people want to know.
  - When is the next bus?
  - How do I get from here to there using transit?
- Get to know the development community, and determine example uses of the data or API.

Step 2. Find the data

- Inventory what is available at your agency.
- Look into potential partnerships.
- See if partners already have an API that you can mirror or integrate.

Step 3. Decide how to format your data

- If you use a standard format, you benefit from all the work that others have done integrating and tooling that standard. This can lead to cost savings in developing documentation, reduced support issues, and increased adoption. In general, if there is a standard, the default should be to develop to it.

Step 4. Decide how to deliver your data

- If you will be offering any real-time data that would require an API instead of any kind of bulk file-based release.
- Make the formatted data available for download. If you are not releasing real-time data at this time, this will be easier than developing a full API for the GTFS data.

Step 5. Release it

- Release the data or API. Make sure that it is properly documented so that people can use it. Talk to the community to make sure that they understand what has been released and that they know that it is actually out.

[46] Adapted from Let's Open Some Government Data! n.d. "Yeah, I want to open transit data." http://jlord.us/open-data-guide/transit/

I want to open transit data[46]

Next Steps

- Reengage the community.
- Consume your own data or API as a means of confirming its quality and utility.
- If facing challenges, it may be appropriate to reevaluate the costs and benefits of sharing. Make sure that:
  - It is *possible* to get the data.
  - It is *feasible* to get the data.
  - It is *cost-effective* to get the data.

3.2.3. Using Open Data

Open data sources are available to transit agencies directly from regional or state agencies or from other public data portals. Data from these sources can be used for a broad range of transit planning and geographic information systems (GIS) applications, including socioeconomic and travel demand analyses. While open data sources are easy to access, many of these datasets require collecting and analyzing large volumes of tabular or spatial data, which may lead agencies to seek technical assistance from cities, Metropolitan Planning Organizations (MPOs), state DOTs, or private consultants.

Socioeconomic Data

Socioeconomic data is publicly available from the U.S. Census, including demographic and employment information for all states and territories at high levels of detail. The most complete datasets come from the decennial Census conducted every ten years, but the Census Bureau offers a large suite of free-to-use data products that are updated more often and cater to specific data topics. Central to many Census datasets is the American Community Survey (ACS), which collects an annual sample of population information, including counts by age, race and ethnicity, income, health, education, employment, and other household characteristics. Employment data is compiled into the Longitudinal Employer-Household Dynamics (LEHD) dataset, which provides details on industries, earnings, job growth, education, commute information, and other employment characteristics. Each of these Census products offers data as tables, with geographic identifiers associated with each piece of data. Census data can be downloaded for geographies as large as states and regions or as small as city blocks. The uses of socioeconomic data for access and equity analyses are discussed in Chapter Six: Access and Equity.

How to use shared or open data to design a service to attract more riders[47]

Many rural providers designed their transit systems to serve one or several specific groups of riders on a demand-response basis. Designing a service that aims to address multiple markets can increase potential ridership and feasibility of more efficient services, including fixed-route services.

---

[47] The National Academies of Sciences, Engineering, and Medicine. 2017. *NCHRP Research Report 861 Best Practices in Rural Regional Mobility*. Washington, DC: The National Academies Press. http://www.trb.org/NCHRP/Blurbs/176823.aspx

How to use shared or open data to design a service to attract more riders[47]

Markets for services include:

- Employment
- Medical appointments
- Community college or high school
- Intercity, regional, or commuter connections
- Shopping and personal business
- Social and recreational needs.

Transit service design usually begins with identifying needs and potential markets. The planning process often includes estimating potential ridership by market segment, which can be estimated by looking at data sources appropriate for each market identified and summing the likely ridership. Open data and data obtained by data-sharing agreements can be used to assess the transit needs of specific target markets include the following:

- **Employment trips:** The Census LEHD data source provides annual employment statistics linking home and work locations to the Census Block level, which can be accessed through the Census Bureau's OnTheMap interface.
- **Medical Trips:** Information on specific groups may be obtained from agencies that fund or contract for these services, such as appropriate agencies that fund Medicaid non-emergency medical transportation (NEMT) or the Veterans Transportation Service mobility manager at the regional Veteran's Administration medical centers.
- **Educational Trips:** Community colleges, colleges, and universities can provide data on commuting students' residence, providing numbers of students by zip code of residence. Potential ridership from an educational institution includes students, staff, and faculty.
- **Connecting Trips:** Publicly available GTFS data on routes, stops, and schedules can point to possible connecting trips. Additionally, publicly available GTFS, GTFS Flex, and GTFS-RT data allow applications to provide information on trips involving more than one carrier. Providing better schedule information (especially if combined with interline ticketing) about making connections is likely to benefit users and providers alike.

Improved ridership estimations may allow for rural transit services designed to meet multiple user groups' needs, making the most effective use of limited resources and attracting ridership from multiple markets. A rural transit provider offering several different types of transportation service, such as fixed-route general public transit, demand-response services, and human service client transportation, will have lower unit costs than if each service were provided by a separate organization. Typically, administrative costs can be spread over more services, and staff, vehicles, and facilities can be used to provide several services, which reduces resource requirements.

Transportation Data

Transit agencies can also take advantage of open transportation data, typically made available by governmental or other regional transportation agencies. This type of data includes travel volumes and commute patterns between origins and destinations and is often inferred from Census data or estimated using a travel demand model. The Census Bureau provides the Census Transportation Planning

Products (CTPP) dataset based on survey questions about Americans' places of work. CTPP data can be analyzed to evaluate the origin, destination, duration, and mode of commute trips, also known as journey-to-work trips. LEHD Origin-Destination Employment Statistics (LODES) offer an alternative source of commute data based on the LEHD dataset, showing both home and work locations paired together. Beyond the journey to work, many state and metropolitan area planning bodies maintain a regional travel demand model of all trips. These travel demand matrices provide a vast dataset of every origin-destination pair within the region and can be used to analyze travel patterns by trip purpose, time of day, and mode. Agencies can access travel demand data online or by direct request from the planning organization.

Transit data is publicly available through the National Transit Database (NTD) and other transit agency websites. NTD's standardized data across agencies make it easy to use for cross-agency studies. Transit agencies can access data from NTD to compare ridership, service hours and miles, or other metrics to similar agencies across the country or to understand trends in public transit performance. For a more detailed analysis of another agency's scheduled services, GTFS data can also be accessed from agencies that publish their GTFS feed online.

Spatial Data

Many socio-economic and transportation datasets have spatial attributes that allow for mapping and geospatial analysis with GIS software. These spatial attributes are often provided as point coordinates or a numeric identifier for the state, county, or zone associated with the data. Spatial datasets in various spatial data types can be found from numerous sources online, including both government-produced and crowdsourced datasets. Some spatial datasets include:

- **Topologically Integrated Geographic Encoding and Referencing (TIGER) or TIGER/Line** are geodatabases used by the U.S. Census to describe land attributes such as roads, buildings, rivers, and lakes, as well as Census tracts. TIGER shapefiles do not include the Census demographic data but contain geographic entity codes (GEOIDs) that can be used to merge Census demographics.[48]
- **OpenStreetMap (OSM)** is a collaborative project to create a free-to-use editable map of the world, with data quality comparable to proprietary data sources in many places. Extracts of OSM data can be used for geospatial analyses and route planning, among other uses.
- **Points of Interest Data**, such as activity centers identified through searches on Google Maps, can be downloaded using browser-based scraping tools without having to write codes or download special software. They are also available through organizations like Policy Map (see Emerging Advances section).
- **Land use and other spatial datasets from public agencies** are downloadable from government data portals, including spatial datasets of transit routes from other nearby agencies.

Analysis Skills

Desirable data and GIS analysis skills to take full advantage of open data sources may not be available for many rural and small transit providers that lack the resources (both in terms of time and staff) necessary to conduct advanced data analysis related to transit planning. With limited funding, agencies also struggle to hire data-savvy staff who can appropriately leverage the data that agencies can access and generate. A rural transit program that is part of a local government entity or regional planning organization is likely to have at least periodic access to GIS mapping and data services through its planning department. On the other hand, private non-profit entities are less likely to have this capability

[48] U.S. Census Bureau. 2020. TIGER/Line Shapefiles. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

available in-house.[49] GIS analysis, for example, may only be feasible as part of a larger planning project conducted by a consulting firm, a regional planning organization conducting a Transit Development Plan, or a local public transit-human service transportation coordination study. Open-source programs like Google Earth allow users to create basic GIS route and stop alignments, and can be used in applications like the RTAP GTFS Builder (see case study in Chapter Two: Data Standards). Other web-based tools detailed in the Emerging Advances section allow for some spatial analysis of things like ACS and LEHD data, or points of interest.

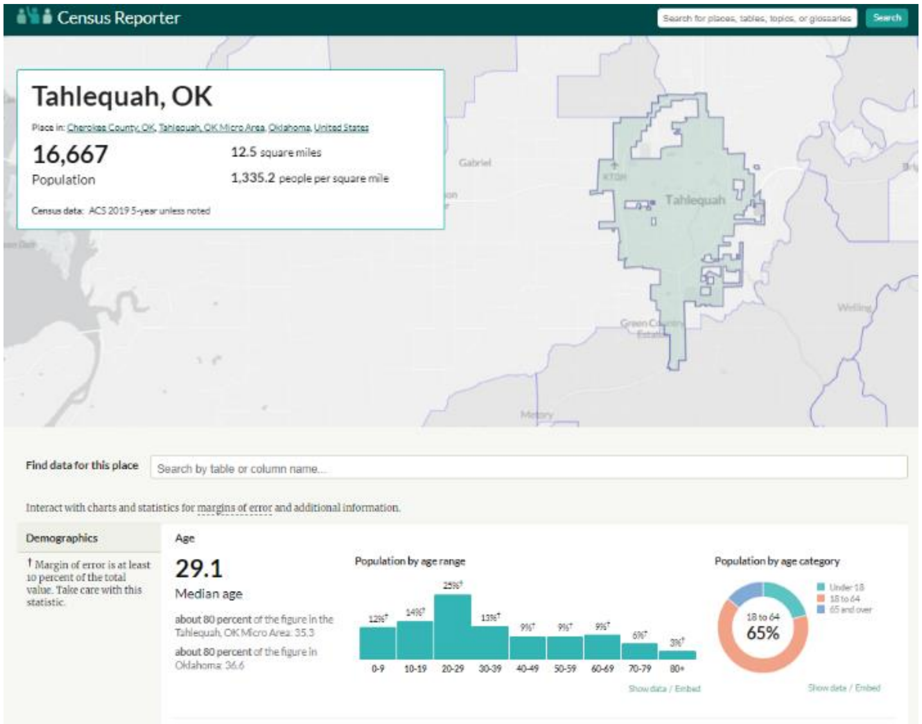
3.3. Emerging Advances

3.3.1. Web Tools to Obtain and Visualize External Open Data

A number of websites and applications allow users to download, view, or visualize open data created by the Census Bureau and other sources. These applications are free-to-use and generally accompany tutorials or user guides to help first-time users navigate them. Applications built for Census data are the most common and include the following.

- **Census Reporter** allows users to view and download ACS data for all available geographies. Users can start by entering the name of the geography or topic they wish to explore or by clicking on pre-populated lists of topics. Data can be visualized in interactive maps in the interface and also downloaded to be used in spreadsheets.

Figure 19: Census Reporter Interface Showing Tahlequah, OK

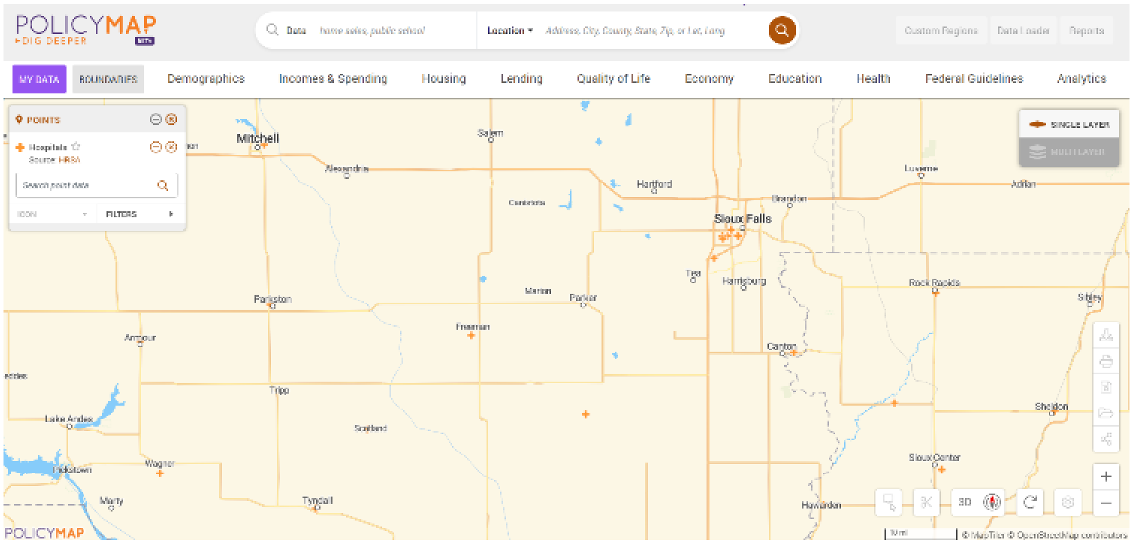


- **Policy Map** has an Open Map interface that allows users to explore data on demographics, housing, quality of life, and healthcare (among other statistics) in an interactive map interface. While much of the available data is based on the ACS, this application can also provide agencies with the locations

[49] The National Academies of Sciences, Engineering, and Medicine. 2012. Data Needs for Assessing Rural Transit Needs, Benefits, and Levels of Service. http://www.trb.org/Publications/Blurbs/167758.aspx

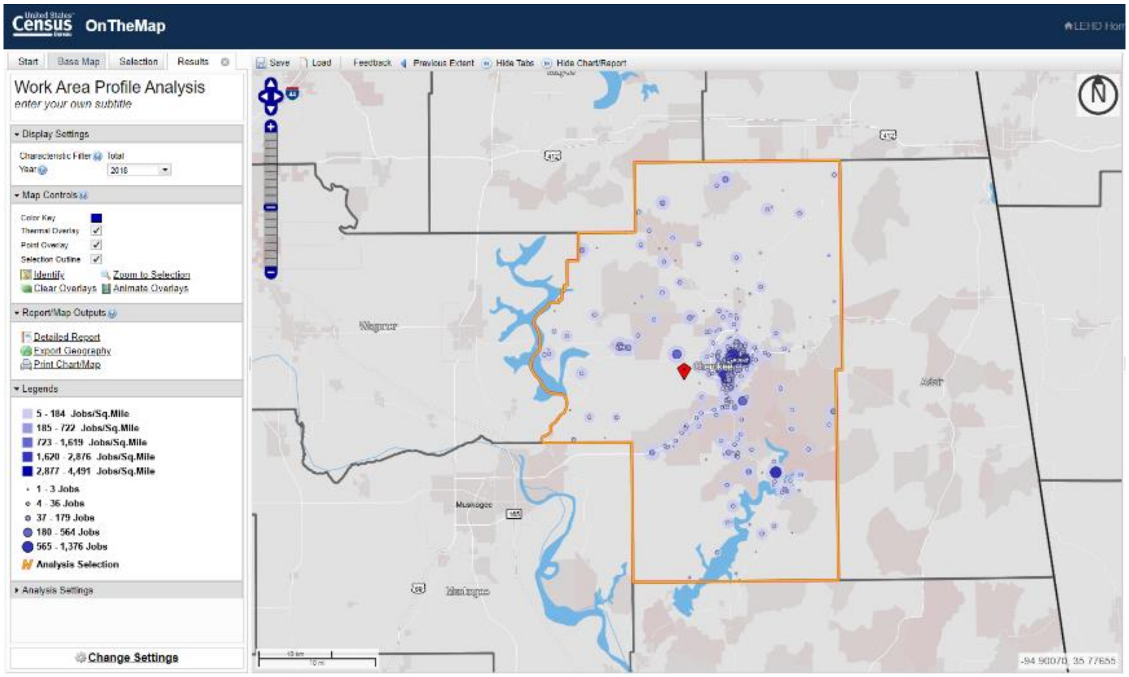
of important destinations such as hospitals and grocery stores and zones like [Rural Food Access Investment Areas] so agencies can plan services accordingly. The information available has been compiled from numerous research and policy organizations around the country.

Figure 20: Policy Map Interface Showing Hospital Locations



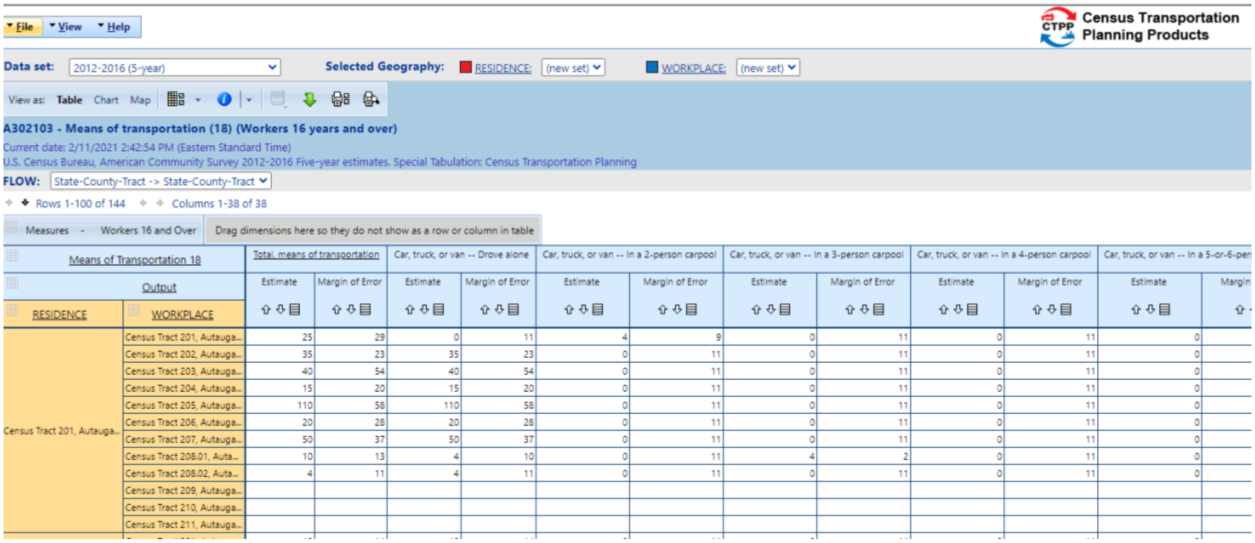
- **[OnTheMap]** allows users to explore LEHD/LODES data. Employment data can be viewed for specific geographies based on where jobs are located or where workers in that geography live and as home to workflows between geographies. Users can view data in the interface or download data in spreadsheet format or spatial format, including for use in Google Earth. Data can also be viewed in the interface in various types of charts that can be downloaded.

Figure 21: OnTheMap Interface Showing Employment in Cherokee County, Oklahoma



- **[CTPP data](#)** can be explored in a user interface maintained by the American Association of State Highway and Transportation Officials (AASHTO). Users can select residence and workplace locations, and view home-to-work travel flows between them, sliced by various ACS demographic characteristics such as mode used and travel time to work. The data can be displayed in interactive tables or maps and downloaded for additional analysis.

Figure 22: CTPP Data Viewer



In addition to Census-based applications that help agencies better understand their service areas and customer bases, other web applications can help agencies with service planning and other tasks. [Remix](#), for example, is a web application that lets transit agencies easily plan routes. [MyMaps](#) is a Google web application for map creation that allows users to upload tabular data (in CSV, XLSX, and GPX formats) and spatial files (in KML format) for easy geocoding, distance measuring, and map creation. The Rural Transit Assistance Program (RTAP) offers a series of [free web-based applications](#) that include a cost allocation calculator, a GTFS builder, FTA procurement procedures guidance, and a website builder.

The Center for Urban Transportation Research (CUTR) [repository at GitHub](#) (a code hosting platform for version control and collaboration) includes a list of transit APIs, applications, datasets, and resources. In addition to other web-based applications, the page lists other open-source or free software tools for transit agencies that include:

- [TransAM](#), an open-source asset management platform for public transportation agencies
- [RidePilot](#), open-source computer-aided scheduling and dispatch (CASD) software system designed for small human service transportation agencies
- [Transit Boardings Estimation and Simulation Tool (TBEST)](#), GIS-based modeling, planning, and analysis tool which integrates socioeconomic, land use, and transit network data into a platform for scenario-based transit ridership estimation and analysis.

3.3.2. New Data Sources

Transit is part of a multimodal transportation system. New data sources, including location data from smartphones, data from trip planning and fare payment apps or mobility-as-a-service (MaaS) apps, and data from new mobility options, such as shared scooters and bicycles and transportation network companies (TNCs), can provide additional insight into how people move around their communities. These datasets can serve many uses, including: helping evaluate overall demand patterns; providing more detailed road speed data to inform bus operations, route alignment decisions, improve bus arrival predictions; and identifying first and last-mile transportation modes and distances.[50]

Data sharing is often critical to building a well-functioning multi-modal transportation network. Partnerships for service provision across modes and with private mobility providers can be facilitated and often require some degree of data sharing. Some public transit agencies have also been working with cities and states to develop and push for regulations to facilitate access to external data streams. In this context, many cities have begun regulating and managing private mobility companies that operate on their public right of way, notably micromobility options. However, a small micromobility or TNC user base can make those datasets less valuable for drawing many conclusions.

Although micromobility or TNC usage may be limited in smaller metropolitan areas and rural areas, cellphone, location-based services, and GPS trace data can be useful for some smaller transit agencies. Cellphone connection data is collected by cellular service companies, while smartphone apps that use users' locations collect location-based services data. Some companies even infer a user's travel mode, helping agencies understand characteristics of alternate modes and demand patterns on alternate modes. These data sources are usually aggregated by analytics companies who derive and sell speed and origin-destination insights. Although generally not open or free-to-use, these data sources might be available to small transit providers through state DOTs or partnerships designed to be replicated in small metropolitan areas.[51]

3.4. Conclusion

Advances in data sources, data standards, and open data and tools can help transit agencies generate more value from their data and external datasets. Transit agencies need to set goals that can be accomplished through open transit data and data analysis and develop staff capabilities and data management processes to work toward those goals. Promoting a culture of open data and data analysis should be a long-term commitment for transit agencies that involves:

- The funding and prioritization of new technologies, data sources, and data analysis at the leadership level
- Training and acquiring qualified employees who have the skills to analyze data at the staff level.

[50] TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.
[51] The National Academies of Sciences, Engineering, and Medicine. 2017. Dubuque Smarter Travel Project. https://trid.trb.org/view/1437892

Checklist: What are your agency's open data needs?

Staffing and Data Needs	Is a Current Need	Want to Prioritize This
Dedicated staff person or division focused on data management	☐	☐
Staff with data administration and maintenance skills, including an understanding of security and permissions	☐	☐
Staff with data analytics, including the ability to use scripts to automate data analysis processes and work with large datasets	☐	☐
Agency has a data catalog	☐	☐
Agency has data-sharing protocols in place	☐	☐
Agency has an open data preparation needs checklist	☐	☐

Key Resources on Open Data

- N-CATT's white paper "Open Source Software and Open Data – What They Are and How to Use Them."
- FTA's Open Data Policy Guidelines
- Open Transit Data Toolkit
- Transit Center's "The Data Transit Riders Want, A Shared Agenda for Public Agencies and Transit Application Developers"
- TCRP Synthesis 115 Open Data: Challenges and Opportunities for Transit Agencies
- TCRP Report 213 Data Sharing Guidance for Public Transit Agencies—Now and in the Future

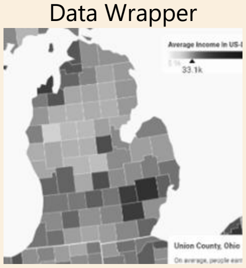
WHAT EXACTLY DO WE MEAN BY "USING DATA?"

Whereas obtaining data is the first step in any data practice, an analysis that leads to useful conclusions about data is the end result. For this Guidebook, data analysis is the practice of working with data for visualization, modeling, or communication about data that informs decision-making. The first section of the Data Practices Guidebook spoke to the ways agencies collect, manage, and share data, while the final section of the Guidebook shows how agencies are using data for planning and performance monitoring, for building new partnerships, and for analyzing access and equity for their transit systems.

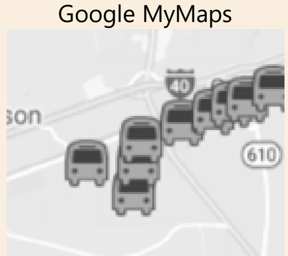
Conventional data analysis practices include loading, cleaning, and visualizing data in Excel and reporting in other documents. Even if your agency does not have access to GIS software, online tools like [Data Wrapper] or [Google MyMaps] can help you create maps in a web browser, while [Google Earth Pro] is an easy-to-use substitute for viewing geographic layers found in many open data portals. Though quantitative analysis of spatial data is not readily performed with these tools, visualizing distributions of data on a map are the start of many spatial data analyses.

User-Friendly Tools for Data Analysis

Data Wrapper



Google MyMaps



Google Earth Pro

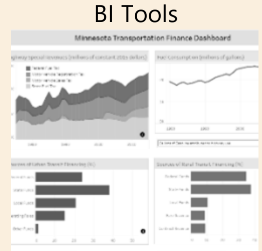
Automatic or passive data sources are complex and large in size, requiring new tools and techniques to use. While vendors of APC and AVL devices offer their own tools to extract and report data, larger agencies may rely on third-party tools to store and analyze data. Large agencies increasingly rely on data warehouses or data lakes to store raw or cleaned data in a way that can be more readily analyzed using third-party applications like Ridecheck Plus or Swiftly, processed with statistical programming languages like Python or R, or visualized with business intelligence (BI) software like Tableau or Microsoft PowerBI.

Tools for Analyzing Larger, Complex Datasets

Third-Party Apps

Programming Languages

BI Tools



Regardless of the data source or method, determining the proper question to ask of the data is key. This Guidebook will show how agencies large and small have addressed questions through data analysis.

4. PLANNING AND PERFORMANCE MONITORING

Transit service planning and performance monitoring are easier with more data. Using this Guidebook, agencies can tap information on their customers, performance, and asset conditions using a variety of data collection and management practices. With this data in hand, agencies can address central questions of transit planning: Who rides transit? Where do they want to go? For some small agencies, the first-hand experience of drivers and agency staff can begin to answer these questions. However, data-driven planning decisions are the key to leveraging limited resources to the greatest benefit for all users of a transit system.

Data-driven planning decisions are the key to leveraging limited resources to the greatest benefit for all users of a transit system.

Planning

Population, Employment, and Land Use

U.S. Census data contains detailed socio-economic information on. Planners identify major trip origins and destinations using this data and then modify service to generate increased transit demand.

Ridership Forecasting

Ridership impacts of service changes can be estimated using just existing ridership and a few common demand elasticity values.

Transit Demand

Existing ridership data is the clearest indicator of demand for transit. Analyzing this data involves aggregating ridership data to observe patterns of demand.

Travel Demand Patterns

Using model data provided by state or regional planning bodies, agencies can identify important connections and design transit services to accommodate those trips.

Capital Planning

Once transit assets have been inventoried and assessed, agencies can put this data to work to plan for investment needs in the years to come.

These findings are from the Data Practices Guidebook. The Guidebook is a resource to assist small urban, rural, and tribal transit agencies in understanding and applying good data practices.

Performance Monitoring

Ridership Data

Ridership performance can be evaluated by identifying trends in ridership totals, measuring passenger load and service capacity, and calculating service productivity metrics.

Schedule Performance Data

Performance monitoring of actual schedule data allows agencies to evaluate service as it is delivered to customers. With these measures, agencies can work to improve services.

Financial Data

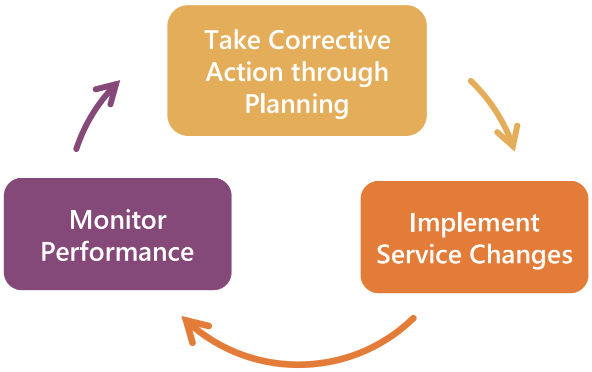
Monitoring financial performance helps agencies keep their budgets in check. Farebox recovery is one of the most important financial performance indicators, among others.

4.1. Introduction

Transit service planning and performance monitoring are easier with more data. The data practices outlined in the previous chapters of this Guidebook provide agencies with the means to tap into information about the ridership, operations, assets, and financials of their transit service, as well as the characteristics and needs of their current and potential customers. With this data, agencies can use analytical practices to inform the service planning process and monitor the performance of existing services.

Planning and performance monitoring are closely intertwined processes linked through data. When agencies conduct service planning to realign routes, change the frequency or span of services, or make other modifications, they can use data on existing ridership, demographics of the areas served, and other performance statistics to guide these decisions. Once services are implemented, performance monitoring data informs evaluation of the changes and suggests routes or portions of the service area that may need further service planning attention.

Figure 23: Cycle of Planning and Performance Monitoring



While the use of data to inform these decisions is not new, the richness of new data sources provides agencies with the ability to conduct more precise evaluations of performance and to plan services with greater clarity. For instance, monthly ridership averages can be supplemented with trip-level overcrowding statistics from Automatic Passenger Counters, while Location-Based Services (LBS) data can provide more insight on passenger travel than Census travel data alone. Even if an agency does not make use of new methods, a facility with data practices can speed the reporting of data for performance-based grants, planning studies, and other requirements.

Even as they provide more insights, the ever-larger and more complex datasets discussed in this Guidebook place new burdens on agencies. Working with data like this requires both technology investments and staff know-how. This Guidebook provides a range of data practices for planning and performance monitoring, some of which may be readily implemented and others that may require the help of outside partners. In Chapter Five, partnership opportunities for building working relationships around data will be discussed in additional detail.

4.2. Planning

For transit service planning, data provides key insights for decision-makers when designing and evaluating service changes. Though some data like ridership and revenue hours will be tabulated by the agency itself, external datasets such as Census population and employment estimates will play a role in many planning decisions. As services change and fleets age, future year capital investment needs can also be made on the basis of asset inventory and condition data. The key data sources discussed in this Guidebook lend themselves to a broad set of planning methods, as outlined in **Table 3**.

Table 3: Data-Driven Service Planning Methods

Category	Input Data	Method	Output Analysis
Ridership Analysis	Ridership data by location and time	Visualize where and when activity occurs using GIS software or in tables and charts	Transit demand: Where are customers using your services?
Ridership Analysis	Existing ridership data, existing and future fare and service level data	Estimate ridership using common elasticities	Ridership forecast: How will changes to fares or service levels affect ridership?
Market Analysis	Census data, land use data	Plot Census variables like population by their density or concentration	Market analysis: Where do potential customers live?
Market Analysis	Travel demand model, LBS data	Visualize travel patterns	Demand analysis: Where do people want to go?
Capital Planning	Asset condition, relevant financial data	Use Decision Support Tools to forecast capital needs, prioritize improvements	Financial forecast: Transit Asset Management Plan, Fleet Plan, and/or Capital Improvement Plan investment priorities

4.2.1. Ridership Analysis

Ridership data is at the heart of understanding utilization and the demand for transit service. When considering service changes to a fixed route or the operational capacity of demand response services, planners can make use of agency-collected boarding and alighting data to evaluate how changes will impact ridership levels and service efficiency. Service changes also impact future ridership, which can be estimated from existing ridership data using common elasticities for fares, span of service, and frequency or wait time.

Existing Transit Demand

Demand for transit varies widely by the time of day, day of the week, and location. Data-driven planning decisions rely on ridership data with both spatial and temporal variation to understand these fluctuations in demand. With this information, planners can evaluate the frequency, span, and geographic extent of service. To evaluate the existing demand, planners may use data from any ridership data source, including ridechecks, electronic passenger counters (EPC), automatic passenger counters (APC), or trip logs for demand response services. Spatial data may be present in the form of stop identifiers or coordinates, while temporal data may be represented by a timestamp, a date, or a specific trip identifier.

At the greatest granularity, fixed-route transit demand can be analyzed for every stop, enabling a comparison of ridership levels between stops along a route. Boardings at the stop level can also be combined for all routes at each stop to gain an understanding of transit demand at a certain location, regardless of route. Demand-response transit demand can be analyzed at the pick-up/drop-off location level, enabling an understanding of where in the service area there is the greatest demand. With less granularity, planners may compare ridership between different routes or services or observe trends in route-level ridership over time.

Data Granularity and Resolution for Transit Planning

How data can be put to use is often a question of its **granularity** and **resolution**.

Data granularity is the level of aggregation in a dataset: each record may represent a unique stop, trip, vehicle, direction, or route. Granularity is important for nearly all planning applications since most measures can be analyzed at any level. Planners must choose the right granularity depending on the question they wish to answer, whether the analysis should be performed separately for each stop (high granularity) or for each route (low granularity). Highly granular datasets are the most advantageous for analysis since they can usually be aggregated to a lower granularity if needed.

Data resolution is the precision of a dataset: a smaller time interval or more precise coordinates would result in a higher resolution. High-resolution data is usually a characteristic of passive data sources, such as automatic passenger counting (APC). Most planning decisions can be made without high-resolution data, but agencies with greater data capabilities are starting to use such data for new insights. For example, a detailed spatial and temporal analysis of APC data enables planners to diagnose potential causes for bus overcrowding, such as dropped trips, demand variability on an hour-to-hour and day-to-day basis, and lack of headway adherence.

For instance, an agency may have highly granular data on stop-level ridership activity from a manual ridecheck, but this data may lack resolution because ridechecks are conducted infrequently. On the other hand, an agency may have high resolution, trip-level boarding counts from driver logs, but the data may lack the granularity of a stop-level ridership activity dataset.

A wide variety of analyses can be conducted to understand transit demand, depending on the level of aggregation and the time frame in question. First, ridership data is aggregated to the desired level of detail by taking the sum of all boardings for a stop, trip, or route. Then, the data is summarized temporally, whether planners need the information broken out by the time of day, day of the week, or at a larger scale such as months or quarters. If coordinates are available, the data can then be plotted using geographic information system (GIS) software to show where transit demand is the greatest.

The planning applications for existing transit demand analysis include evaluating changes to the span and frequency of service, capacity analysis, and bus stop planning. Bus routes with high demand may warrant higher frequency or larger vehicles to meet capacity needs. Conversely, routes with low ridership may require new approaches to attract riders or may be candidates for elimination or conversion to demand-response, micro-transit style services. At the stop level, planners can analyze demand to prioritize implementation of stop amenities or to identify stops for removal when consolidating stops on a fixed route. At the trip level, high ridership at the beginning or end of the span of service may suggest extending the hours of service or modifying the frequency of service during particular time periods.

Visualizing Transit Data

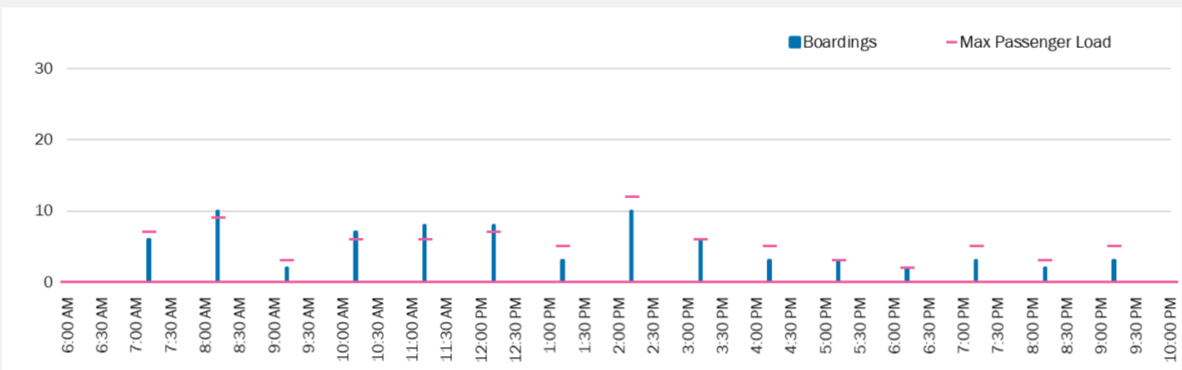
Transit data is often multidimensional, containing quantitative measures, geographic locations, and temporal information with every piece of data. Agencies must distill this data to communicate clearly with stakeholders and the public and inform decision-making. The richness of transit data often requires more information-dense graphics, charts, and maps, though these can be more challenging to produce. Effective data visualizations are:

- Easily digestible by the intended audience
- Understood independently of other visualizations
- Clear in purpose and message
- Aesthetically attractive.

Most transit data contains two key components for visualization: **time** and **space**. Combined, these two elements produce rich illustrations that communicate many layers of information.

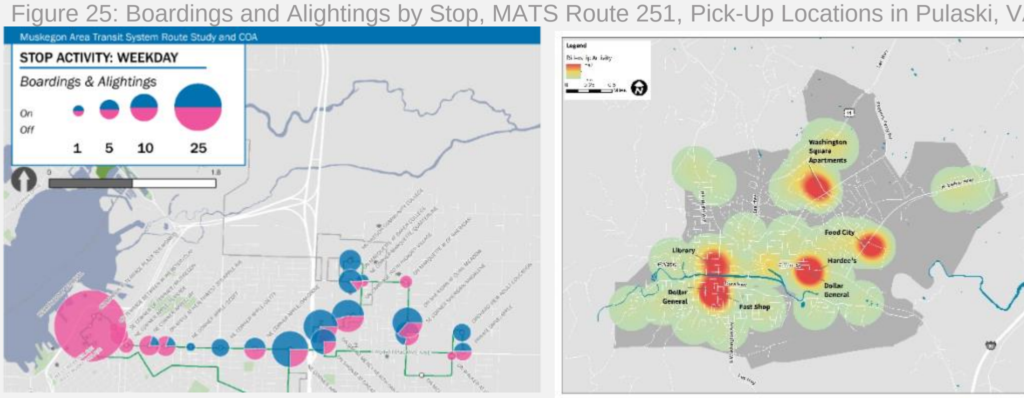
For high-resolution data, **time** can be visualized using one axis of a graph or table, such as in **Figure 24**. For low-resolution (or aggregate) data, variations in data over time may be shown by repeating a visualization for each time period. For example, a map of ridership levels may be repeated three times to show weekday, Saturday, and Sunday data separately.

Figure 24: Average Total Boardings by Trip, Muskegon Area Transit System (MATS) Route 251



Space is often—but not always— visualized on a map. As demonstrated in **Figure 25**, a map provides geographic context to data, enabling readers to relate the data to their own familiarity with a location. Alternatively, the spatial element of a dataset is sometimes one-dimensional, allowing for space to be represented by one axis of a graph or table. For example, the boardings at stops along a route may be visualized using a bar graph since there is one spatial dimension: distance traveled along the route.

Figure 25: Boardings and Alightings by Stop, MATS Route 251, Pick-Up Locations in Pulaski, VA



Ridership Forecasting

When riders choose to use transit, they often compare the cost of transit to other transportation options—or not making a trip at all. The cost to the rider is not just the fare but also the time spent in transit. In addition, the service must be available when and where the rider needs it. Ridership depends on these factors, and service changes that alter the fare, wait times, or overall availability of service will result in impacts to ridership. While there are many ways to estimate future ridership, changes caused by fares, wait times or frequency, and span of service can often be forecast by simply applying a common demand elasticity to existing ridership data. More information on demand elasticities can be found in the following call-out box.

Other methods of ridership forecasting are available that require more data and technical expertise. Software such as STOPS (Simplified Trips-on-Project Software), created by FTA, and TBEST, created by the Florida Department of Transportation, enable agencies to include the impacts of population, land use, and geographic coverage in ridership estimates for fixed-route services. For demand response services, some research has also provided methods for calculating ridership based on population, fares, reservation policies, and other service characteristics.

Estimating Ridership Changes Using Elasticities

Elasticity is an economic concept that measures the percentage change of one economic variable in response to a change in another. In transit, changes in the service supplied, and fares for service can lead to changes in ridership.

For both demand response and fixed-route services, ridership and fares typically have an inverse relationship, such that fare increases result in ridership decreases and vice versa. This relationship is represented by a theoretical ratio called the "fare elasticity," typically ranging from -0.1 to -1. For example, a fare elasticity of -0.5 suggests that a 10 percent increase in fares results in a 5 percent decrease in ridership. Fare elasticity values vary according to the size of ridership impact—for some agencies, a fare increase will result in a significant ridership decrease, whereas others may experience little to no change in ridership.[52]

Demand elasticity values can also be used for estimating the ridership impacts of span and frequency changes to fixed-route services. Span elasticities are positive since an increase in the number of hours or days of service causes an increase in ridership. For frequency, demand is measured relative to the headway, or the number of minutes between arrivals on a route. Headway elasticities are negative, indicating that as the headway increases, ridership will likely decrease[53]. Research on span and frequency elasticities for demand response is limited, but studies have shown that demand response services experience similar relationships. An increase in the hours or days of service will increase ridership on a demand response service. Analogous to increasing frequency on a fixed route, a

[52] Todd Litman. 2004. "Transit Price Elasticities and Cross-Elasticities." Journal of Public Transportation,7 (2): 37-58. https://www.nctr.usf.edu/wp-content/uploads/2010/03/JPT-7-2-Litman.pdf.
[53] The National Academies of Sciences, Engineering, and Medicine. 2004. "TCRP Report 95: Traveler Response to Transportation System Changes Handbook, Third Edition: Chapter 9, Transit Scheduling and Frequency." Washington, DC: The National Academies Press. http://www.trb.org/Publications/Blurbs/154748.aspx.

Estimating Ridership Changes Using Elasticities

reduction in the amount of time between booking and taking a trip has also been shown to increase ridership.[54]

Existing research may indicate appropriate elasticity values for transit services with a given set of characteristics. Such elasticities are generally calculated using existing data from other transit agencies. To select an elasticity value for analysis, agencies must consider the applicability of the source, including:

- Geography, such as service area population, density, and regional similarities
- Type of transit service, including mode and quality or attractiveness of service
- Cost and availability of other transportation options, such as the presence of other transit services or the price of gas
- Ridership impact of past fare changes to evaluate whether an elasticity value matches observed conditions
- Rider characteristics, including demographics such as age, income, and vehicle ownership.

When performing ridership estimation with demand elasticities, agencies should consider riders that may exhibit different demand behaviors than other riders. For example, riders who do not have access to a personal vehicle and rely on transit may be less sensitive to service changes. Passengers who pay a reduced fare (or none at all) or with no other options may not change their behavior due to a fare change. In some cases, it may be appropriate to segment the ridership data and analyze the impact of service changes separately for each group. To identify and isolate such groups, planners may examine on-board customer surveys or fare collection data.

4.2.2. Market Analysis

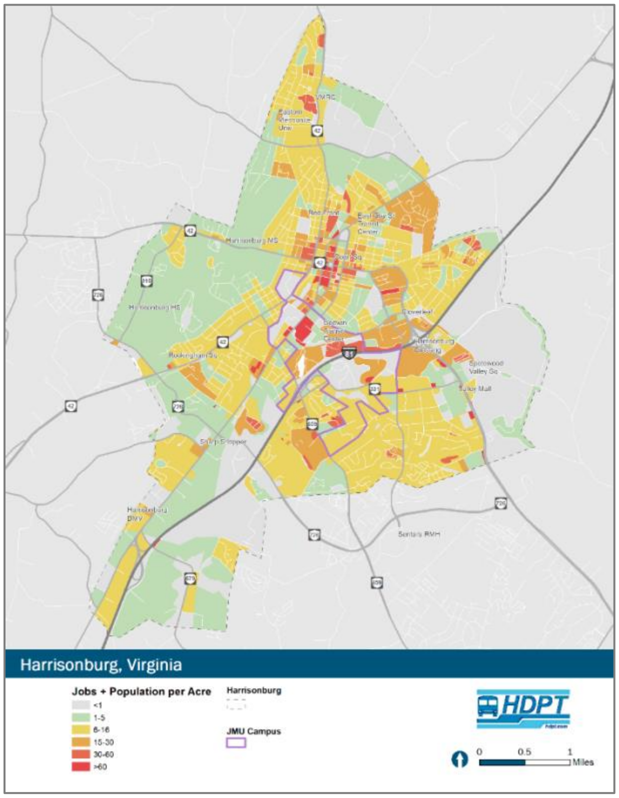
In addition to understanding the existing demand for transit, planners often aim to increase ridership by leveraging the potential market for transit. A market analysis evaluates who would be likely to ride transit, where they live, and where they want to go. Typically, many people within the market who would ride transit do not currently do so. Since agencies only collect data about current users of their services, planners must turn to external data sources, such as the U.S. Census and regional transportation demand models.

Population, Employment, and Land Use

The transit market consists of two types of locations: places of residence and places of employment. Most transit trips begin or end at home. At the other end of the trip is almost always a place of employment, whether it is the rider's own workplace or a business, medical center, or another point of service where others work. Therefore, understanding the geographic distribution of population, employment, and their corresponding land uses is critical to measuring the potential of the transit market.

[54] National Center for Transit Research. 2016. "Estimating Ridership of Rural Demand-Response Transit Services for the General Public." Fargo, ND: Small Urban and Rural Transit Center. https://www.nctr.usf.edu/wp-content/uploads/2016/09/21177060-NCTR-NDSU08.pdf.

Figure 26: Jobs and Population Density, Harrisonburg,



The most consistent, reliable, and accessible data source for a transit market analysis is the U.S. Census, including the American Community Survey (ACS) and Longitudinal Employer-Household Dynamics (LEHD) datasets (see Chapter Three: Open Data for more detailed information). These surveys contain a wealth of information about population and employment estimates in the United States, with details on demographic information such as age, gender, income, race, and vehicle ownership. ACS data is available for small geographic areas called "block groups," such that planners may use this data to analyze population and employment at bus stops or for individual neighborhoods. However, for more rural jurisdictions, the geographic size of block groups will still be impractically large for detailed analysis. In these cases, a simple examination of population counts at the "block" level will suffice. LEHD employment data is also available at a detailed block level, but many transit services will not be focused on the commuter market. In those cases, the locations of major destinations like grocery stores, hospitals and clinics, and other services can be used to identify the possible destinations of trips.

Market analysis must consider the unique makeup of the market for different transit services. For example, the market for a commuter service may consist of working-age adults, whereas the market for demand response or local bus service may be oriented toward non-work (shopping, medical, or recreational) trips by residents with limited access to a personal vehicle. A valuable source of data on the target market for transit service is on-board customer surveys. By evaluating what demographics of riders and which trip purposes are currently served, planners can conduct a market analysis that is most relevant to their community and make equitable planning decisions.

Transit planners use GIS or other mapping software to measure the geographic distribution of certain demographics and places of employment. The density of both people and jobs are important market analysis metrics, as they indicate where there is a concentration of potential transit riders. The result of population and employment analysis is a picture of the areas with potential for transit demand. In some cases, these areas may already have high transit ridership. Comparing existing services and current ridership levels to these results enables planners to identify geographic gaps in the transit network or opportunities for increased level of service.

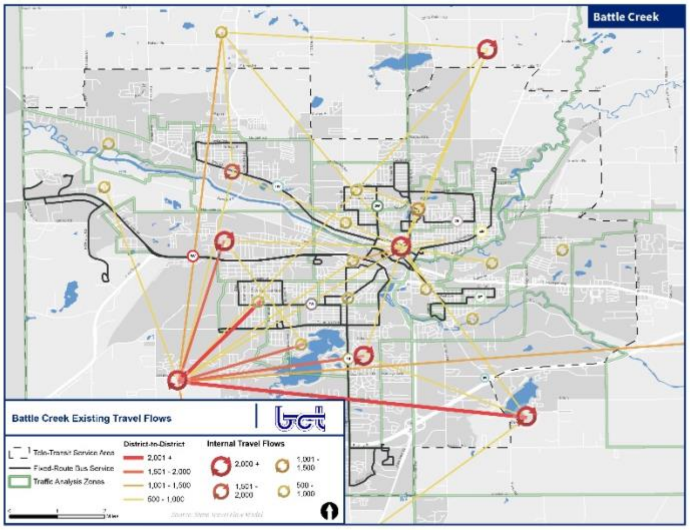
Travel Demand Patterns

If population and employment analysis shows the potential origins and destinations for transit trips, travel demand data shows the connections between those trip ends. Many states and regions maintain a travel demand model that produces matrices of the number of trips between each origin and destination zone within the region. Trips are categorized by purpose and mode, enabling transit planners to visualize and evaluate the potential demand for a transit service along a corridor or between two locations.

Transit ridership is highest when a service gets the most people where they need to go, and travel demand model data provides a wealth of information about the most common trips made by residents or workers in a certain area. Just like population and employment analysis, most of the flows in the travel model represent trips that are not currently made with transit but could be if there was a service to those origins and destinations. Planners can use travel demand data to identify key origin-destination pairs that could be served by transit, or at a higher level, entire corridors with high travel demand.

To analyze travel demand data, spatial visualizations are used to highlight connections with the greatest volumes of flows. When overlaid on the existing transit network, planners can more easily identify geographic gaps where many people are traveling, but there is limited or no transit service available (see example in **Figure 27**). Similar to U.S. Census data, travel demand data is available at a small scale, which allows for analysis at varying granularity. Many models also produce data by trip purpose or time of day, so that planners can drill down on travel patterns according to the specific types of destinations that should be served, and when during the day they should be served.

Figure 27: Travel Flows in Battle Creek, MI from a Regional Travel Demand Model



While calibrated based on real-world data, travel models do not necessarily produce results that match real conditions. Especially in future year forecasts, these estimates have non-negligible limitations and uncertainties. To address this shortfall, an emerging data source for analyzing travel demand patterns is aggregated cellphone location data (also known as location-based services datasets, or LBS), which produces travel volumes with actual observations for every individual trip captured. While this data must be purchased from private companies, this data can be used to understand travel patterns for a transit market analysis.

Capital Planning

Once transit assets have been inventoried and assessed, agencies can put this data to work to plan for investment needs in the years to come. Processes of investment prioritization vary widely across the transit providers; for rural, tribal, and small urban ones, state and regional government bodies have a large role to play in this process. Common to each prioritization process is an assessment of the cost to maintain agency assets in a State of Good Repair, as well as consideration of other priorities and criteria. These criteria may include regulatory requirements, new needs, equity, customer service enhancements, or other criteria that may inform decisions. These decisions are often supported through software such as TERM Lite or the Transit Asset Prioritization Tool (TAPT) and codified in Transit Asset Management Plans and Capital Improvement Plans.

Outside of the investment decision-making process, asset and maintenance data have a role to play in reporting on the condition of the health of the fleet and its ability to deliver services to the public. A common metric reported at the agency or sub-fleet level is the Mean Distance Between Failure (MDBF). This figure is calculated by totaling the number of component failures onboard vehicles and dividing against the number of revenue miles traveled by the fleet in a given period.[55] The resulting rate is a tangible indicator of how well maintenance procedures are preventing disruptions to service given the age of the fleet.

While asset condition is well understood to play a role in service delivery, to date, few agencies have attempted to tie asset conditions to on-time performance, ridership, or other service performance metrics. Often this is because data on a vehicle's condition, a vehicle's assignment to a route, and a route's performance are tracked in a variety of separate data sources that are difficult to connect to one another at matching levels of granularity, if the necessary data exists at all. In time, agencies may be able to differentiate what share of on-time performance are due to mechanical issues and not due to traffic or dispatch challenges or understand the effect on ridership from the adoption of new, more attractive vehicles. While that trend continues to emerge, plenty of data remains for agencies to analyze in their performance monitoring programs.

4.3. Performance Monitoring

As communities grow, travel behaviors change, and agency conditions fluctuate, so too do metrics like ridership, on-time performance, service efficiency, revenues, and costs. Transit performance monitoring is the process of reporting a set of performance measures repeatedly over time. While rural, tribal, and small agencies already report annual performance data to the National Transit Database (NTD) or for other state and regional transportation plans, performance monitoring programs are typically geared towards reporting on a more frequent, targeted basis to support adjustments to operations or near-term planning decisions. In this sense, the design of a performance monitoring program is a concerted effort involving not only data but also an agency's broader goals and objectives.

However rural, tribal, and small urban agencies structure their performance monitoring efforts, new data sources provide new opportunities to gauge performance. Sensor-based data from automatic passenger counters enable nuanced capacity analysis, while high-resolution data from automatic vehicle location devices allow for detailed evaluation of on-time performance and runtimes. Automatic fare collection data facilitates analysis of revenue and cost-efficiency measures. Even without passively collected data, agencies can take advantage of regularly collected ridership, trip logs, and fare revenue data to analyze a robust set of performance metrics. The approaches for performance monitoring enabled by these datasets are summarized in **Table 4**.

Table 4: Data-Driven Performance Monitoring Methods

Category	Input Data	Method	Output or Analysis
Ridership data	Ridership data (e.g., APC) by route and stop for fixed-route services	Calculate the passenger load leaving each stop	Capacity analysis: Whether your vehicles have enough space and run frequently enough to meet demand and meet social distancing requirements
Ridership data	Ridership data by hour for demand-response services	Visualize the number of trips per hour (e.g., heatmaps)	Capacity analysis: Whether driver shifts or passenger pick-up times need to be adjusted

[55] For some agencies, failure of some component on a vehicle (e.g., the air conditioner) may not in fact result in a road call or delays to passengers. Instead, these agencies may also report the Mean Distance Between Delay (MDBD), which reflects only events that cause delays.

Category	Input Data	Method	Output or Analysis
Ridership data	Ridership data by month or year	Create a line chart of ridership over time	Trend analysis (by route or systemwide): How is demand changing over time?
Ridership data	Ridership data relative to peers	Use iNTD database to gather performance data about peers	Peer analysis: How does your system compare to peers?
Schedule data	Schedule adherence and runtime (e.g., AVL) data	Calculate rate of on-time performance	OTP analysis: Are routes arriving at time points on time?
Schedule data	Service produced and service consumed	Calculate common performance measures	Performance analysis: Do performance measures support changes to services?
Financial data	Fare revenue and operating costs	Analyze the difference between costs and revenues	Farebox recovery analysis: How much do fares offset operating costs?

From a data perspective, effective performance monitoring makes use of metrics that are consistent, accurate, and frequently available. Before conducting a performance monitoring analysis, agencies should define a set of performance metrics that they will calculate for every analysis period, whether on a monthly, quarterly, or annual basis. Some agencies will set targets for each metric, such as an on-time percentage goal or a minimum farebox recovery rate. With consistent source data and consistent metrics, trends or patterns can easily be identified over time. Accuracy in performance metrics requires data with small uncertainty and ensures that high-quality data is compared to high-quality data. Frequently-collected data allows agencies to repeat performance analysis regularly. Together, these considerations help reduce the level of effort required to perform the analysis since the same calculations can be repeated each time a new performance monitoring period has been completed.

4.3.1. Ridership Data

Many key performance metrics use ridership data to give insight into changing transit demand and the productivity of routes or services. Ridership performance metrics may indicate the need for a change in the level of service, highlight areas with frequent issues, or reveal low-productivity routes that should be improved. For the collection of ridership data, agencies may use manual methods, such as ridechecks or electronic passenger counters (EPCs), or passive methods, including automatic passenger counters (APCs). Ridership performance can be evaluated over a variety of time periods, including service periods, days, months, quarters, and years.

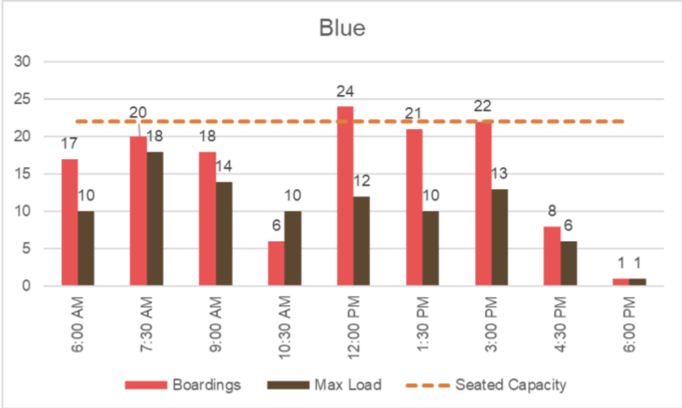
Ridership Totals

Any agency that reports data to NTD is accustomed to calculating total ridership on an annual basis, but analyzing ridership totals at a variety of levels can provide new insights when monitoring performance. Annual totals, both systemwide and for individual services, are especially useful for long-term trend analysis. With this type of Rural Integrated National Transit Database (iNTD) data, agencies can compare ridership levels to peer agencies by calculating ridership per capita. However, if data is collected more frequently, short-term trends such as seasonal variations in ridership may become apparent. Additionally, agencies can quickly identify the impacts of service or operational changes on demand for their services. For demand response services, ridership performance can also be analyzed by plotting the geographic distribution of trips using GIS, including the frequency of pickups and drop-offs in each area.

Passenger Load and Capacity

With boarding and alighting data at the stop and trip level, agencies can evaluate overall service capacity. Passenger load, which measures the number of passengers on a bus at a given time, is calculated by subtracting alightings from boardings at each stop as a vehicle travels along a route. When passenger loads meet or exceed vehicle capacity, a route may require larger vehicles or increased frequency (see example in **Figure 28**). Low passenger load may be important for social distancing, but many empty buses on a route may indicate a lack of demand. For demand response services, the service capacity can be measured by analyzing trip request data to identify the maximum number of trips that can be provided. Additionally, capacity can be measured using the average wait time for non-scheduled trips – excessive wait times can also indicate that there is greater demand than can be provided. If this is the case, agencies may consider hiring additional drivers or implementing computer-aided dispatch (CAD) for more efficient routing.

Figure 28: Example of Passenger Loads and Seated Capacity on a Fictional "Blue" Route



When demand outweighs supply: Pulaski Area Transit's shift to fixed-route service

[Pulaski Area Transit]() (PAT) in rural Pulaski, VA, recently switched from operating general-public demand-response to a fixed-route system. When operating only demand-response service, PAT received in excess of 700 calls per day for service and completed nearly 500 trips per day. This put tremendous stress on their fleet of nine vehicles and led to excessive wait times for passengers – often exceeding 45 minutes on busy days. Tracking their daily ridership totals and wait times enabled the agency to conclude that the demand in their service area far outweighed the amount of service they were able to supply. Subsequent studies, including a fixed-route analysis and a Transit Development Plan, recommended implementing the two fixed routes with a 30-minute headway each, which enabled a more efficient operation and less wait time for customers.

Figure 29: Pulaski Proposed Fixed-Route Network

Service Productivity

The productivity of a transit service measures the number of passengers served relative to the amount of service provided. Common metrics for service productivity include boardings per revenue hour and boardings per revenue mile, as well as boardings per trip for fixed-route services. Higher productivity indicates a more efficient service that is able to provide more trips with the same resources. However, for demand response services, productivity may be directly related to how far passengers are traveling rather than service efficiency. Productivity can be measured at any granularity for which ridership data is available. By comparing the productivity of fixed routes, agencies can also more easily make planning decisions such as prioritizing routes for service improvements, planning level of service changes and identifying services for elimination.

4.3.2. Schedule Performance Data

Performance monitoring of schedule data allows agencies to evaluate service as it is actually delivered to customers. Schedule performance metrics can be compared to planned service, other routes, and peer agencies to determine the quality of service and identify areas for improved operations. The primary sources for schedule data are trip logs or automatic vehicle location (AVL) data. Using this data, planners can measure the actual hours and miles of service provided, on-time performance, delay, dwell time, travel speeds, and actual runtime.

Revenue Hours and Miles

Actual vehicle revenue hours and vehicle revenue miles are some of the most important metrics for any transit service. Beyond using this information for reporting, these measures are central to understanding differences between planned and actual service. Dropped trips, variations in runtimes, and other unpredictable influences on delivered service are often evident in revenue hours and miles. In addition, these metrics can be compared to peers to evaluate whether agencies of a similar size or in a similar geographic context are providing more or less service. Overall trends in revenue hours and miles also provide important context to other performance monitoring metrics. For example, increased revenue hours will typically correlate with an increase in operating costs. Moreover, revenue hours and miles are the foundation of many performance metrics, including service productivity and cost-efficiency.

On-Time Performance

Differences between scheduled and actual arrival times impact riders significantly, who can experience long wait times or miss the bus entirely when schedule performance is poor. The percentage of arrivals that fall within an agency's defined on-time window, or the on-time performance (OTP), is a key schedule performance metric that centers on the passenger experience. Arrivals include every stop made by a fixed route service or all pickups and drop-offs for an on-demand service. On-time performance can be reported for individual trips or summarized across a longer time period. By measuring on-time performance trends over time, agencies can make operational improvements such as reducing boarding delays, schedule adjustments, or transit signal priority to improve schedule adherence.

Runtime and Speed

In addition to on-time performance, agencies use schedule data to measure actual runtimes and vehicle speeds. For fixed-route and demand-response services, accurately planned runtimes are critical for scheduling and maintaining on-time performance and accurate pick-up/drop-off times. Over time, road conditions, traffic patterns, and congestion cause travel speeds and runtimes to change, and agencies rely on recorded vehicle location data to determine how long it takes to run a route. Vehicle revenue miles divided by vehicle revenue hours may also serve as a proxy for measuring average travel speed while in service. As a performance metric, agencies may use runtime and speed data to identify where transit signal priority, stop consolidation, or other speed improvement measures are needed.

4.3.3. Financial Data

Basic fare revenue and operating cost data are often used to evaluate the financial performance of transit services. Understanding the financial state of individual transit services, as well as an entire system, are key to maintaining a balanced budget and planning future service within cost constraints. Financial performance monitoring can also help agencies compare their transit investments against peers and identify routes or services that have the lowest return on investment.

Cost Efficiency

Operating cost data at the route, service, or system-level enables agencies to calculate exactly how efficiently they are using resources to provide service. Three common financial performance metrics for cost efficiency are operating costs per revenue hour, revenue mile, and passenger trip. Over time, operating costs depend on the price of fuel, maintenance costs, wage increases, and other changes to operational procedures as applicable. Between routes, the cost per hour or mile may vary depending on travel speeds, delay, layover duration, and deadhead time and distance. By measuring costs on a per-hour and per-mile basis, planners can identify sources of costs and may be able to implement measures to reduce delay or minimize operational inefficiencies. However, some routes with higher operating costs may have a higher return on investment if ridership is also high. The operating cost per passenger, therefore, reveals a side of cost efficiency that is focused on serving the most passengers with the same resources. Along with reducing operational inefficiencies, agencies may also aim to increase ridership to best leverage their resources.

Farebox Recovery

In order to keep a balanced budget, agencies must track revenues as well as operating costs. Farebox recovery is one of the most important financial performance indicators, representing the ratio of fare revenue divided by operating costs. Although most agencies do not charge passengers the full cost of service, farebox recovery is useful for evaluating what percentage of the cost of service is covered by fares and using the data to forecast future revenues from new or changed services by estimating the farebox recovery.

Another metric for measuring farebox recovery is the average subsidy per passenger or the dollar amount of operating costs that are not offset by passenger fares. Since many agencies offer reduced or free fares to some passengers, keeping track of subsidy per passenger allows agencies to plan for financial trends in the relationship between fare revenues and operating costs over time for a balanced budget.

Identifying Peers for Comparison

"Compared to what?" is the key question underlying most data analysis. For many agencies, the value of a performance measure will be tracked over time, across routes, or relative to agency targets in order to provide insights on how the agency is performing. For the agency as a whole, comparisons to peers can provide a more holistic picture of how the agency performs, especially when factors outside an agency's direct control (such as gas prices) are changing passenger behavior.

When conducting a peer comparison, determining the set of peer agencies to examine is often the most difficult step. While every agency faces a unique set of constraints and customers that make a true "apples-to-apples" comparison impossible, the Rural iNTD website provides a tool to make identifying peer agencies straightforward. For a given agency, the tool will assess the similarity of other agencies based on nine metrics in three categories:

Screening factors:

- *Agency Type*: Whether the agency is part of a tribe, local government, non-profit, private provider, or another agency type
- *Operates Commuter Bus*: Whether the agency provides commuter bus service
- *Percent Motorbus Revenue Hours*: The amount of service provided that is fixed route relative to demand-response service
- *Headquarters City in an Urbanized Area*: Whether the agency is based in an urbanized area

Similarity factors:

- *Annual Vehicle Revenue Miles Operated*: A metric that encapsulates each agency's level of service
- *Percent 5310 Funding*: Percent of operating funds sourced from federal support for the mobility of seniors and persons with disabilities
- *Percent local funding*: Percent of operating funds from local sources
- *Population of Headquarters City*: Because the specific extent of an agency's service area is not reported in NTD, the population within the agency's headquarters' city is used

Proximity factor:

- *Proximity*: The distance between the home ZIP code of the given agency and a possible peer agency.

The likeness scoring method used by NTD will then weight and aggregate these metrics to produce a ranked set of peers.

Once a set of peer jurisdictions have been identified, statistics on each agency can be readily compiled from the Rural iNTD database or from the National Transit Database (NTD) directly. Because peer agencies will still differ from your agency in the extent of their services and populations served, it is important to compare rates rather than levels. For instance, because peer agencies may serve different-sized populations, consider making comparisons of ridership per revenue mile rather than ridership alone. Examining these rates can lead to further questions to consider: does your agency require more operator hours relative to the amount of service delivered? If so, it may be worth examining whether services are scheduled efficiently. After you've made comparisons, high-performing agencies in your agency's peer group can then be directly contacted for insights into their performance.

4.4. Conclusion

In the cyclical process of transit service planning, implementation, and performance monitoring, transit data sources offer a wealth of information to make decisions easier. Data on ridership, population and employment, travel demand, customer surveys, agency assets, schedule performance, and finances each lend themselves to a variety of analyses as agencies plan service changes and monitor all aspects of transit performance.

Planning makes use of both internally collected data and data from external sources, such as the U.S. Census. With ridership data, agencies can evaluate existing transit demand and calculate the potential growth of ridership with new services. Census data on population and employment characteristics, as well as travel demand model data, reveals where potential new transit riders live and want to go, such that planners can alter or expand service to capture that demand. Capital planning efforts rely on asset inventory and maintenance data to evaluate the condition of assets and prioritize future investments.

Performance monitoring leverages regularly collected data to track the success of service changes and identify trends or patterns to plan a more efficient and productive service. Ridership data can be used to monitor overall ridership, as well as evaluate service capacity and productivity. Schedule performance can be evaluated with metrics including revenue miles, revenue hours, on-time percentage, and runtimes. Financial data, including passenger fare revenues and operating costs, can be tracked with cost efficiency and farebox recovery measures to ensure a balanced budget. In all areas, agencies can compare performance metrics to peer agencies to gain context for their service's performance.

Data-driven planning and performance monitoring require both technology investments and staff know-how. While this data provides new insights, it also increases the level of effort required to manage data and perform analyses. In the next chapter, learn how partnerships with other organizations can broaden agencies' abilities to gather, analyze, and share data to use internal resources and create better services efficiently.

Checklist: What types of analysis does your agency perform in planning efforts?

Analysis	Currently Perform this Analysis	Want to Perform this Analysis
Ridership geographically	☐	☐
Ridership temporally	☐	☐
Ridership forecasting	☐	☐
Population and employment geographically	☐	☐
Locations of major services	☐	☐
Asset conditions	☐	☐

Checklist: What types of performance data does your agency monitor?

Type	Metrics	Currently Monitor	Want to Monitor
Ridership totals	Daily ridership	☐	☐
	Annual ridership	☐	☐
	Ridership by stop or pick-up/drop-off location	☐	☐
Passenger Load and Capacity	Passenger loads by trip	☐	☐
	Passenger loads by stop	☐	☐
	Passenger loads compared to vehicle capacity	☐	☐
	Number of demand-response trips that can be provided	☐	☐
	Wait time for non-scheduled demand-response trips	☐	☐
Service Productivity	Passengers per revenue mile	☐	☐
	Passengers per revenue hour	☐	☐
	Passengers per trip	☐	☐
Schedule Performance	Actual revenue hours operated	☐	☐
	Actual revenue miles operated	☐	☐
	On-time performance	☐	☐
	Actual runtimes	☐	☐
	Actual vehicle speeds	☐	☐
Financial	Cost per revenue hour	☐	☐
	Cost per revenue mile	☐	☐
	Cost per trip	☐	☐
	Farebox recovery	☐	☐
	Subsidy per passenger	☐	☐

Key Resources on Planning and Performance Monitoring

- Rural iNTD Database
- "Estimating Ridership of Rural Demand-Response Transit Services for the General Public," National Center for Transit Research, 2016.
- TCRP Report 124: Guidebook for Measuring, Assessing, and Improving Performance of Demand-Response Transportation
- TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System
- TCRP Report 95: Traveler Response to Transportation System Changes Handbook, Third Edition: Chapter 9, Transit Scheduling and Frequency
- Performance Evaluation for Rural Transit Systems, National RTAP
- "Transit Price Elasticities and Cross-Elasticities," Journal of Public Transportation, Vol. 7, No. 2, 2004.

5. PARTNERSHIPS

Many small agencies do not have the resources to collect or analyze all the data they need to. Partnerships with other organizations—another transit agency, a non-profit, a state agency, a business, local community college/university, or other entity—can help. These engagements can provide access to novel datasets, help analyze data for decision-making, and share useful technology. Data can also strengthen existing relationships and support new ones. Through partnerships, agencies can capitalize on the collective knowledge of the industry and their respective regions.

Partnerships broaden an agency's capabilities to collect, analyze, and share data, creating better services and increasing efficiency.

Partnerships Providing Access to Data

Local School Districts

Schools can provide information on where students live and school bell times so agencies can plan services to meet demand. They can also provide information on important student demographics.

Metropolitan Planning Organizations (MPOs)

MPOs (and Regional Planning Agencies) can provide a host of datasets to aid in decision-making and planning, including demographic data, employment data, travel data, future projections, and infrastructure data. They can also assist with data collection.

Colleges and Universities

Colleges and universities can provide assistance with data collection as well as information on class times and locations of off-campus student apartments so agencies can plan services that adequately meet demand.

State Departments of Transportation (DOTs)

State DOTs can provide datasets on roadway infrastructure and often maintain travel surveys that can provide additional context for agencies when service planning.

Municipal Departments

Municipal Departments can provide access to local demographic and infrastructure data to help agencies with decision-making and planning.

Other Transit Agencies

Other transit agencies can provide ridership information at connection points to help agencies plan connecting services and also pool resources to purchase new data collection technologies more efficiently.

Major Employers

Employers can provide information on shift start and end times and approximate employee home locations so agencies can plan services that adequately meet demand.

These findings are from the Data Practices Guidebook. The Guidebook is a resource to assist small urban, rural, and tribal transit agencies in understanding and applying good data practices.

Partnerships for Data Analysis and Technology

Municipal Departments

Municipal departments can provide IT assistance and work with agencies to create mobile apps or web interfaces to display data such as real-time vehicle locations.

Colleges and Universities

Colleges and universities can provide assistance with developing mobile apps and other interfaces to display agency data to the public.

Other Transit Agencies

Other transit agencies can share programming language used to create mobile apps or other interfaces and can do joint procurements for new technologies.

Using Data to Build Partnerships

Ridership and Performance Data

Data analysis can help justify service partnerships with local funding partners or justify on-demand services.

Transit Demand

Transit agencies can analyze transit demand to see where partnerships with TNCs or mobility-on-demand companies to operate microtransit or on-demand services may be best suited.

Customer Satisfaction

Customer satisfaction data is important to funding partners so they can be sure their constituents are satisfied with the service they are paying a transit agency to operate.

5.1. Introduction

The success of a transit agency depends on many circumstances outside of its control. From the distribution of jobs and housing in a community to the amount of funding available from the state government, engaging with these factors is often a matter of working with data. Partnerships with other organizations—another transit agency, a non-profit, a state agency, a business, or other entity—can help. Partnerships allow agencies to capitalize on the collective knowledge and resources of the industry and their respective regions to better address the circumstances outside of an agency's direct control.

Obtaining data and drawing insights from it may require resources outside your organization. Smaller agencies in particular often have staffing constraints that limit the amount of data that can be collected and analyzed and what technology can be deployed. Small agency staff wear many hats, and with basic scheduling, operations, and administrative tasks taking up much of each day, data collection and analysis often take a backseat. Partnerships with other transit agencies or local, regional, and state planning agencies can provide access to novel datasets, help analyze data for decision-making, and share useful technology.

Data can also help strengthen existing relationships and support new funding or operations partnerships. Community stakeholders like local schools, colleges and universities, major employers, and others rely on transit, and data can help make a case for improving service to these constituencies or receiving financial support from them. In cases where it may be impractical for a transit agency to provide more service, partnerships help to complement an agency's services to improve the mobility of residents and workers. One recent trend in this area involves collaborating with Transportation Network Companies (TNCs) not only for data sharing but also for making use of new technologies.

5.2. Partnerships Providing Access to Data

Not all data is "open," and not all open data is readily used. As a result, agencies must rely on other organizations to obtain datasets vital to transportation planning. Both public and private sector entities can be potential partners, including local school districts, colleges and universities, major employers, metropolitan planning organizations (MPOs), regional planning agencies (RPAs), municipal departments, state departments of transportation (DOTs), and other transit agencies. **Table 5** outlines potential partnerships and types of data available through them.

Table 5: Summary of Partnerships that Provide Access to Data

Partner	Partner Contribution	Output or Analysis
Local School Districts	Bell times, school locations, anonymized student home locations, and student demographic information	Plan services that will adequately serve the student population and plan where targeted outreach may be necessary for low-income populations and non-English speaking populations
Major Employers	Shift times and approximate employee home locations	Plan services that will adequately serve the employment site
Colleges and Universities	Class schedules and locations of major off-campus student apartments, student help with data collection	Plan services that will adequately serve the student population, collect data efficiently
Metropolitan Planning Organizations and regional planning organizations	Provide demographic, social service, travel, employment, building permit, and infrastructure data, as well as technical assistance	Plan services that will adequately serve the community currently and in the future, adequately plan for infrastructure improvements

Partner	Partner Contribution	Output or Analysis
Municipal Departments	IT support, building permit data	Troubleshoot issues with data-producing technology, maintain websites, adequately plan for future service
State Department of Transportation	Travel data, infrastructure data, location-based services (LBS) data	Plan services that adequately meet demand and adequately plan for infrastructure improvements
Other transit agencies	Ridership data at connection points, new technologies to collect data	Plan services that provide connectivity to other regions and collect new data in an automated fashion

Local School Districts

When transit agencies provide service to local schools, understanding when students travel to school and where they travel from are key inputs to service design. Funding provided by the school district to an agency will often be predicated on service to schools before and after certain bell times (see Funding Partnerships section for details on funding arrangements and special fares). Even without funding partnerships in place, it is often in a transit agency's best interest to provide service between schools and areas with high concentrations of students, as this will increase ridership on the system.

> Local school districts can provide data on school locations, bell times, and the location of student residences so agencies can better plan their school services

To adequately serve the student population, school districts must provide agencies with a few important datasets: each school's bell times for the beginning and end of the day, school locations, and the home addresses of students, anonymized to protect privacy to the extent possible. With this important data, the transit agency can then plan services that pick up and drop off most students within walking distance of their home and serve the schools before and after bell times in the morning and afternoon, respectively.

Even if services are oriented towards schools, federal regulations require this service remain open to the general public. According to FTA, agencies may not operate "service that a reasonable person would conclude was primarily designed to accommodate students and school personnel and only incidentally to serve the non-student general public."[56] However, agencies can operate services for local school districts that are a part of their service available to the general public. Many medium and large transit agencies operate special school service on days schools are in session, with extra trips intended to serve students and limit overcrowding on other parts of the system.

In addition to data that can be used to plan service, local schools can also provide important information on student demographics, including the number of students receiving free or reduced-price lunch and the number of students with limited English proficiency (and their language spoken at home). This information can be important to transit agencies for determining where to reach low-income populations in their service area, and also to better understand where different types of language outreach may be needed.

[56] Federal Register. 2008. 73 (180): 53384-53390. https://www.govinfo.gov/content/pkg/FR-2008-09-16/pdf/E8-21601.pdf

Major Employers

Similar to local school districts, major employers often partner with transit agencies to provide transportation for their employees to and from work, with a similar funding agreement typically in place (see Funding Partnerships section for details on funding arrangements and special fares). As with arrangements with school districts, even without a funding agreement, it is often in a transit agency's best interest to provide service between major employers and areas with high concentrations of employees, as this will increase ridership on the system.

To help improve services to work sites, major employers can provide two important datasets to transit agencies: shift start and end times and employee home addresses, if need be anonymized to larger geographies such as US Census Block Groups or ZIP codes. With this data, the transit agency can plan services that operate between areas with high concentrations of employees and the employment site prior to shift start times and after shift end times.

> Major employers often partner with transit agencies to provide special services directed at their employees. Shift times and approximate employee home locations are critical datasets in this partnership

Colleges and Universities

> Colleges and Universities partner with transit agencies to provide service for students to campus. Class times and student apartment locations are crucial to this partnership

Higher education institutions, including colleges, universities, and community colleges, and transit providers are often close collaborators for a good reason: students may live in densely-populated apartment complexes, may face on-campus parking restrictions, and are in an age category that is generally more likely to ride transit. While larger colleges and universities typically provide funding for transit services or provide their own services, even without a formal partnership, it is worthwhile for a transit agency to provide services to these institutions to help generate ridership.

Colleges and universities can provide valuable information to transit agencies to help them plan services that will maximize ridership and efficiency, including class schedules, population figures for specific student housing complexes on campus (as applicable), and off-campus locations where students are likely to live. With this data, the transit agency can plan services that operate between areas with high concentrations of student apartments or dorms and the major academic, athletic, and administrative buildings during times when they may be in the highest use. Similarly, community colleges can provide class schedules and students' home addresses anonymized to larger geographies, so transit agencies can plan services that students can utilize.

Additionally, students at colleges and universities are a potential source of help for data collection. Particularly at those colleges and universities with urban planning programs, students interested in transit planning may be willing to help conduct on-board surveys, ridechecks, or even community outreach in order to fulfill internship requirements or prepare for future job opportunities.

Strong University Partnerships: Blacksburg Transit and Pullman Transit

Blacksburg Transit in Blacksburg, Virginia and Pullman Transit in Pullman, Washington have strong relationships with Virginia Tech and Washington State University, respectively. Both agencies structure their services, so they adequately transport students from large student apartment complexes to various buildings on campus. During school holidays, when demand is low, service is scaled down significantly. Both agencies work with their local universities to make annual revisions to service to ensure that routes are adequately serving the student population. Additionally, Blacksburg Transit often uses students from Virginia Tech's Urban Affairs and Planning program to work as interns, to help collect data through onboard surveys, and even to develop new mobile apps that communicate real-time bus location information to riders.

Metropolitan Planning Organizations and Regional Planning Agencies

Regional bodies like Metropolitan Planning Organizations (MPOs) and Regional Planning Agencies (RPAs) generate, collect, obtain, or maintain numerous datasets that would be useful to transit agencies in their jurisdictions. They can also provide technical assistance to transit agencies within their regions and help with data collection, service planning, APC/AVL data maintenance, and grant applications.

Some datasets that MPOs/RPAs collect that may be useful to transit agencies include demographic information from the Census American Community Survey (ACS), social services locations, employment data from the Longitudinal Employer-Household Dynamic (LEHD), demographic and employment projections from regional travel demand models, travel flows from regional travel demand models, building permit data for new developments, and GIS databases on roadway and sidewalk infrastructure.

- **Demographic information** can be useful in helping an agency identify where there are high concentrations of transit-oriented populations, such as seniors, persons with disabilities, and households without vehicles (see Chapter Three: Open Data for more information). Institutional knowledge of local demographics can also be valuable, including an understanding of what languages or dialects are commonly spoken in the region. Similarly, they may have information on community partners that deal with vulnerable populations that can provide information on where constituents live and work. While some of this data is accessible online through other sources (such as the Census American Community Survey), MPOs often compile, summarize, and analyze this data in ways that may be easier for transit agencies to use.
- **Locations of social services** can be useful in helping an agency identify what destinations should be served outside of traditional employment centers or activity centers. Agencies such as councils on aging, senior services, unemployment offices, child and family services, and job training are important destinations for many transit riders who do not have access to vehicles. MPOs often have GIS databases of these locations that they can provide to transit agencies.
- **Employment data** can be obtained from the US Census' Longitudinal Employer-Household Dynamics (LEHD) dataset or from proprietary sources such as data axle USA (formerly InfoUSA). This data not only allows an agency to see where jobs are located in their region but, in the case of LEHD data, also allows agencies to see home-to-work travel flows—a valuable dataset when planning transit routes (see Chapter Three: Open Data for more information). Like demographic information, some job data is accessible online but can be difficult to process and use effectively. MPO/RPAs may have compiled and summarized this data into easier-to-use formats and may have access to proprietary sources of job data.
- **Demographic and employment projections** developed by MPOs for travel demand modeling allow agencies to see where growth is projected in their region. This, in turn, enables them to plan services that will be viable for years to come and plan for services to implement in the future.

- **Travel flows** from MPO regional travel demand models allow agencies to see future travel patterns, which like demographic and employment projections, let agencies plan services that will be viable in future years. Travel flows are often broken out by trip purpose and travel mode as well, allowing transit agencies to gain insight as to the specific service types that would be most successful.
- **Building permit data** allow agencies to see where new major developments are planned so they can plan services for the near-term future. Though regional travel demand model projections can forecast growth over a ten-to-thirty-year horizon, they have less specificity about developments pending within shorter timeframes.
- **Roadway and sidewalk GIS** databases can provide context to inform capital improvements for transit services. For example, they may show that certain roadways have wide shoulders that could be converted to bus-only lanes or sidewalks that could provide connections to future bus stops. Though certain roadway data is available through the US Census, local jurisdictions typically maintain more detailed databases on these assets.

Technical assistance provided by MPOs and RPAs can help transit agencies with collecting and analyzing data to be used in service planning efforts. For example, MPO staff can assist with ridechecks or distributing surveys on transit vehicles to help inform a transit development plan. Technical assistance can also be provided to help agencies clean and maintain data generated by APC and AVL systems. Since federal funding for transit agencies in urbanized areas flows through MPOs, it is easy to allocate pockets of funding to MPOs for this type of assistance.

A Regional Partnership: Southeastern Massachusetts MPO (SMMPO) and Southeastern Regional Transit Authority (SRTA) Partnership

The Southeastern Regional Transit Authority (SRTA) has a partnership with the Southeastern Massachusetts MPO to provide technical assistance in a number of areas. Some of the major areas of assistance include data collection, Title VI analyses for major service changes, mapping, modeling, and demographic projections[57]. Data collection assistance typically includes on-board surveys, ridechecks, bus stop inventories, and bus stop amenity inventories. Mapping assistance typically includes creating transit system maps, individual route maps, and public-facing schedule brochures. Assistance with modeling and demographic projections usually involves providing transit market data so SRTA can conduct service planning activities. This could include prevalent travel flows in the service area and areas with high current and projected transit demand. In addition to these specific areas of assistance, as a regional entity, the SMMPO provides opportunities for coordination discussions with other regional partners and adjacent transit agencies.

Municipal Departments

Like MPOs and RPAs, departments of local governments often create and maintain important datasets on roadway and sidewalk infrastructure, employment and demographic statistics, and local building permits. Notably, when it comes to things like bus stop placement, roadway and sidewalk infrastructure data are necessary for the agency to plan for the placement of new bus stops adequately, as well as zoning and permitting data to identify upcoming trip generators. Additionally, where local transit agencies are, in fact, municipal departments themselves, other departments may be hosting the agency's website or maintaining mobile apps. This necessitates a strong partnership between the agency and these departments so that data can be updated regularly and websites and apps can be properly maintained.

[57] Southeastern Regional Planning and Economic Development District (SRPEDD). 2020. "FFY 2021 Unified Planning Work Program." https://srpedd.s3.amazonaws.com/wp-content/uploads/2020/07/13213504/FFY-2021-UPWP-Full-122220.pdf

State Departments of Transportation

Like MPOs, RPAs, and local municipal departments, state DOTs can often provide access to roadway infrastructure datasets that can help with planning for transit priority treatments or even route realignments. State DOTs also often maintain statewide or regional travel surveys that could provide an additional source of information to transit agencies exploring travel patterns in their region. The Massachusetts Department of Transportation conducted a household travel survey in 2011, which included over 15,000 responses and provided valuable information for transit agencies, including data on household travel patterns, vehicle access, age, and modes used.[58]

One of the most important datasets state DOTs can provide is location-based services (LBS) datasets. LBS datasets include anonymized cell phone location pings for any day or month of the year, allowing users of the data to see travel patterns across a given area. These datasets are typically made available to all public entities within the state for planning purposes, and state DOT staff can often provide technical assistance with querying and downloading data. The data allows for fine-grained travel analysis at smaller geographies and more specific time periods than traditional travel models and thus can be useful for planning specific trip times (see Chapter 1: Data Sources for more information).

Other Transit Agencies

Partnerships between adjacent agencies that have connection points between their services are particularly important. Adjacent agencies can provide access to their own schedules, ridership data, or other regional data that they may have collected. Schedules would be critical for the agency to provide timed connections to transfer points between the agencies, while ridership data could give a sense as to how many people may want to transfer between each agency's service. For example, Radford Transit in Radford, VA, has a route that connects to the Blacksburg Transit system. Coordination between the two agencies through regional planning efforts has led to better-timed connections at several shared stops in Blacksburg and Christiansburg.

> Other transit agencies can be valuable partners by providing ridership data on connection points

Regional datasets often contain data for areas outside of a particular agency's service area, including the service areas of adjacent agencies. By sharing all the regional datasets the agency has access to, adjacent agencies can make sure that everyone in the region has the same data. Additionally, adjacent agencies may be able to pool resources to collect data more cost-effectively by conducting simultaneous ridechecks or by buying APC/AVL systems in bulk.

5.3. Partnerships for Data Analysis and Technology

ITS devices and thorough data analysis come at the cost of both capital outlays and staff expertise. Transit agencies need not conduct all procurement or data analysis themselves: often, there are partners that can help agencies procure, develop, or maintain technology or who can process datasets that new technologies produce. In some cases, agencies are sharing their own proprietary technology with other agency partners and helping them troubleshoot issues with new technology. All of these methods of collaboration allow transit agencies to expand their technical capabilities without having to hire new staff or undertake difficult procurement processes themselves. **Table 6** shows a summary of these partnerships.

---

[58] Massachusetts Department of Transportation. 2010-2011. "Massachusetts Travel Survey." https://www.mass.gov/doc/2010-massachusetts-household-travel-survey/download

Table 6: Summary of Partnerships on Technology

Partner	Partner Contribution
Municipal Departments	IT support, app or interface development, website development, technology procurement, GIS assistance
Colleges and Universities	Development of new apps and interfaces
Other Transit Agencies	Sharing technology, help with technical specifications for new technology procurements

Municipal Departments

When transit agencies are departments of local municipalities, municipal Information Technology (IT) departments can be a valuable partner for agencies wishing to deploy new technologies or develop their own applications. These departments typically have staff familiar with website design and can provide assistance to agencies dealing with issues within these realms. Agencies with AVL systems can work with IT departments to develop interfaces that display real-time information to customers and operations staff – much like the Eastern Panhandle Transit Authority (EPTA) did with the Berkeley County, WV IT Department (see case study below).

Other departments can also be helpful partners to transit agencies. Maintenance departments can help when transit agencies are trying to procure new technologies, providing guidance as to which components would be necessary to make the proposed system work within the agency's current fleet. Municipalities also typically have GIS departments or GIS specialists that can aid transit agencies with compiling their spatial data into maps or creating components of GTFS feeds.

Municipal IT Departments can help agencies deploy new technologies or create their own

Colleges and Universities

Colleges and universities are often a great source of technical knowledge and can help with developing things like mobile apps and website interfaces. Those that have urban planning, engineering, or computer science programs often have class projects or internship requirements for students. Transit agencies can make good partners for these efforts and develop new customer information apps, real-time information apps, or other public-facing web interfaces with students and faculty through them.

Deploying Technology: Blacksburg Transit and Virginia Tech

Blacksburg Transit in Blacksburg, Virginia, and Virginia Tech University collaborated to create a mobile app for the agency that would contain important information for riders. The main goal of the app was to disseminate real-time bus location information to riders using the agency's AVL feed. Other important information was added, including passenger loads on each active bus, links to bus schedules, and a trip planner. The passenger load information was particularly important to riders, as buses can become overcrowded during peak times right before classes begin in the morning at Virginia Tech. The original app was called BT4U, but with subsequent updates, it was changed to just the "BT App."

Other Transit Agencies

> Other transit agencies can share technology and collaborate in troubleshooting issues with new technologies

Partnerships between transit agencies can also be beneficial to the development and deployment of new technologies such as APC, AVL, EPC, or AFC systems, mobile ticketing apps, or mobile customer information apps.

Funding partnerships may also be formed with other transit agencies through joint procurements of new technology. In this scenario, two or more transit agencies would pool funding in order to get bulk discounts from vendors when purchasing new technologies such as AVL or EPC systems. Then when these devices start generating data, the agencies can work with each other to troubleshoot data processing issues or even create shared data templates for analysis.

When developing a new real-time information app, agencies using the same AVL vendors can share the backends used to create the app so another agency can customize it for their own system. The sharing of this kind of information is typically beneficial to all agencies involved, as deploying them to additional agencies often uncovers minor errors or expands their capabilities. When procuring new technologies, transit agencies can call on other transit agencies that have deployed the technology before to ensure that it will work well with their existing systems and even help develop technical specifications.

AGENCY PROFILE: EASTERN PANHANDLE TRANSIT AUTHORITY AND BERKELEY COUNTY, WV "WHERE'S MY BUS" APPLICATION

THE EASTERN PANHANDLE TRANSIT AUTHORITY (EPTA) AND THE BERKELEY COUNTY, WV IT DEPARTMENT COLLABORATED ON A REAL-TIME BUS LOCATION PLATFORM CALLED "WHERE'S MY BUS."

Context

Over the past several years, EPTA has expanded its services significantly while upgrading its fleet, branding, marketing, and customer service. EPTA operates deviated fixed-route service in Berkeley County and Jefferson County, West Virginia, with its main hub in downtown Martinsburg. Approximately 80 miles west of Washington, DC, Berkeley County has seen rapid population and employment growth over the past ten years. In reaction to this, EPTA expanded from six routes in 2014 to 13 in 2020.

Context (cont.)

After EPTA installed Zonar Systems AVL technology on its vehicles, they sought a platform to easily share vehicle locations with their dispatching staff and the general public. After conversations with Bluefield Transit, another small transit agency in West Virginia, EPTA learned how they could create a real-time bus location web interface with their AVL data. EPTA enlisted the Berkeley County Information Technology (IT) department's help to develop a similar interface. Using the Zonar API, Berkeley County staff created a web page that tracks each bus in operation on a Google base map with an automatic refresh. The site can be viewed easily on a mobile phone in addition to EPTA's website

Figure 30: An EPTA Van in Service

Resources Needed

EPTA staff provided the Berkeley County IT with the API received from Zonar, and then two Berkeley County IT staff members were able to create the website in about six weeks for a total cost of only $2,500. After EPTA staff reviewed the initial version, EPTA launched it to the public in April 2018. When Zonar made updates to their API, EPTA paid Berkeley County to update the website, which only took a few hours. Overall, EPTA noted that using Berkeley County staff saved them considerable money over using a private vendor. Since the webpage was created in the [PHP](#) programming language, which is free and open-source, there were no additional costs to create and maintain the website outside EPTA's annual contract with Zonar and Berkeley County staff hours.

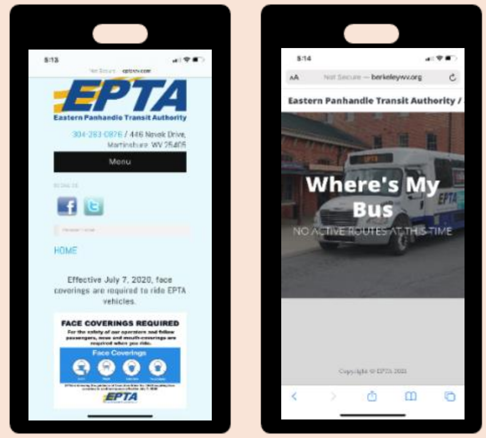
Figure 31: Where's My Bus Webpage

EPTA and Berkeley County did not note any major barriers to creating the webpage outside of having a public partner like Berkeley County with staff skilled in PHP, the use of APIs, and website building. The webpage was set up under the Berkeley County website domain to make updates to it easier. Zonar Systems set up their hardware on EPTA's vehicles as part of a state procurement that included installation costs and the hardware itself.

Results

The webpage created, known as "Where's My Bus," has reduced the number of customer service calls into EPTA's office and has been well received by customers. The webpage also allows EPTA administrative and operations staff to track bus locations, which aids in troubleshooting potential transfer issues with customers who call in for assistance.

Figure 32: Where's My Bus Mobile App



Very little maintenance of the webpage is required, and when needed, Berkeley County IT staff has taken the lead to provide any updates necessary. As was previously mentioned, Zonar changed their API in 2020, and Berkeley County staff were able to update the webpage in only a few hours. Switching to a new AVL vendor would require a more significant update, however.

EPTA also uses in-vehicle cameras provided by Zonar Systems. In the future, EPTA and Berkeley County hope to partner again to bring live camera feeds from each vehicle into "Where's My Bus" so customers and EPTA staff can monitor passenger loads in real-time. This will allow customers to make trip planning decisions on busier routes that at times can get overcrowded and allow EPTA staff to verify overcrowding issues reported by drivers and customers

Lessons Learned

There were several lessons learned when implementing the "Where's my Bus" webpage, including:

- Different AVL vendors have significant differences in the format of their data and their APIs. Initially, EPTA thought they might be able to deploy the Bluefield Transit code with minimal effort, but Berkeley County staff had to rework much of it to get it to work with the Zonar formatting.
- Using a public partner like a local government's IT department meant significant savings over using a private vendor.

Key Takeaways

- With inspiration from another small agency in the same state, EPTA was able to envision a real-time vehicle tracking application it could easily deploy to its customers without significant procurement efforts.
- The partnership between EPTA and Berkeley County led to the successful deployment of the real-time vehicle tracking application called "Where's my Bus."
- Using county staff, EPTA was able to save costs over hiring a private vendor to deploy a similar application.

For More Information

Eastern Panhandle Transit Authority (EPTA)

Where's my Bus Webpage

PHP programming language

5.4. Using Data to Build Partnerships

Even though transit agencies are relied upon for transportation services, it can be difficult for transit agencies to prove their importance to the agencies that fund or oversee them. When building a new relationship with a partner organization, providing proof of the effectiveness of an agency can be even more difficult. Certain types of data analyses can help illustrate a transit agency's importance to a community and make a case for more funding support from current and existing partners. These analyses include ridership and performance data analysis, transit demand analysis, and customer satisfaction analysis. **Table 7** summarizes the ways to use data to strengthen partnerships.

Table 7: Summary of Ways to Use Data to Strengthen Partnerships

Partnership	Type of Data	Key Result
Major employers	Ridership and performance data, customer satisfaction data	Ensure that services intended for their employees are performing well, continue partnership
Local school districts, colleges, and universities	Ridership and performance data, customer satisfaction data	Ensure that services tailored to their students are performing well, continue partnership
Local funding partners	Ridership and performance data, customer satisfaction data	Ensure that services are efficient and passengers are satisfied with the service
TNCs or mobility on-demand companies	Transit demand analysis, ridership on the partnered service	Ensure that service provided has strong ridership and partnership is viable

Ridership and Performance Data Analyses

Ridership and other performance data can provide key information to justify partnerships with existing or potential partners, particularly those that provide funding or grants to agencies.

- **Major employers that are partnered with transit agencies** typically want to see how the service they are contributing to is performing in terms of ridership and on-time performance, and also whether it is well-timed to serve their employees. Providing these employers with summaries of this information helps keep them informed and actively aware of the benefits of the service. Where partnerships include free fares for employees, tracking them allows the agency to ensure that the funding received from the employer is sufficient. Major employers that could be potential partners often will want to see data on how many potential employees could use the service proposed along with data on the agency's reliability performance before committing to a partnership.
- **Other major employers such as hospitals, retail establishments, restaurants, and hotels** can become better partners when data on how connecting transit services are performing. These employers want to know that both their employees and customers are able to reach their sites efficiently and easily, and can often become funding partners if they feel the service is directly benefiting them.
- **Local school districts, colleges, and universities** that are partnered with transit agencies also typically want to see how their partnered services are performing, so ridership data and on-time performance data are important to justify and continue partnerships.
- **Other local funding partners, such as municipalities, counties, or special districts** that provide funding matches to transit agencies typically want to see that services are operated efficiently, and passengers are satisfied with the service they are getting. Ridership and performance data can demonstrate that these factors are true; for example, high ridership per revenue hour figures can prove that the agency is operating efficiently.

Transit Demand Analyses

Transit demand analyses demonstrate where transit service would be most successful. Demographic characteristics and travel patterns often show that certain types of transit service would be more viable than others. In particular, areas with high concentrations of transit-oriented populations but lower population and employment densities may be better suited for demand-response or microtransit service instead of fixed-route service. In this scenario, the agency could enter into an operations partnership with a Transportation Network Company (TNC) or a mobility on-demand (MOD) company that provides this type of service and can operate it on behalf of the agency in a more efficient manner.

Partnerships to operate service can increase efficiencies and improve mobility in a community

In an operations partnership, an agency typically pays a TNC or mobility on-demand company a portion of its operating budget to operate daily service in certain portions of their service area. These companies then provide a trip request and routing platform that is used to request trips directly from customers and route vehicles efficiently to their desired pick-up and drop-off locations. Agencies can either use their own vehicles and drivers, or the company they are partnering with can provide vehicles and drivers. Key to this partnership is a data exchange; the company operating the service must know where demand exists, and the transit agency must know what trips are being made and what fares are being collected to ensure proper reporting. Good transit demand data will ensure that the service areas defined for these companies will yield strong ridership, and strong ridership will ensure that the partnership will be viable and long-lasting.

Baldwin County, AL Microtransit Service Partnership

Baldwin County, Alabama, recently converted its entire network of fixed-route services (known as BRATS) to microtransit service in a partnership with the mobility on-demand company Via.[59] Customers request trips using a mobile app, and then Via's routing platform develops dynamic routing that allows riders to go anywhere in the service zone on-demand. Passengers usually need to walk a block or two to board service. Ridership data, reported through trip logs, are important to maintain the viability of the partnership. In addition to being reported to NTD, spatial ridership data can help the agency make decisions on service zone expansions or reductions or portions of the service area that may need additional vehicles to meet demand.

Customer Satisfaction Analysis

Funding partners of transit agencies providing service to their constituents want to know that they are satisfied with the service they are paying for; this makes customer satisfaction data important to share with funding partners. Whether they are major employers, local school districts, colleges, universities, or local governments, partners need to know that their investment is valued by their students, employees, or residents. Customer satisfaction data is typically collected by agencies through on-board surveys (some with online counterparts), mobile apps, website forms, or even specialized programs where passengers can provide feedback. Virtual public engagement can also be an important tool for gathering feedback on customer satisfaction. For more information on virtual public engagement, see the N-CATT Guidebook for Virtual Public Engagement.

Agencies that have mobile apps often allow riders to send feedback through the app. Pullman Transit, for example, collects customer satisfaction data this way. Larger agencies like the Maryland Transit Administration in Baltimore, Maryland, deploy a specialized application called "Rate My Ride" that allows

[59] Baldwin County. n.d. BRATS Transit. https://baldwincountyal.gov/departments/BRATS

riders to fill out a short survey on their website. What is most important is that the agency analyzes this data regularly and works to solve areas of their service that are poorly rated by riders. This will help demonstrate that the agency is committed to customer service and ensure that there is a desire to continue the partnership.

5.5. Conclusion

Partnerships with outside agencies can help small transit agencies provide better service, operate more efficiently, implement new technologies, and acquire new data. Partnering with entities such as local schools, colleges and universities, and major employers can unlock new potential ridership and revenue sources for an agency. Partnering with state and regional agencies can provide new data sources that are critical to an agency understanding its customer base and the characteristics of its service area, allowing them to operate more efficiently and better serve their community.

Partnerships with municipal departments and colleges and universities allow transit agencies to develop or acquire and maintain new technologies, including real-time information apps or websites, and mobile fare payment apps. They also provide resources to ensure that technology procurements will work with the agency's existing systems.

Ridership, performance, and customer satisfaction data are critical to proving the importance of the transit agency's service to current and potential partners. When partners are providing funding to a transit agency, they need to know that their funding is being used by the agency efficiently and that their constituents are satisfied with the service they are receiving. Data can also be used to guide the development of partnerships with companies that operate services on behalf of an agency, such as TNCs or mobility on-demand companies.

Checklist: What types of partnerships does your agency already have, and what types would you like to explore further?

Partnerships With…	Partnerships Intended To…	Currently Use This Type of Partnership	Want to Use This Type of Partnership
Local School Districts	...share data to operate school service	☐	☐
Major Employers	...share data to operate service to major employment sites	☐	☐
Colleges and Universities	...share data to operate service for students	☐	☐
	...help develop technology	☐	☐
MPOs/RPAs	...acquire new data	☐	☐
Municipal Departments	...acquire new data	☐	☐
	...help develop and maintain technology	☐	☐
State DOTs	...acquire new data	☐	☐
Other Transit Agencies	...share data	☐	☐
	…pool resources to purchase technology	☐	☐
	…share technology	☐	☐
TNCs and Mobility On-Demand Companies	…share data to operate on-demand or microtransit service	☐	☐

6. ACCESS AND EQUITY

Transit riders are diverse and have diverse needs. Different groups—minorities, low-income, seniors, students, or others—may ride different routes, travel for different reasons or at different times of the day, or pay for transit differently. Conducting equity analysis or accessibility analysis can help agencies understand the varied needs and travel patterns of riders in their service area. Equity analysis ensures that service changes provide the same service levels to minority and low-income riders and that fare changes do not harm any group. Accessibility analysis examines the number of destinations, often jobs, that are accessible via transit within a certain time frame, which allows an agency a better understanding of how riders use their services. Data allows this analysis to happen, and new tools and data sources provide richer answers to these questions.

> Focusing on access and equity can help agencies understand how different riders use transit services and how to meet their needs better.

Access and Equity

Service Equity Analysis

When service changes are proposed, conducting a service equity analysis ensures that minority and low-income riders receive the same service as other riders.

Customer Profiles

Different riders may use transit differently. Understanding customer profiles helps planners understand how different groups ride, including the routes they use or the fare media they rely on.

Accessibility Analysis

Accessibility analyses determine which destinations or jobs a person can access within a certain time frame on public transit, which helps planners understand how the service can be used by a potential rider.

Fare Equity Analysis

Different groups may also pay fares differently. Fare equity analysis examines the impact that fare changes have on different groups to ensure that minority or low-income riders are not harmed.

Trip Purpose Analysis

Trip purposes extend beyond commuting. Collecting trip purpose information from riders can help agencies plan routes to meet riders' needs.

Trip Planning Accessibility

Offering access to trip planners helps riders plan their trip using transit. Providing a GTFS feed to trip planners or including a trip planner on an agency website helps riders easily understand the service and plan their trip.

6.1. Introduction

Transit agencies seek to provide transit services across their service area, but the people living, working, and traveling in the service area are not uniform. Different groups of people may take different trips, use different routes, or pay for transit differently, and transit service may serve some trips or needs better than others. Access in transit typically examines the number of destinations, often jobs, that riders can access within a fixed amount of time. Equity in transit is the value that degree of need should be considered when determining how resources are allocated. Data provides a means to assess who transit serves and how well it serves them. New tools and data sources provide the means to answer questions of access and equity with richer detail and actionable insights.

Evaluating accessibility to key destinations and jobs is an emerging trend in the industry. Transit agencies have begun to pay more attention to accessibility metrics, which examine how well transit provides connections to key destinations. Although equity analysis is required through Title VI, there are opportunities for agencies to more thoroughly address these questions. Equity reports that go beyond federal requirements are being conducted in places like San Francisco and may soon be repeated in other jurisdictions. Access is one lens through which equity can be evaluated. In addition, asking the right questions is the most important part of an equity analysis. Consider the many dimensions of equity as you frame your research questions. Equity considerations include service offered, fare payment types, trip purpose analysis, and trip planning accessibility. If there is a riders' advocacy group in your service area, they may be helpful in highlighting the right questions to ask.

6.2. Practices to Increase Access and Equity with Data

Transit agency data allows agencies to examine access and equity questions. This includes service equity analysis and fare equity analysis, understanding customer profiles and the reason they take trips, accessibility analysis, and making trip planning accessible to riders. **Table 8** shows the topics covered in this chapter, the questions that each asks, the data required for each, and the method.

Table 8: Access and Equity Analyses

Category	Input Data	Method	Analysis
Service Equity Analysis	Schedule data, Census data	Title VI-style analysis of demographics at stop locations with proposed service changes	Will service changes disproportionately affect minority or low-income customers?
Accessibility Analysis	GTFS schedule data, OpenStreetMap data	Generate travel time isochrones using GIS software	What areas can you reach on public transit in a limited amount of time?
Fare Equity Analysis	Fare payment data	Estimate the amount of fares paid by demographic group under current and proposed fares	Will fare changes result in a disparate impact on protected populations?
Customer Profiles	Customer survey data	Create crosstabs of survey data by demographic characteristics at the route and systemwide level	Are certain routes critical to high-need groups? Are certain types of fare media (senior discounts, student fare cards) used effectively?
Trip Purpose Analysis	Demand Response customer demographic and trip information data	Create crosstabs of survey data and make inferences on trip purpose from other available data	Are there needs to increase services for particular demographic groups or trip types (e.g., Medicaid trips)?

Category	Input Data	Method	Analysis
Trip Planning Accessibility	GTFS schedule data	Upload feed to trip planning providers or host trip planning service on the agency website	More accessible trip planning for customers.

6.2.1. Service Equity Analysis

Service areas are not homogenous, and neither are the populations within them. The Federal Transit Administration (FTA) requires transit agencies that provide fixed-route service, operate 50 or more fixed route vehicles during peak service, and operate in an urbanized area with a population of at least 200,000 analyze the effects of service changes specifically for minority and low-income populations, as stipulated in FTA's Title VI Circular.[60] The purpose of a Title VI analysis is to determine if minority or low-income populations bear more of the burden of service reductions or if they receive fewer of the benefits of service additions. Even if an agency is not required to evaluate service changes, analyzing service offerings and changes and how this affects minority and low-income populations can help an agency understand if its services are equitable.

If an agency is required to complete a Title VI Program by FTA, the Title VI Program must include the following definitions:

- **Major Service Change Policy:** This defines which service changes warrant a Title VI analysis. This may include a percentage of a route that changes, a system-wide change, and the number of people affected by the change. This applies to changes that add or reduce service.
- **Adverse Effects Definition:** This explains the type and magnitude of changes that represent adverse effects. This can include the elimination of a route or stop, shortlining a route, rerouting, headway changes, or span changes.
- **Disparate Impact (DI) Policy:** This is the threshold at which an impact is considered substantially different from the service area average for minority populations. For example, if the percentage of minorities in a service area is 30 percent, but service is removed from an area with 50 percent minority populations, this likely exceeds the threshold for a disparate impact.
- **Disproportionate Burden (DB) Policy:** This is the threshold at which an impact is considered substantially different from the service area average for low-income populations. For example, if the percent of low-income households in a service area is 30 percent, but service is removed from an area with 50 percent low-income households, this likely exceeds the threshold for a disproportionate burden.

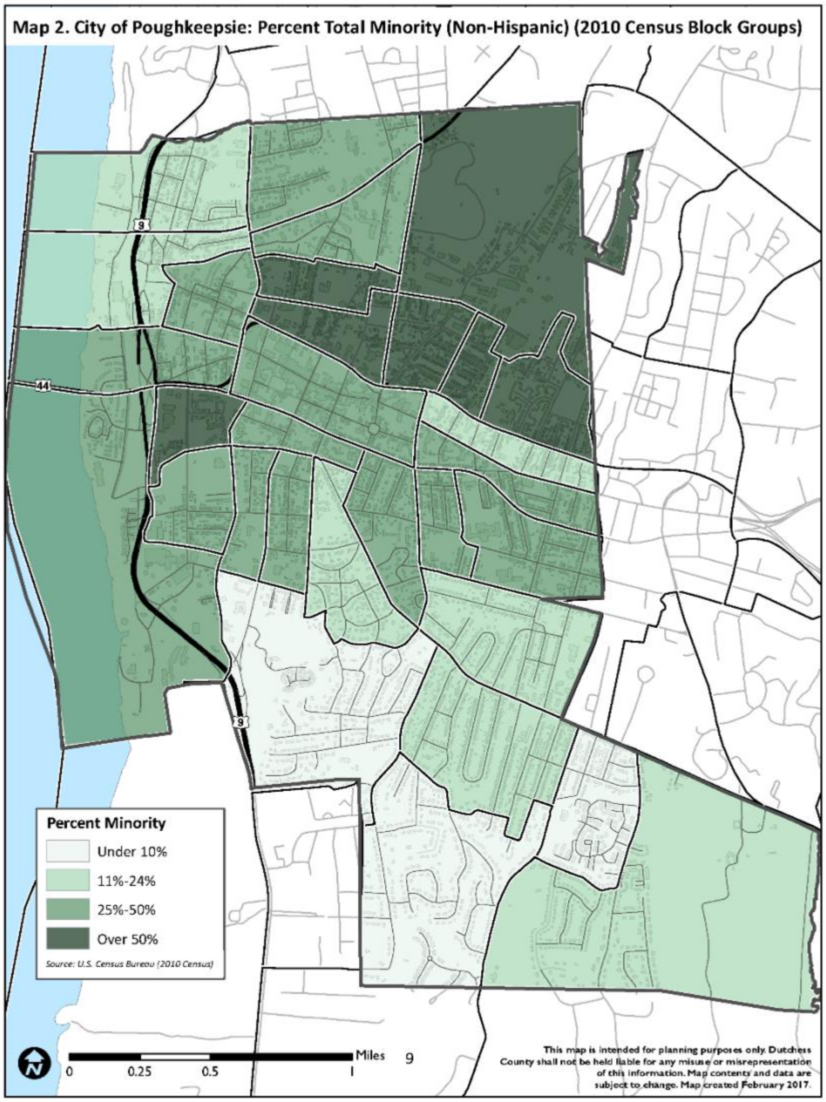
To complete a service equity analysis (SEA), an agency needs to collect census data about populations in the service area, including minority and low-income populations. Each agency can determine the threshold for low-income households based on the federal poverty level, median income in the service area, or other criteria chosen by the agency. Minority populations consist of people who identify as:

- American Indian and Alaskan Native
- Asian
- Black or African American
- Hispanic or Latino, regardless of race
- Native Hawaiian or Pacific Islander.

[60] Federal Transit Administration. 2012. Circular. transit.dot.gov/sites/fta.dot.gov/files/docs/FTA_Title_VI_FINAL.pdf

Supporting maps illustrating where these populations live in a service area can be useful for illustrating where changes to service could imposed a disproportionate burden or a disparate impact, as shown in **Figure 33**. Whenever possible, the color breaks used to illustrate these figures should be changed to include the DB and DI thresholds appropriate for the service area.

Figure 33. Example Map of Minority (Non-Hispanic) Populations for Service Equity Analysis[61]



The agency also needs to collect information about service levels before and after a proposed service change. This information includes route headways, span of service, number of trips, and stop locations. The analysis can be completed by route, if only a few routes are changing, or at the system level if more than a few routes are changing.

---

[61] Dutchess County Department of Planning and Development. "Title VI Service Equity Analysis: Dutchess County Public Transit Service Expansion in the City of Poughkeepsie". March 1, 2017. https://www.dutchessny.gov/Departments/Public-Works/Docs/service-expansion-analysis.pdf. Accessed March 2, 2021.

To do this, data from a GTFS feed for existing services can be compared to a GTFS feed that includes the proposed service. Though other data sources can be used for this purpose, using GTFS feeds allows for more granular analysis of service changes and a more repeatable analysis. For example, this analysis can examine service levels at the stop level, taking together the effect on non-changing routes and multiple routes that change simultaneously to understand the resulting service at the stop level.

Follow these steps to complete an SEA:

1. **Determine if the change in the route meets your major service change in policy.**
2. **Prepare a GTFS feed that includes the proposed service.** This may require coordination with IT or other staff who build GTFS feeds or use of a tool like the RTAP GTFS Builder (see Chapter Two: Data Sources).
3. **Quantify the changes that are happening at the route or stop level.** How much span is it gaining or losing? By how much are headways becoming more or less frequent? How many stops along the route are being added or removed? These can be quantified either by the raw change (number of minutes or hours) or by the percent change, according to the definition of adverse effect in your Title VI program. Use the existing and proposed GTFS feeds to quantify span and headway. Then determine how much span and headway are changing at each stop.
4. **Evaluate whether the changes meet your adverse effects definition.** For example, a span change of 15 minutes might not meet the definition of adverse effect, while a span change of two hours probably will.
5. **Evaluate if minority or low-income populations are disproportionately bearing the burden where adverse effects occur using your DI and DB policies.** Use a GIS program to select the block groups surrounding the changing routes, typically within ¼ mile or ½ mile of the route. Compare the percent minority and percent low-income populations of the affected block groups with the service area's average. For service reductions, are the percent minority or low-income *above* your defined threshold (for example, 10 percent above the service area median)? For service additions, are the percent minority or low-income *below* the defined threshold (for example, 10 percent less than the service area median)? If so, you may have a DI (for minority populations) or a DB (for low-income populations).

If impacts are identified, this is called a disparate impact (DI) when minorities are affected and a disproportionate burden (DB) when low-income communities are affected. If a DI is found, transit agencies must either enact an alternative plan that minimizes the effect on minority populations or provide a substantial legitimate justification for the chosen service plan. If a DB is found, the transit agency must evaluate alternative policies and mitigate the burden where possible.

The FTA requires service equity analyses to ensure that minority and low-income populations receive equitable transit services. Conducting a service equity analysis allows an agency to understand the equity implications of service changes. Using GTFS feeds for existing and proposed services allows agencies to examine service changes at a more granular level, providing a more thorough understanding of the impacts of service changes.

6.2.2. Accessibility Analysis

While a basic analysis of a transit system might include the number of people a route serves or even the number of jobs accessible along a route, an accessibility analysis can deepen the rigor of the analysis. Analyses of these kind estimate the time a person must spend on the transit system in order to access jobs or destinations. An estimate of accessibility from a single location, or an isochrone, is the most basic form of this analysis. An isochrone is a summary of the travel paths that a person could take from or to a point under a certain time limit, such as 60 minutes. The result is a boundary shape that shows the total area that a person can access given the existing transit service in a specific amount of time. When combined with information about destinations like jobs, the amount of opportunities accessible to a person from a given point can be quantified.

Accessibility analysis provides a more realistic perspective on how riders will actually use the system. In practice, many riders will find that travel time limits their ability to use transit, especially if they also have the option of using a personal vehicle or another mode of transportation to complete their trip. Compared to simply examining route coverage, isochrone analysis can help synthesize the accessibility generated from route transfers.

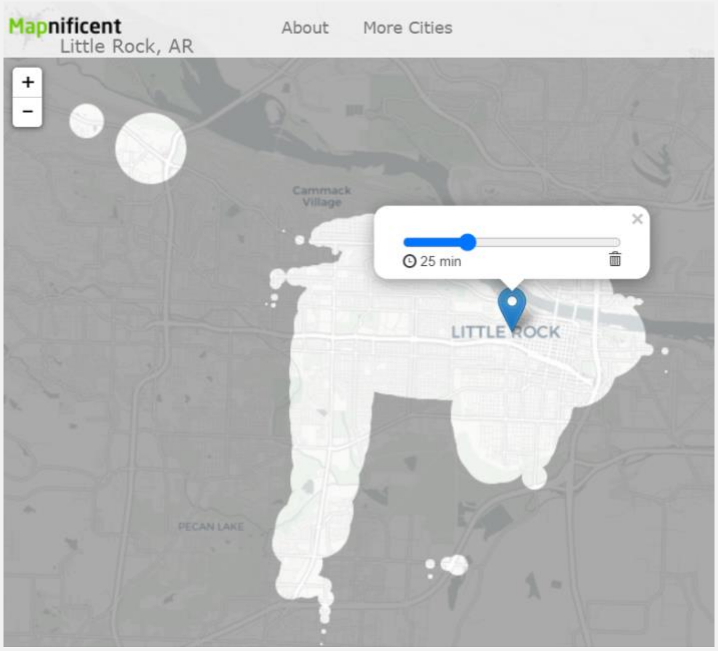
To conduct this analysis, a transit agency needs to collect GTFS feed information about the transit service, the number of jobs available throughout the service area, and GIS software to conduct the analysis and an underlying network. There are a number of options, using both free (QGIS) and proprietary (ArcGIS) software, to perform isochrone analyses, including:

- TravelTime plugin for QGIS
- Hqgis for QGIS
- Network Analyst in ArcGIS by ESRI
- Generate Service Area tool in ArcGIS Pro by ESRI.

Easy to Use Isochrones: Mapnificent

Mapnificent.net shows areas accessible by transit in a number of cities throughout the United States. A user can move the starting point to a variety of areas and change the total travel time. The resulting polygon indicates areas that are reachable by transit from the starting point in the given time.

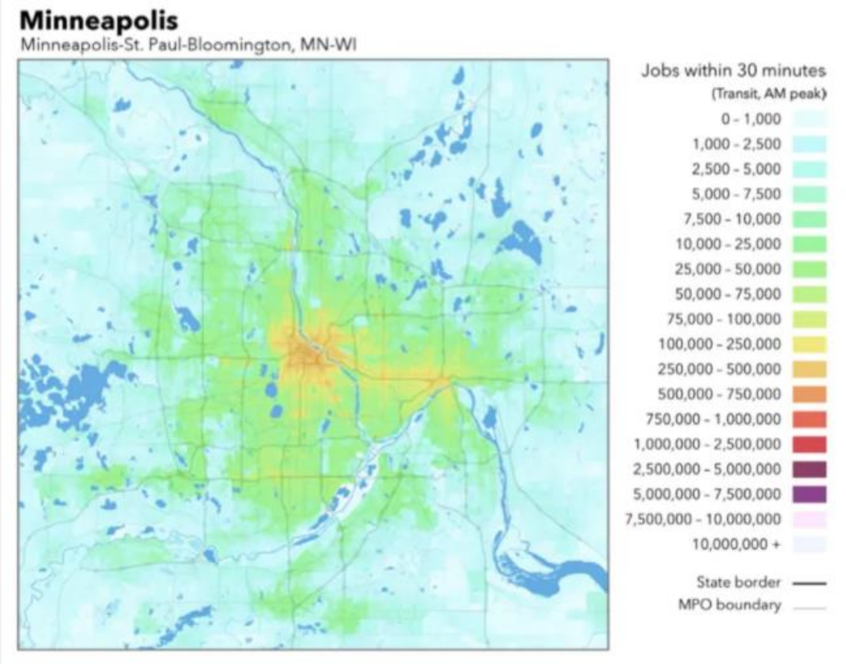
Figure 34: Mapnificent in Little Rock, AR



Accessibility Statistics: Access Across America

The University of Minnesota produced a set of maps showing job accessibility by transit for a number of metropolitan areas.[62] The maps show the total number of jobs that can be accessed within 30 minutes of travel time during the AM peak.

Figure 35: Minneapolis's Job Accessibility Map



6.2.3. Fare Equity Analysis

Different groups of riders use transit and pay for transit differently. Agencies can use data from smart fare cards to understand these different groups' needs, and this can also help agencies understand how a fare change may have different impacts on different users. To better understand the equity impacts of a fare change, agencies can conduct surveys or other analyses to evaluate the impact that a fare change may have on low-income or minority riders.

Low-income riders are one example of a group that is likely to use transit and pay for transit differently than other riders. Low-income riders are more likely to:

- Travel shorter distances
- Take more frequent transit trips[63]
- Make more transfers[64]

---

[62] University of Minnesota. 2017. "Access Across America: Transit 2017." *Accessibility Observatory*. https://access.umn.edu/research/america/transit/2017/.
[63] Cornelius Nuworsoo, Aaron Golub,and Elizabeth Deakin. 2009. "Analyzing Equity Impacts of Transit Fare Changes: Case Study of Alameda-Contra Costa Transit, California."*Evaluation and Program Planning* 32 (4), 360-368. https://pubmed.ncbi.nlm.nih.gov/19616299/
[64] Jeffrey Rosenblum. 2019. "How Low-Income Transit Riders in Boston Respond to Discounted Fares: A Randomized Controlled Evaluation." *Massachusetts Institute of Technology, Department of Urban Studies and Planning.* http://equitytransit.mit.edu/wp-content/uploads/2019/06/whitepaper_v8.pdf#page=5

- Pay for each ride separately rather than use a multi-day or monthly unlimited pass[65]
- Be unbanked or underbanked and rely on cash transactions.[66]

This means that raising flat fares, cash fares, and transfer fees will disproportionately harm low-income riders. Low-income riders, especially those that are unbanked or have low savings, may have trouble paying the up-front cost of a pass, even if they would benefit from using the pass rather than paying per ride. Fare capping, which is a pay-as-you-go option that automatically issues a pass after paying the equivalent cost through single ride fares, has the potential to address this need but requires technology to enable this on smart fare cards or through mobile apps. An agency that uses smart fare cards can analyze the need for fare capping by examining data they collect with smart fare cards. Are there riders who pay more per month in single fare rides than the cost of a monthly pass? If so, fare capping may be a strategy to address this.

Fare Capping on The Rapid

The Rapid in Grand Rapids, Michigan, offers daily, weekly, and monthly fare capping.[67] After riding the bus twice in one day, a rider using a Wave card will receive a daily pass and will not be charged for additional trips on that day. The same is true for weekly and monthly passes. This benefits low-income riders who may not be able to pay for the upfront cost of a pass but also benefits riders who do not know how often they will ride.

Corpus Christi Fare Change Analysis

In an analysis of proposed fare changes, the Corpus Christi Regional Transportation Authority in Nueces County, Texas conducted a survey to better understand how minority and low-income riders would be affected. They collected data about the trips that riders in each group took and then calculated the price each group would pay under the existing fare structure as well as under the proposed fare structure. They found that raising off-peak fares and charging for transfers disproportionately harmed low-income populations, although it did not disproportionately harm minority populations.

6.2.4. Customer Profiles

While agencies may have general ideas about who rides their services, it is helpful to collect more specific data about ridership: who is riding and for what purposes. One way to do this is by conducting rider surveys periodically to gather demographic and trip data. Important demographic information to collect includes age, household income, race, language spoken at home, ability to speak English, and employment status (including retired and student). Travel information includes the type of fare paid (e.g., regular fare or discount), fare media used (e.g., smart fare card, mobile app, cash), route used, number of transfers taken, trip purpose, and trip frequency.

Customer profiles allow agencies to extract basic information about how their service is used so they can make changes that make customers' travel easier. For example, are there specific improvements that

[65] Rosenblum. "How Low-Income Transit Riders in Boston Respond to Discounted Fares: A Randomized Controlled Evaluation."
[66] Federal Deposit Insurance Corporation. 2020. "How America Banks: Household Use of Banking and Financial Services." https://www.fdic.gov/analysis/household-survey/2019execsum.pdf
[67] The Rapid. n.d. "The Wave Benefits." https://www.ridetherapid.org/the-wave#benefits

many riders would like to see in the service? Are there common transfer points between routes that might warrant combining routes?

It is possible to create crosstabs to gain valuable information about how specific riders use transit services. For example, are students and seniors using fare discounts that they are eligible for? Are low-income riders more likely to pay in cash or transfer more often? Are there particular routes that are more heavily used by speakers of languages other than English? Depending on the answers to these questions, an agency might determine to post notices in other languages on certain routes, advertise fare discounts more widely, or avoid raising transfer fare prices.

6.2.5. Trip Purpose Analysis

Transit riders use transit for many different purposes. While riders on larger systems primarily use transit to access a job, many riders on smaller systems are instead riding to access services, such as retail, medical, educational, or social services. Understanding riders' reasons for travel can help an agency better meet those needs. There are several ways to obtain information on customers' trip purposes or infer trip purpose:

- On-board surveys or surveys conducted at busy transit stops
- Type of fare used
- Type of service used
- Pick-up or drop-off location
- Customer demographic information.

Surveys

Surveys conducted on-board transit vehicles or at major transit stops can ask questions about riders' trip purpose. This data can then be cross-tabbed with other information gathered in the survey, including the riders' age, employment status, disability status, or other demographic information. This cross-tabulation can help transit planners better understand their customers' needs so services can be tailored better to their needs.

Figure 36: Example Survey Language on Trip Purpose

7. What is the purpose of your trip today?

☐ School	☐ Social/Recreational
☐ Work	☐ Other Errands
☐ Shopping	☐ Government Service (Social Security, etc)
☐ Medical/Dental	☐ Other: ____________

Fare Type

Most agencies have multiple fare types that can offer clues as to customers' trip purposes. Specific fare types may only be eligible for certain trip purposes or may only be available to certain populations who typically have only a few common trip purposes. Fare options may include:

- Half-fares: most agencies offer half fares to seniors and persons with disabilities. Seniors are typically retired and not attending school, so it can be assumed that they are riding to access retail, medical, or social services.
- Medicaid: these trip purposes are required to be medical by law.
- Fare Zones: many agencies have fare zones where passengers pay higher fares for longer travel within the system. Where fare zones contain only one major destination (e.g.., a rural hospital), it can be assumed that trips are medical or, if during peak shift change times, trips to and from work.
- School Passes: school pass passengers can generally only use their passes to ride to and from school.
- Employer Passes: employer passes used during peak periods or shift change times can be assumed to be to and from work.

Type of Service Used

There are numerous ways to infer customers' trip purposes based on the type of service they are riding.

- Peak period commuter service: riders using more expensive express services are more likely to commute for work.
- School tripper service: riders using trips designed for school service are likely going to and from home and school.
- Demand-response Medicare or Medicaid service: riders using these services are likely traveling to and from medical appointments.
- Evening or weekend service to specific destinations: riders using these services are likely essential or service workers, particularly on smaller systems. Riders could also be attending events or accessing retail or entertainment.

Pick-Up or Drop-Off Location

Trip purpose is inherently tied to a customer's destination. On demand-response trips, drop-off and pick-up locations are the easiest way to infer information on trip purpose if they are locations that provide specific services. Similarly, trip purposes can be inferred on fixed-route services if destinations at stops are fairly uniform. For example, if a stop serves a shopping center, it can be assumed that most customers' trip purpose is retail, although, for some customers, it could be their place of employment.

Customer Demographic Information

Customer demographic information can offer clues to their trip purpose. For example, most seniors are retired and use transit instead to access medical care, senior services, or retail. Middle and high school-aged youths typically use transit to get to and from school. Using survey information, agencies can create simple correlations between demographic information and trip purpose to build simple spreadsheet models that can help define trip purpose based on this information.

Providing Access for Different Trip Purposes

Different trip purposes require different types of service, and it is important for transit agencies to understand the different types of locations customers in their service area are trying to access.

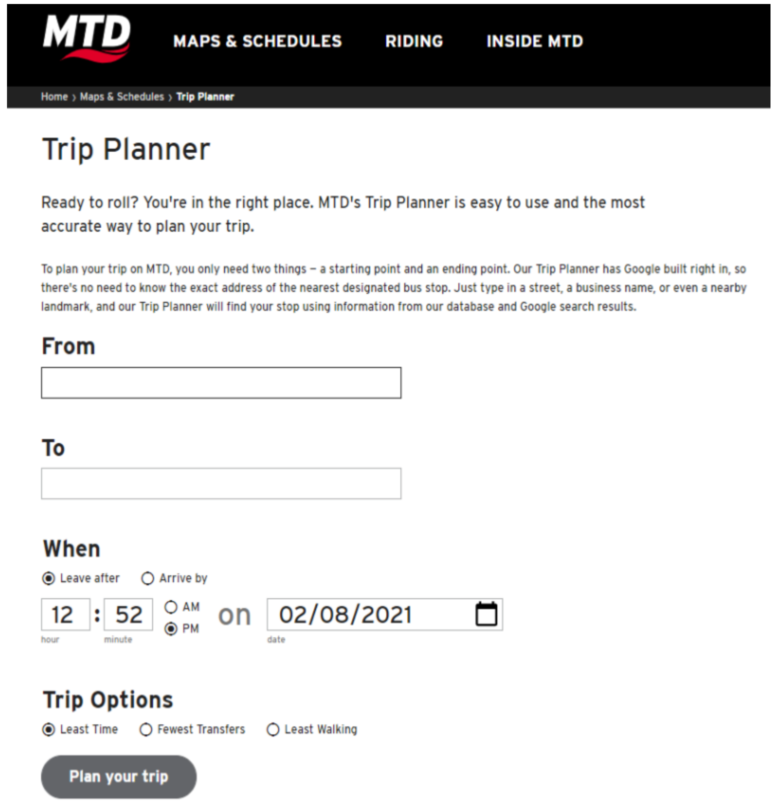
- Peak period commuter service is often needed to provide access to jobs on the traditional "9 to 5" schedule. These types of services are typically oriented in a single direction during each peak period.
- Commuter trips for second and third-shift essential and service workers typically take place during the PM peak, evening, or late-night periods and also on weekends; therefore, peak period only service is not adequate for these workers.
- Medical and social service trips typically require service between 8 a.m. and 5 p.m. on weekdays when these services are open for business; therefore, peak period only service is not adequate for these purposes.
- Shopping trips typically take place during off-peak times, PM peak times, or on weekends and therefore require service during these times.

There are a number of other important destinations for vulnerable populations, particularly those with limited English proficiency and low incomes. These destinations are often governmental buildings, language services organizations, or community colleges. Transit agencies should work with their local partners to identify where these services are located and what their hours of operation are to ensure that they are properly served. Often, providing service to these types of destinations during off-peak times creates inefficient routes, and that is often expected; these types of "lifeline" routes are important to ensure vulnerable populations have access to the services they need.

6.2.6. Trip Planning Accessibility

Transit agencies often provide information about their services through schedule brochures that show the route pattern and scheduled arrivals at each timepoint. While this is helpful information, it requires that the user understands how to extract the information that they need from the schedule. Providing data for trip planners can make understanding the service and planning a trip much easier for a rider.

Figure 37: Champaign-Urbana Mass Transit District (MTD) Trip Planner



Trip planners allow users to enter a starting location and destination, and the trip planner will find the best way to make the trip using transit services. The trip planner may also offer additional options, such as : choosing a date or timeframe for the trip, which is important if service is substantially different in peak versus off-peak periods; choosing travel preferences, like minimizing transfers or reducing walking distance; and choosing street addresses rather than existing stops, which is helpful if the user does not know where the nearest bus stop is located.

An agency can produce a GTFS feed that can be published online for use in existing trip planners. Alternatively, an agency can build and host its own trip planner on its website. Either way, providing schedule data to a trip planner increases riders' ability to understand and use the service.

Trip planners work from GTFS schedule data, but providing real-time data can also be valuable information for riders. Real-time data can estimate when a vehicle will arrive based on the actual vehicle location, which provides riders more realistic information about their trip. Real-time data generally comes from automatic vehicle location (AVL) data.

Real-Time Signs in Victor Valley

Victor Valley Transit Authority (VVTA) is a rural transit provider with a service area of 425 square miles in California.[68] Their large service makes providing accurate information difficult, but they implemented real-time signage at a major transfer point in Apple Valley using solar power and existing CAD and AVL data. The new sign provides valuable real-time information to riders, improving the rider experience and making it easier to use the service.

Figure 38: VVTA's New Solar-Powered, Real-Time Sign

AGENCY PROFILE: PULLMAN TRANSIT MOBILE APP

PULLMAN TRANSIT IMPLEMENTED A NEW MOBILE APP THAT PROVIDES USERS WITH A TRIP PLANNER, VEHICLE LOCATIONS, ESTIMATED ARRIVAL TIMES, PASSENGER LOADS, AND THE ABILITY TO PROVIDE FEEDBACK TO THE AGENCY.

Context

Pullman Transit, located in eastern Washington, deployed a new mobile app in 2019 when it updated its AVL system. The agency operates fixed-route services in the city of Pullman, which is home to Washington State University (WSU), a PAC-12 school. Most routes connect to major destinations on campus and also to off-campus student apartments. Over 75 percent of riders on the system are WSU students, so Pullman operates a reduced "community service" schedule when classes at WSU are not in session. The agency also operates school services for the city's middle and high schools.

Figure 39: A Pullman Transit Bus in Operation

When it came time to procure a new AVL system, Pullman Transit issued a request for proposals (RFP) for one that would include a mobile app as part of its system. They eventually procured a system from GMV Syncromatics. The mobile app has many capabilities, including real-time vehicle tracking, passenger load tracking, a trip planner, automated announcements, and the ability to receive feedback. The passenger load tracking is critical to riders, as buses often experience overcrowding on certain trips just before class

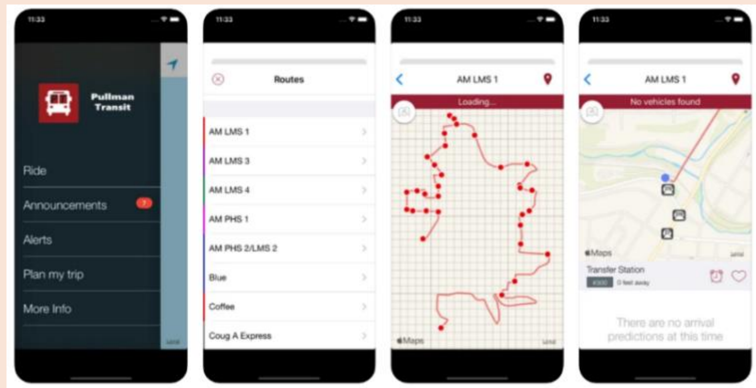
[68] Victor Valley Transit Authority. 2015. "VVTA Tests Real-Time Transit Sign in Apple Valley." *Mass Transit Magazine.* https://www.masstransitmag.com/technology/press-release/12069296/victor-valley-transit-authority-vvta-vvta-tests-high-deserts-first-realtime-transit-sign-in-apple-valley

Context (cont.)

start times, leading to "pass-bys." The feedback Pullman receives from riders through the app is also important as it helps identify real-time issues. The Syncromatics suite also includes a program called Transitcheck, which allows drivers to do their pre-trip and post-trip maintenance checks digitally.

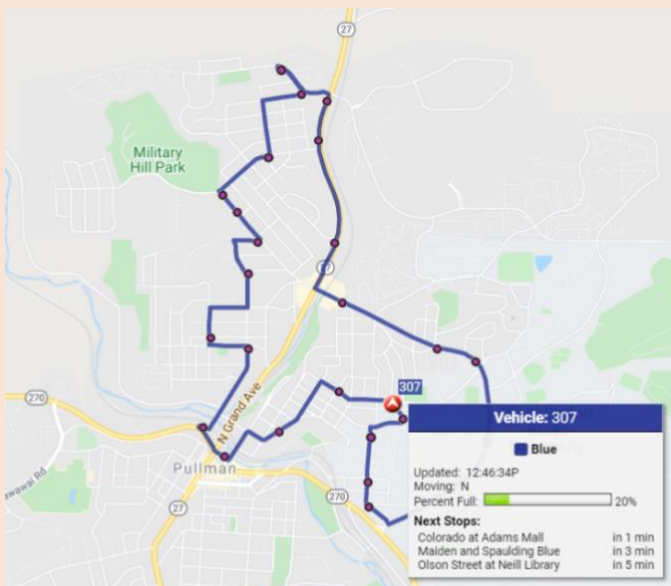
Resources Needed

Figure 41: Pullman Transit's Mobile App



Pullman Transit staff wrote an RFP for a new AVL system and accompanying mobile app in-house and then scored and selected a final vendor. They used 100 percent local funding for this procurement, following city guidelines and procurement requirements. The agency had staff on-site when the systems were being installed to learn about the systems in the process. Because the agency has WiFi in its garage and bus yard, it did not need a cellular package to perform regular software updates.

Figure 40: Vehicle 307's Routing and Current Load



GMV Syncromatics, the vendor for the AVL system and mobile app, provides troubleshooting guides for common issues and a designated contact for when issues arise. They also provided training for operators shortly after the system was installed. Replacement parts could be sent whenever necessary and installed by Pullman Transit maintenance staff. Updates and upgrades to the system are automatically available to the agency without additional cost. Like most vendors, financing was provided available over a five- to six-year period, so that the agency did not have to find a large funding source in a single fiscal year.

Pullman Transit staff only needed to enter new schedule information to see updates displayed in the mobile app, including the trip planner. The app also is set to update automatically when the agency switches from its Community Plus service to Community Service when WSU isn't in session.

Results

Pullman Transit tracks the number of mobile app users, which had more than 14,000 in February 2020, the last full pre-COVID 19 pandemic month for which data was available. They also include questions about the app in their onboard survey efforts. Users of the app can send direct feedback through it as well. Overall, agency staff noted that since the implementation of the app, customer service calls have dropped significantly.

Lessons Learned

Pullman Transit staff noted several lessons learned throughout the process of implementing their new AVL system and accompanying mobile app. They include:

- Prior to procuring systems from new vendors, agencies need to research vendors' customer service ratings or capabilities to ensure that ongoing technical support is offered.
- Planning the installation of new systems can be difficult when your agency does not have many spare vehicles. Systems may need to be installed during off-peak periods.
- It is important to be present when new systems are being installed so that agency maintenance staff can learn as much as possible to troubleshoot issues that may arise in the future. When agencies have several different types of vehicles, installations on each type may differ.

Key Takeaways

- In implementing a new AVL system with an accompanying mobile app, Pullman Transit was able to provide better information to its customers.
- Ongoing technical support and an understanding of the technology are important in selecting a new system.

For More Information

Pullman Transit

Pullman Transit Mobile App

AGENCY PROFILE:
BLACKSBURG TRANSIT MOBILE APP

BLACKSBURG TRANSIT (BT) IMPLEMENTED A REAL-TIME BUS LOCATION APP AND LIVE BUS MAP USING A PARTNERSHIP WITH VIRGINIA TECH UNIVERSITY

Context

Blacksburg Transit (BT) provides service in Blacksburg and Christiansburg, VA. Much of the agency's service is centered on Virginia Tech University, the main employer and economic driver in the area. Because over 90 percent of BT riders are Virginia Tech students or staff, BT operates reduced service when classes are not in session. Service in Christiansburg is more community-oriented, serving major residential developments and employment concentrations.

Driven by requests from students for more real-time information, BT first partnered with a computer science program at Virginia Tech to launch a real-time bus location texting service. Users could text a bus stop id to a specified number and receive a text back with the next three bus arrival times at that stop. Each bus stop sign in the system had the bus stop id included on it. Eventually, this service was formalized into an actual mobile app first called "BT4U." The app added additional capabilities, including live bus maps, a link to the BT website, bus stop information, a trip planner, and a feedback option. In addition to information for passengers, the app also sends the unique ID and app-based data to the content-management software on BT's server in order to collect a record of passenger behavior. For instance, if a rider looks up a specific route, that event is captured, and if they actually complete that trip, that event is also collected. Subsequent upgrades to the app-enabled riders to see bus passenger loads on active vehicles – adding further trip planning capabilities. With these updates, the app was also renamed the "BT App."

Figure 42: BT Sign with Stop ID

The latest version of the BT App has a number of new features, including:

- See stop(s) closest to the user's current location on a live map
- Select a route by clicking on a bus on the live map
- Select a stop to view all routes serving that specific stop, including departure information
- Plan a trip by a preference of quickest route, fewest transfers, or least amount of walking
- View all active BT alerts regarding detours, route changes, and emergency information
- View departure times for an entire day by stop

Resources Needed

The major resources that were required to develop BT's real-time information services were the AVL/GPS and APC systems installed on vehicles, internal information technology staff, and the partnership with Virginia Tech. BT's partnership with Virginia Tech's computer science program required extensive knowledge of not only the AVL technology used by the agency but also of programming in general. Without this internal knowledge, BT would have needed to hire an outside consultant or worked with an organization like Code for America to help build all stages of the app.

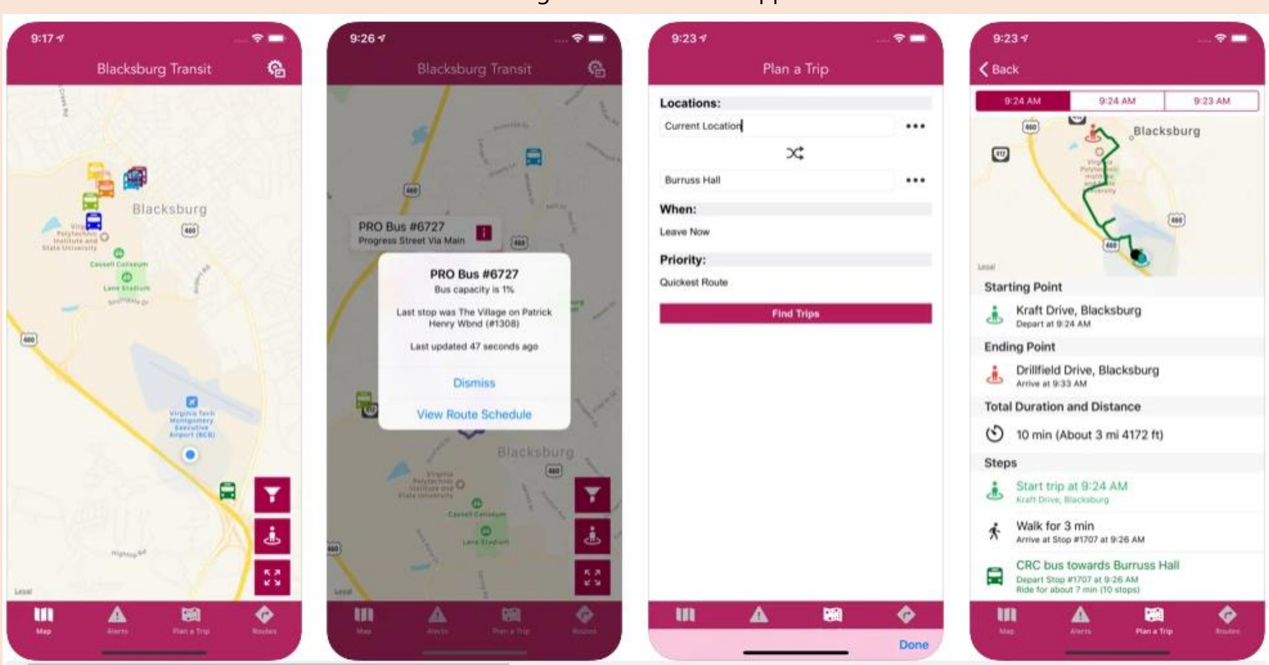
Additionally, staff knowledge needs to include the ability to update and maintain the app, particularly when smartphone manufacturers like Apple or Android make changes to their platforms. Having the ability to store data in the cloud is also important to ensure that data and code are always backed up, and cloud-computing is enabled to speed up the data processing time.

Results

The BT App drastically improved BT's ability to disseminate information to riders and improve its customer service.

- The number of customer service calls was significantly reduced once the app was launched.
- While the majority of BT riders are affiliated with Virginia Tech, the app can also be used for the agency's routes that are focused on Christiansburg, benefiting local riders in this area.
- The inclusion of passenger loads in the app has helped reduce BT's "pass-by" issue (when full buses lead to customers being left at stops), as it allows customers to see in advance when buses are full and alter their trip plans.
- Feedback received through the app has led to its continued improvement, including several improvements to the trip planning component.

Figure 43: BT Mobile App



Lessons Learned

BT learned important lessons when deploying the various iterations of its real-time notification system and mobile app, including:

- When partnering with universities, it is important to account for the differences in expected project timelines between the agency and the university. University students may only have a semester or two to devote to a specific project, whereas technology procurements and planning staff resources at agencies can take two to three years. Setting expectations and allowing for continuing the partnership through agency internships or through specific professors is vital.
- Before deploying new technology, it is important to first define what problem or issue you are hoping to solve with it. Too often, agencies see new tools or technology that they find interesting and procure it before really knowing how it will help them solve problems. In BT's case, the issue was first identified as customers not having real-time information about bus arrivals, leading to frustration when waiting at bus stops.

Key Takeaways

- Agencies providing service in university towns can form beneficial partnerships with university programs to develop new technologies, including mobile apps.
- Mobile apps that include real-time bus locations and passenger loads let customers plan for potential disruptions to their planned trips – particularly when trip planners are also included in the same app.
- In addition to the hardware needed on-board vehicles and a partner with the capability to build a mobile app, staff knowledge at the agency is key to ensuring that mobile apps are sustainable and able to be continually updated through service changes and mobile app vendor updates.

For More Information

Blacksburg Transit

BT App

6.3. Conclusion

Considering issues of access and equity can help a transit agency more thoroughly understand the diversity of its riders and their needs. Conducting accessibility analyses to understand the destinations and jobs accessible by public transit is becoming more widespread as agencies seek to understand how a rider can use their services. Equity and accessibility are intrinsically linked, as differences in transit accessibility between different groups raise issues with equity. Although equity considerations are required by Title VI, there are opportunities to conduct a more thorough analysis. This can include detailed service equity analysis, understanding different customer characteristics and how they relate to the transit services they use, and understanding the range of trip purposes customers use transit for. Collecting and analyzing accessibility, fare, demographic, and trip purpose data enables transit agencies to consider access and equity more thoroughly.

Checklist: What types of equity and accessibility analyses does your agency already conduct, and what types would you like to explore further?

Analysis	Will Help Determine…	Have Conducted This Analysis	Want to Use This Analysis
Service Equity Analysis	…if service changes harm minority or low-income populations.	☐	☐
Accessibility Analysis	…what areas a rider can reach on public transit in a limited amount of time.	☐	☐
Fare Equity Analysis	…if fare changes result in a disparate impact on minority or low-income pouplations.	☐	☐
Customer Profiles	…if routes or fare media are critical to high-need groups.	☐	☐
Trip Purpose Analysis	…if there are unmet needs for particular demographics	☐	☐
Trip Planning Accessibility	…if trip planning is accessible to riders.	☐	☐

KEY THEMES

When it comes to making decisions, data is one of the most important tools that agencies have to ensure that their decisions are beneficial to their customers and the agency. Obtaining data is the first part of the process, whether it comes from traditional pen and paper or automated technology that exports data directly into the hands of analysts. By following data standards, agencies create efficiencies in data sharing and analysis both within their organization and when sharing with outside entities. Increasingly, stakeholders of transit agencies expect agencies to publish open data, and most agencies already make use of open data sources to facilitate both day-to-day operations and planning.

How data is used is just as important as how it is collected, stored, and disseminated. Planning and performance monitoring are two of the biggest uses of transit data. Both internal agency data, open data, and other external datasets allow agency staff to make data-driven planning decisions. Partnerships can also provide novel datasets, and in turn be strengthened when an agency shares its own data effectively. Besides creating more efficient services, data is also critical in ensuring that transit service is accessible and equitable in the agency's service area and that the service provided is useful for different population groups to make different types of trips.

The following are key themes among the data practices summarized in this Guidebook.

The adoption of new technology can not only improve data collection and analysis, but also improve customer access to data

The adoption of most new technologies not only enables agencies to collect data passively but also to provide that data to customers to help with their trip planning. Technologies like AVL systems provide agencies with large samples of runtime, speed, and on-time performance data, and greatly reduce the need for manual ridechecks that provide only small samples of data. At the same time, these systems enable real-time vehicle location information to be disseminated to customers through mobile apps or on agency websites, thereby improving their ability to plan their trips. Similarly, APC systems provide agencies with large samples of ridership data and, at the same time, can in some cases disseminate real-time passenger load data to customers.

Adopting industry data standards let agencies easily use tools that consume common datasets

The adoption of standards such as GTFS opens up new tools for agencies to provide information to customers and conduct different types of analysis. GTFS allows riders to use free trip planning tools such as Google Maps or an agency-branded mobile app. The schedule and spatial data included in a GTFS feed, such as route alignments and stop locations, also allows agencies to more easily conduct Title VI service equity analyses and accessibility analyses using GIS software.

The more data an agency has available, the better it can monitor performance and make planning decisions

Making good planning decisions is dependent on having good data to back those decisions up. Similarly, effective performance monitoring provides an early signal of problems that need an agency's attention—issues that can then be solved through good planning decisions. The more data are available to an agency, the better decisions it can make. For example, when adjusting service levels, agencies can review performance monitoring statistics based on APC or ridecheck data to see if any of their current routes are experiencing overcrowding. Similarly, data showing declining ridership on a specific route or service would allow an agency to make the decision to reallocate its resources away from that service and into areas of greater need.

Agencies can make better decisions using external data, and partnerships can be key to accessing this data

Data collected or created by other entities can be useful to transit agencies when planning new routes or making service modifications. While datasets created by entities like the Census Bureau are open and available online, many datasets maintained by other entities may not be publicly available. For example, regional or statewide travel demand models can provide valuable information on the demand for certain trip connections in a region. Unless an agency forms a partnership with these entities, it may be difficult to access and use this data regularly, let alone interpret it. Similarly, partnerships with major employers or colleges and universities can open up new datasets for agencies, including where workers or students live. This allows agencies to better tailor services to the needs of these unique groups.

Additional Considerations for Decision-Making

Despite the usefulness of data, it should not be the only source of decision-making when it comes to transit planning. Certain datasets, particularly when in aggregate form, can smooth over signs of poor performance that many riders encounter on their day-to-day journeys. For example, on-time performance may be fine when averaged over the course of an entire peak period; however, there may be one or two trips that are consistently late. Those trips hide in the data yet frustrate riders nonetheless. Therefore, things like public outreach, feedback loops, and stakeholder outreach can be just as important as data collection and analysis.

REFERENCES

Baldwin County. n.d. *BRATS Transit.* Accessed January 2021. https://baldwincountyal.gov/departments/BRATS.

Craig, Thomas. 2020. "GTFS-flex." *Trillium & Minnesota Public Transit Association.* October. https://www.mpta-transit.org/sites/mpta/files/uploads/documents/events/2020-10/TechnologyGTFS-flex%20for%20Minnesota.pdf.

Cubic. n.d. "NextBus: Real-Time Passenger Information." Accessed January 20, 2021. https://www.cubic.com/solutions/transportation/nextbus, https://www.nextbus.com/#!/ttc/5/5_1_5B/1405.

Del Peterson, Jeremy Mattson, Kenechukwu Ezekwem. 2020. *ITS Technology Usage and Feasibility in Small Urban and Rural Transit.* Fargo, ND: Small Urban and Rural Center on Mobility (SURTCOM 20-01).

Federal Deposit Insurance Corporation. 2020. "How America Banks: Household Use of Banking and Financial Services." https://www.fdic.gov/analysis/household-survey/2019execsum.pdf.

Federal Register. 2008. "Rules and Regulations." 73 (180): 53384-53390. https://www.govinfo.gov/content/pkg/FR-2008-09-16/pdf/E8-21601.pdf.

Federal Transit Administration. 2016. *Asset Management Guide for Small Providers.* FTA Report No. 0092. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/Asset%20Management%20Guide%20FTA_Report_No._0092.pdf.

Federal Transit Administration. 2012. "Circular." https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/FTA_Title_VI_FINAL.pdf.

Federal Transit Administration. 2016. *FTA Open Data Policy Guidelines.* Washington, DC: FTA Report No. 0095. https://www.transit.dot.gov/research-innovation/fta-open-data-policy-guidelines-report-0095.

—. 2020. "TAM for Tribes." May 15. Accessed January 20, 2021. https://www.transit.dot.gov/TAM/Tribes.

—. 2015. "TERM-Lite." September 29. Accessed January 20, 2021. https://www.transit.dot.gov/TAM/TERMLite, https://www.transit.dot.gov/TAM/TAMPlans/SmallProviderV2_0.

—. 2020. "Transit Asset Management." May 14. Accessed January 20, 2021. https://www.transit.dot.gov/TAM, https://www.transit.dot.gov/sites/fta.dot.gov/files/2020-10/TAM-Systems-Handbook.pdf.

Github. n.d. "Conveyal / gtfs-lib." Accessed January 21, 2021. https://github.com/conveyal/gtfs-lib/.

—. n.d. "MobilityData / gtfs-flex." Accessed January 21, 2021. https://github.com/MobilityData/gtfs-flex.

Google. n.d. "Create a GTFS-realtime feed." *Transit Partners Help.* Accessed January 20, 2021. https://support.google.com/transitpartners/answer/6368391.

—. n.d. "Google Transit APIs." *Static Transit: Reference.* Accessed January 20, 2021. https://developers.google.com/transit/gtfs/reference.

—. n.d. "GTFS Static Overview." *Google Transit APIs.* Accessed January 20, 2021. https://developers.google.com/transit/gtfs#SubmitFeedToGoogle.

—. n.d. "Testing GTFS Feeds." *Transit APIs - Static Transit.* Accessed January 20, 2021. https://developers.google.com/transit/gtfs/guides/tools.

Keenan, Sean Richard. 2019. "MARTA is turning to the Internet of Things to lower costs, reduce system hiccups." *Curbed Atlanta.* February 28. Accessed January 13, 2021. https://atlanta.curbed.com/2019/2/28/18243947/marta-ibm-transit-internet-things-transportation.

Litman, Todd. 2004. "Transit Price Elasticities and Cross-Elasticities." *Journal of Public Transportation* 7 (2): 37-58. https://www.nctr.usf.edu/wp-content/uploads/2010/03/JPT-7-2-Litman.pdf.

Massachusetts Department of Transportation. 2010-2011. "Massachusetts Travel Survey." https://www.mass.gov/doc/2010-massachusetts-household-travel-survey/download.

McHugh, Bibiana. 2013. "Chapter 10: Pioneering Open Data Standards: The GTFS Story." *Beyond Transparency.* Accessed January 22, 2021. https://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story/.

National Academies of Sciences, Engineering, and Medicine. 2020. *TCRP Report 213: Data Sharing Guidance for Public Transit Agencies Now and in the Future.* Washington, DC: The National Academies Press. doi:10.17226/25696.

National Academies of Sciences, Engineering, and Medicine. 2015. *TCRP Synthesis 115: Open Data: Challenges and Opportunities for Transit Agencies.* Washington, DC: The National Academies Press. doi:10.17226/22195.

National Academies of Sciences, Engineering, and Medicine. 2020. *TCRP-210: Development of Transactional Data Specifications for Demand-Responsive Transportation.* Washington, DC: The National Academies Press. doi:10.17226/25800.

National Center for Transit Research. 2016. *Estimating Ridership of Rural Demand-Response Transit Services for the General Public.* Fargo, ND: Small Urban and Rural Transit Center. https://www.nctr.usf.edu/wp-content/uploads/2016/09/21177060-NCTR-NDSU08.pdf.

N-CATT. 2020. "GTFS Flex: What is it and How is it Used?" December 10. https://n-catt.org/tech-university/gtfs-flex-what-is-it-and-how-is-it-used/.

Nuworsoo, Cornelius, Aaron Golub, and Elizabeth Deakin. 2009. "Analyzing equity impacts of transit fare changes: Case study of Alameda-Contra Costa Transit, California." *Evaluation and Program Planning* 32 (4): 360-368. doi:10.1016/j.evalprogplan.2009.06.009.

OneBusAway. n.d. "Homepage." Accessed January 20, 2021. https://onebusaway.org/.

Open Data Handbook. n.d. "Glossary: Open Data." http://opendatahandbook.org/glossary/en/terms/open-data/.

—. n.d. "How to Open up Data." Accessed January 2021. http://opendatahandbook.org/guide/en/how-to-open-up-data/.

—. n.d. "What is Open Data?" https://opendatahandbook.org/guide/en/what-is-open-data/.

OpenTripPlanner. n.d. "GTFS-Flex routing." Accessed January 20, 2021. http://docs.opentripplanner.org/en/v1.5.0/Flex/.

Rosenblum, Jeffrey. 2019. "How Low-Income Transit Ridership is Boston Respond to Discounted Fares: A Randomized Controlled Evaluation." *Massachusetts Institute of Technology, Department of Urban and Regional Planning.* http://equitytransit.mit.edu/wp-content/uploads/2019/06/whitepaper_v8.pdf#page=5.

Southeastern Regional Planning & Economic Development District (SRPEDD). 2020. "FFY 2021 Unified Planning Work Program." https://srpedd.s3.amazonaws.com/wp-content/uploads/2020/07/13213504/FFY-2021-UPWP-Full-122220.pdf.

The National Academies of Sciences, Engineering, and Medicine. 2012. *Data Needs for Assessing Rural Transit Needs, Benefits, and Levels of Service.* Washington, DC: The National Academies Press. doi:10.17226/22715.

The National Academies of Sciences, Engineering, and Medicines. 2004. *TCRP Report 95: Traveler Response to Transportation System Changes Hanbook, Third Edition: Chapter 9, Transit Scheduling and Frequency.* Washington, DC: The National Academies Press. doi:10.17226/23433.

The National Academies of Sciences, Engineering, Medicine. 2017. *NCHRP Report 861: Best Practices in Rural Regional Mobility.* Washington, DC: The National Academies Press. doi:10.17226/24944.

The National Acadmies of Sciences, Engineering, and Medicine. 2017. "Dubuque Smarter Travel Project." Washington, DC. https://trid.trb.org/view/1437892.

The Rapid. n.d. *The Wave Benefits.* Accessed January 2021. https://www.ridetherapid.org/the-wave#benefits.

TheTransitClock. n.d. "Homepage." Accessed January 20, 2021. https://thetransitclock.github.io/.

TransitCenter. 2018. "The Data Transit Riders Want: A Shared Agenda for Public Agencies and Transit Application Developers." December. Accessed January 2021. https://transitcenter.org/wp-content/uploads/2018/12/TC_TransitData_Final_FullLayout_121718-1.pdf.

Transportation Research Board (TRB). Pending. *TCRP G-18: Improving Access and Management of Transit ITS Data.* Accessed January 20, 2021. https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4687.

U.S. Census Bureau. 2020. "TIGER/Line Shapefiles." https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

University of Minnesota. 2017. "Access Across America: Transit 2017." *Accessibility Observatory.* https://access.umn.edu/research/america/transit/2017/.

Victor Valley Transit Authority. 2015. "VVTA Tests Real-Time Transit Sign in Apple Valley." *Mass Transit Magazine.* April 28. https://www.masstransitmag.com/technology/press-release/12069296/victor-valley-transit-authority-vvta-vvta-tests-high-deserts-first-realtime-transit-sign-in-apple-valley.

Wilson et al. 2017. "Lecture 2. Public Transportation Systems: Data Collection Techniques and Program Design." *MIT.* Accessed January 11, 2021. https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-258j-public-transportation-systems-spring-2017/lecture-notes/MIT1_258JS17_lec02.pdf.

n.d. "Yeah, I want to open transit data." *Let's Open Some Government Data!* http://jlord.us/open-data-guide/transit/.

APPENDIX – LIST OF INTERVIEWEES

Interviewee Name, Position	Organization	Date
Elaine Bartoldson, Director Gary Wine, IT Director	Eastern Panhandle Transit Authority, West Virginia; Berkeley County, West Virginia	January 13, 2021
Roman Steichen, Director	Frederick County TransIT, Maryland	January 15, 2021
Brad Rader, Operations Supervisor	Pullman Transit, Washington	January 22, 2021
Chelsea Youngs, Planning Specialist	Crawford Area Transportation Authority, Pennsylvania	January 26, 2021
Tim Witten, ITS Manager	Blacksburg Transit, Virginia	February 5, 2021