

NATIONAL
ACADEMIES

Sciences
Engineering
Medicine

TRB TRANSPORTATION RESEARCH BOARD

TRB Webinar: Improving Access, Quality, and Management of Public Transit ITS Data

June 1, 2023

12:00 – 1:30 PM

NOVEMBER 2022 UPDATE

AICP Credit Information

1.5 American Institute of Certified Planners Certification Maintenance Credits

You must attend the entire webinar

Log into the American Planning Association website to claim your credits

Contact AICP, not TRB, with questions

Purpose Statement

This webinar will discuss various tools for transit data as well as the development and adoption efforts for new specifications that will allow agencies to better organize information.

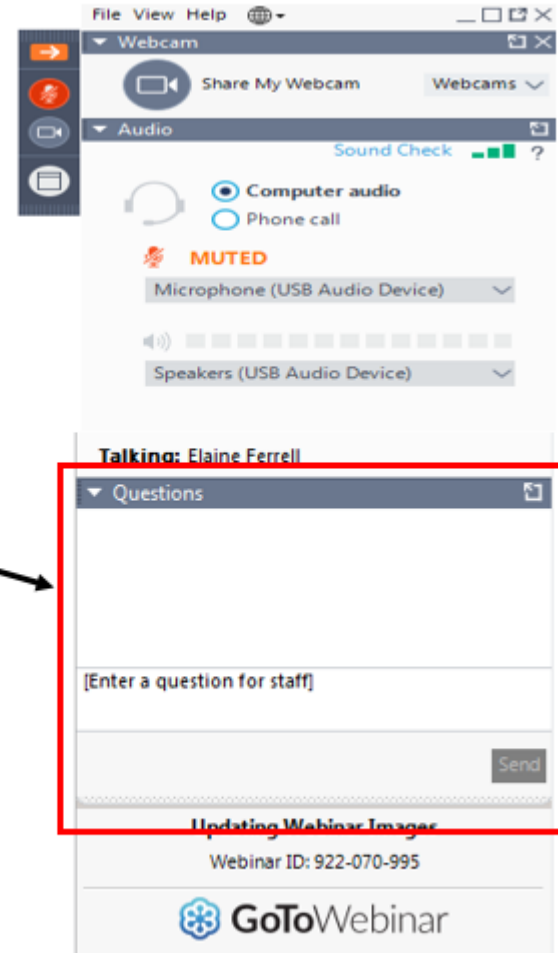
Learning Objectives

At the end of this webinar, you will be able to:

- (1) Access proposed data specification for AVL, APC, and fare data
- (2) Follow the development and adoption of the proposed data specification for AVL, APC, and fare data
- (3) Identify strategies, resources, and processes transit agencies are adopting to become more data driven

Questions and Answers

- Please type your questions into your webinar control panel
- We will read your questions out loud, and answer as many as time allows



Today's presenters



John Levin
Metro Transit
John.Levin@metrotransit.org



Polly Okunieff
GO Systems and Solutions LLC
okunieff@gmail.com



GO Systems and Solution LLC



Cecilia Viggiano,
EBP
cecilia.viggiano@ebp-us.com



Sandy Brennan,
Foursquare ITP
sbrennan@foursquareitp.com



FOURSQUARE ITP

NATIONAL
ACADEMIES Sciences
Engineering
Medicine

TCRP Synthesis Report 153: The Transit Analyst Toolbox: Analysis and Approaches for Reporting, Communicating, and Examining Transit Data

1 June 2023

TRB Webinar: Improving Access, Quality, and Management of Public Transit ITS Data

Presented by
Polly Okunieff



GO Systems and Solutions LLC

Transit Analyst Toolbox Focus

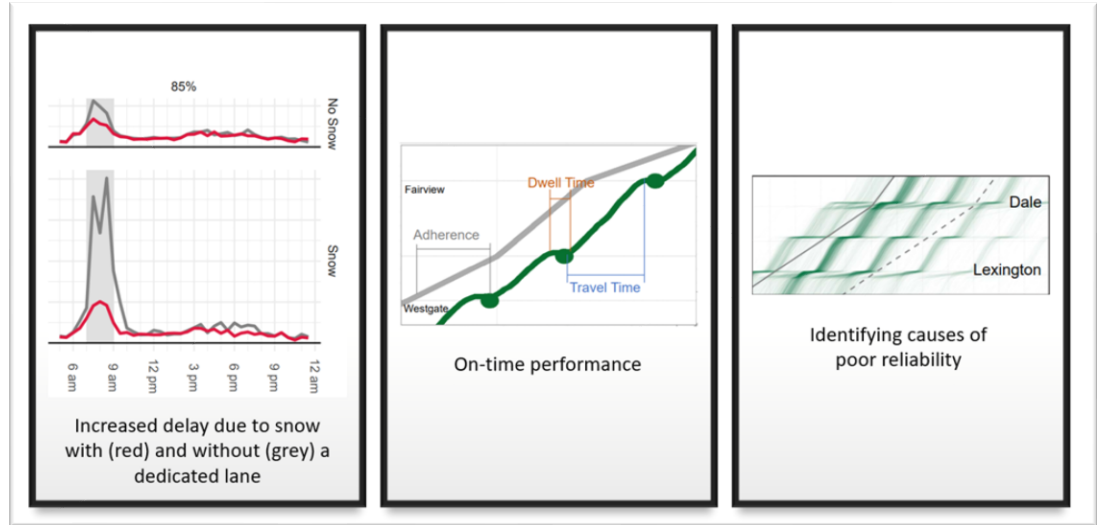
Transit Data

- Transit Service Data and Performance Measures

Transit Data Management

- Data Collection and Management Tools

Transit Data Governance



Transit Data Categories

STATIC SERVICE DATA BY MODE

■ Schedule data ■ Stop and station locations and maps ■ Special event schedules

REALTIME DATA COLLECTED BY MODE

■ Travel times ■ Travel events ■ Passenger Wait times ■ Dwell times ■ Boardings and alightings

PERFORMANCE MEASURES

■ Reliability / On-time performance % ■ Headway Performance % ■ Ridership % ■ Load % ■ Crowding %

Data Management Challenges

• Key Challenges

- Volume
- Quality
- Integration
- Access
- Resources (skills / people)



Example Challenge: Responsibility for managing and synchronizing duplicate data sets

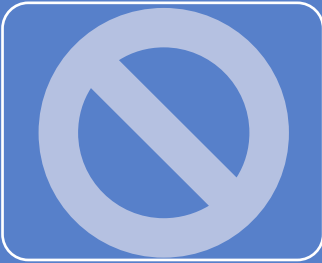
Stop / Station Data responsibilities are shared by several organizational units – facilities, service planning, GIS, operations, and information technology.

- the role may fall to facilities, planning, operations, or a multi-organization data team

Schedule Data does not appear to be well synchronized once it is used by other organizational units or ingested by other tools

Boardings – although planning manages most synchronization for boardings, IT may share the role of managing and synchronizing the data

Data Governance



Data Governance is NOT:

Management of data

The collection, operational and maintenance tasks and tools associated with data.



Data Governance is

Rules of engagement for how institutions (people and policies) manage and sustain data across the enterprise, over the data lifecycle.

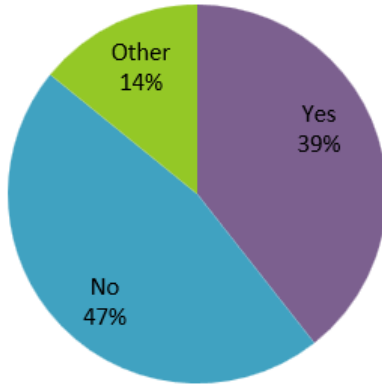
The formal processes applied to ensure that data is

Comments on Data Governance in Transit

- **Data is siloed and inconsistent across datasets**, problems our data warehousing effort is designed to improve. We have no data dictionary, nor a good tool for one. Our IT department is working on this, and in the meantime, we are building a data dictionary for the data warehouse using the **Agile Data Governance framework** using the best tools available to us right now.
- [Data collection and management] lacks governance
- Data governance is a new initiative for ... this upcoming year.

Data Governance in Transit

31. Do you have internal cross-disciplinary committees or groups that focus on managing and sharing service data?



32. Which data sets are governed within scope of the committee?

Value	Percent	Count
Reliability / on-time performance	93%	13
Schedule (including GTFS)	71%	10
Ridership	64%	9
Special event schedules	57%	8
Facilities	36%	5
Travel times	36%	5
Travel events	29%	4

Case Example Categories



Building Blocks to Create a Data Management Ecosystem

- Kitsap Transit (Bremerton, WA)
- Alameda-Contra Costa Transit District (AC Transit) (Oakland, CA)
- Metro Transit (Minneapolis, MN)



Transit Data Governance

- King County Metro
- AC Transit
- Utah Transit Authority



Open-Source Software: Multimodal tools and analysis methods

- Current approaches
- Transit Data Analysis

Data Governance Case Example Lessons Learned

- Initiated governance processes while deploying a new enterprise data management tools (e.g., warehouse)
- Eliminate duplicate data sets prior to establishing data governance
- Store curated data in an enterprise database, of which the rawest data was accessible to everyone.
- Start with a slice of key data (service data is a good start).
- Take an agile approach to implement governance, along with targeting a slice of data, each agency focused on an incremental approach to establishing formal data governance processes, meetings and structures. All three efforts applied “just enough governance” approach so that the organization was eased into changing behaviors.
- Create a role that mediated between the technology and business – collecting needs and quality issues from the business and communicated the information to the technology groups.
- Describe and assign roles and responsibilities to data stewards and data domain stewards (data subject matter experts / business analysts) early in the process.
- Integrate existing processes and people responsible for governance into the data governance framework. Use existing groups and meetings to handle data issues.
- Educate staff responsible for managing data (e.g., IT) on their role and the impacts of their actions on downstream users.

Suggestions for Further Research

Promulgate Best Practices and Guidance on Transit Data

- Guidelines for Transit Agencies for Implementing Data Governance
- Transit Stop Level Data Management Guidebook
- Best Practices in managing data for small and rural transit agencies

Promote research in Big Data and prediction analytics for transit

- Big Data and Transit Planning Opportunities
- Machine Learning and Artificial Intelligence (ML/AI) for Transit Analytics using Big Data

Questions



Thank You

Paula (Polly) Okunieff
GO Systems and Solutions LLC
okunieff@gmail.com

Improving Access and Management of Public Transit ITS Data

Summary of Research

Sponsored by the Transit Cooperative Research Program (TCRP)

Research conducted by:

EBP US, Inc. | IBI Group | Foursquare ITP

Background

Transit agencies are collecting increasingly granular data on vehicle movement, service performance, ridership, customer behavior, and financial recovery:

- **Automated vehicle location (AVL)**
- **Automated passenger counters (APC)**
- **Automated fare collection (AFC)**

However, transit agencies face many challenges in accessing, validating, storing, and analyzing these data sets.

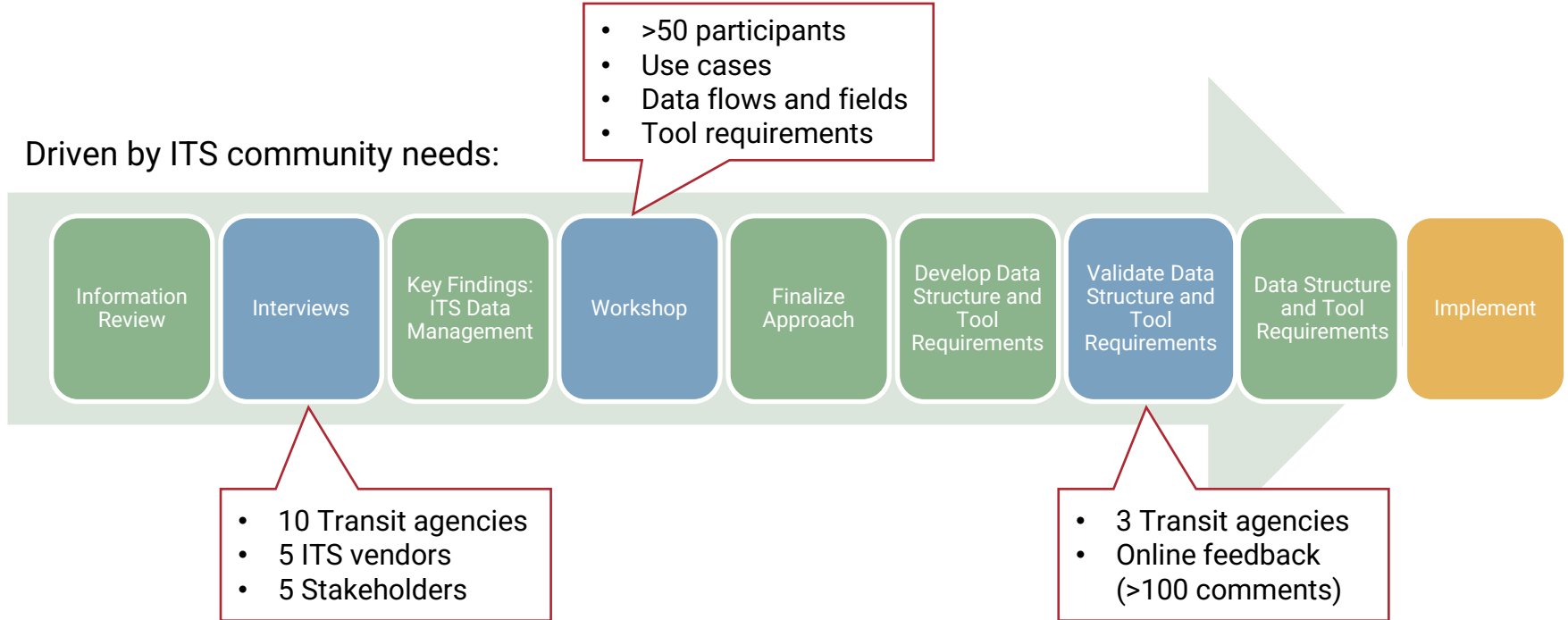
Motivation

This project developed a **standard structure** for historical fixed route ITS data, along with a set of supporting **tool requirements**, to:

- Promote open-source tool and third-party tool development
- Enable transit agencies to share data support protocols and practices
- Encourage vendors to provide data in consistent formats
- Allow transit agencies to more easily generate accurate key performance indicators

Research Process

Driven by ITS community needs:



Key Findings

ITS Data Needs

- Address range of agency capabilities
- Challenges in management, validation, cleaning
- **Support estimates of:**
 - Boardings & alightings (ridership, load, fare revenue)
 - Travel time (runtime, dwell time, on-time performance)

Standard Success Factors

- Flexibility to accommodate diverse transit and data collection systems
- Collaborative end-user-oriented development
- Consideration of cost and effort of implementation
- Support for **Key Performance Indicators**
- Alignment with **GTFS**

Conditions for Adoption

- Ongoing advocacy, education
- Compatibility with procurement cycles
- Community of practice and clear governance processes for management and evolution

Objectives and Approach

Feature	Definition
Outcome-oriented	<ul style="list-style-type: none">• Supports calculating KPIs• Provides documentation of required fields for downstream analysis• Provides best practices for collecting data (including granularity and coverage)
Implementable	<ul style="list-style-type: none">• Does not conflict with and builds on existing standards, in particular GTFS• Is as simple as possible, while meeting transit agency needs• Clearly defines data elements and requirements• Provides documentation for use, including cases with different levels of data availability• Is open source and accessible to all transit agencies
Flexible	<ul style="list-style-type: none">• Can be implemented by a variety of agencies and vendors• Does not require coverage of all three ITS data types• Does not require the same types of data and data availability from all vehicles or modes• Accommodates both event-level and summary data
Extensible	<ul style="list-style-type: none">• Can evolve to accommodate additional modes (e.g., on-demand transit) and types of analysis• Can evolve as transit ITS data evolves

Objectives and Approach

**Designed to
Support High-
Priority KPIs**
(by time period)



On-time performance

- Stop-, timepoint-, or trip-level with adjustable early/late thresholds



Headway spacing

- Average/median/percentile on a route or at a stop



Speed/Runtime

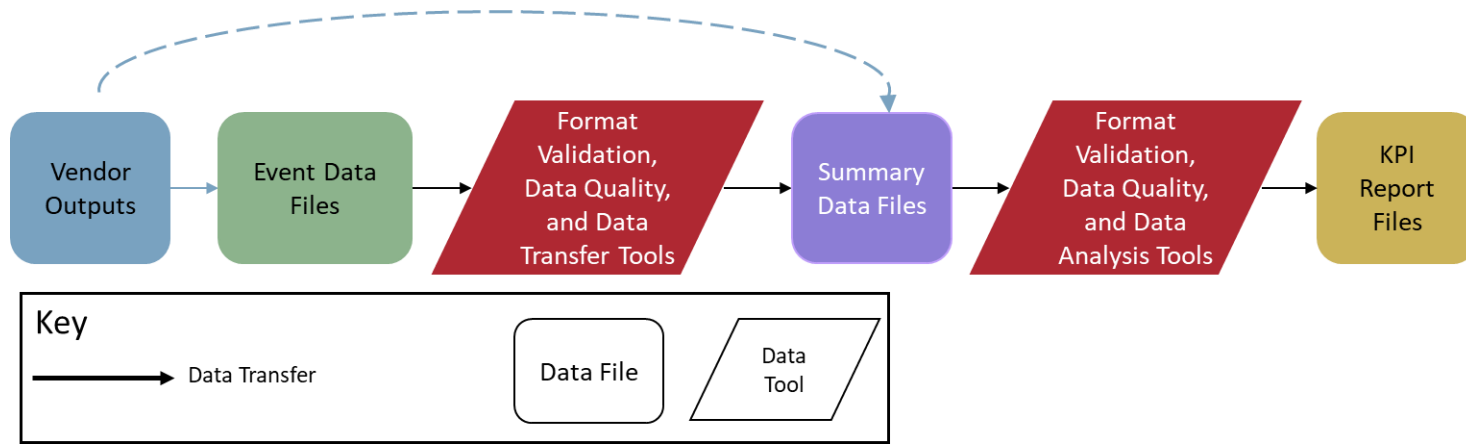
- Average/median/percentile by route or between timepoints



Boarding/Alightings/Load

- Average/median/percentile by stop/trip/route

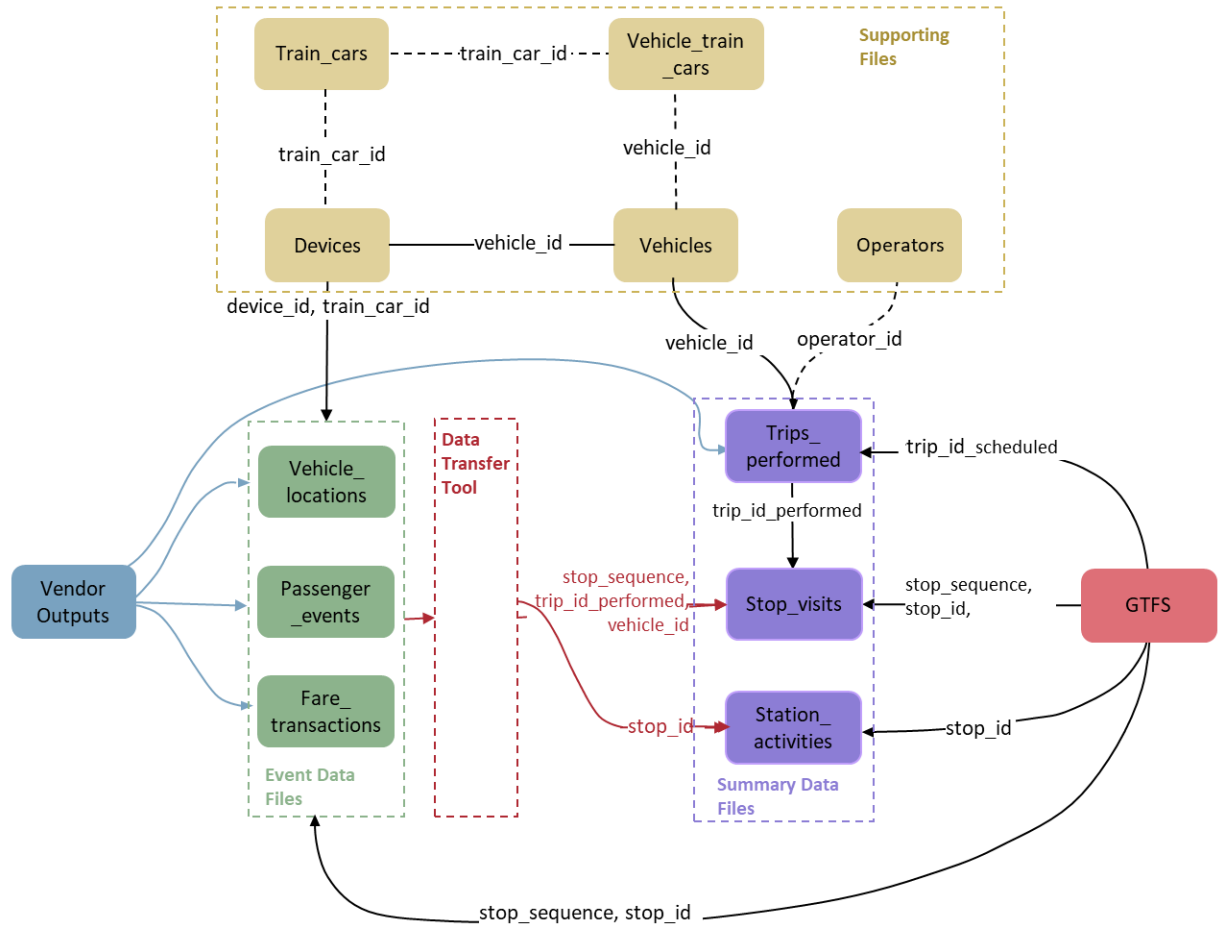
Roadmap



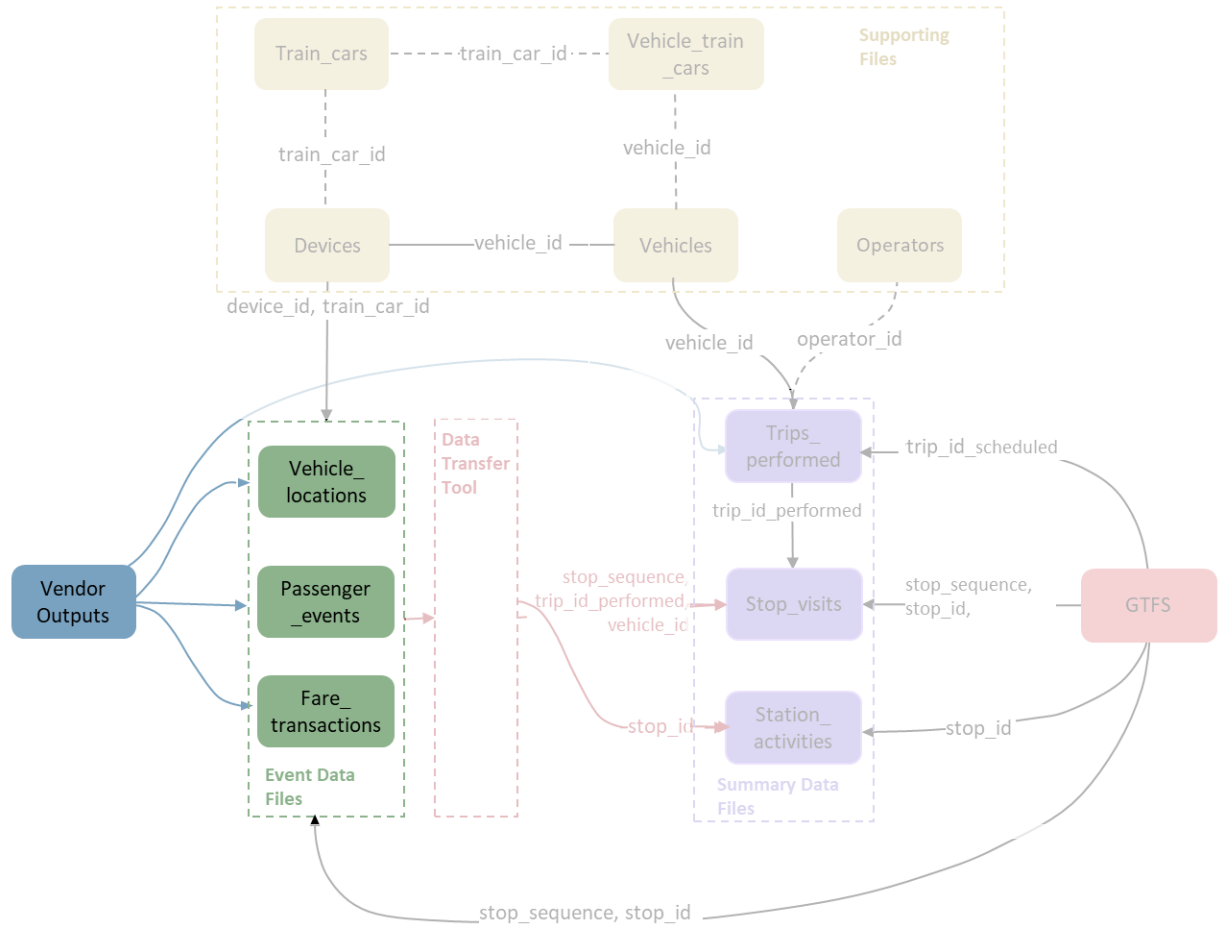
Supports two different paths – one for transit agencies that receive discrete **event data** and an alternative for transit agencies that receive only **summary data files**.

Note: This research developed tool requirements. Tool development is left for follow-up work.

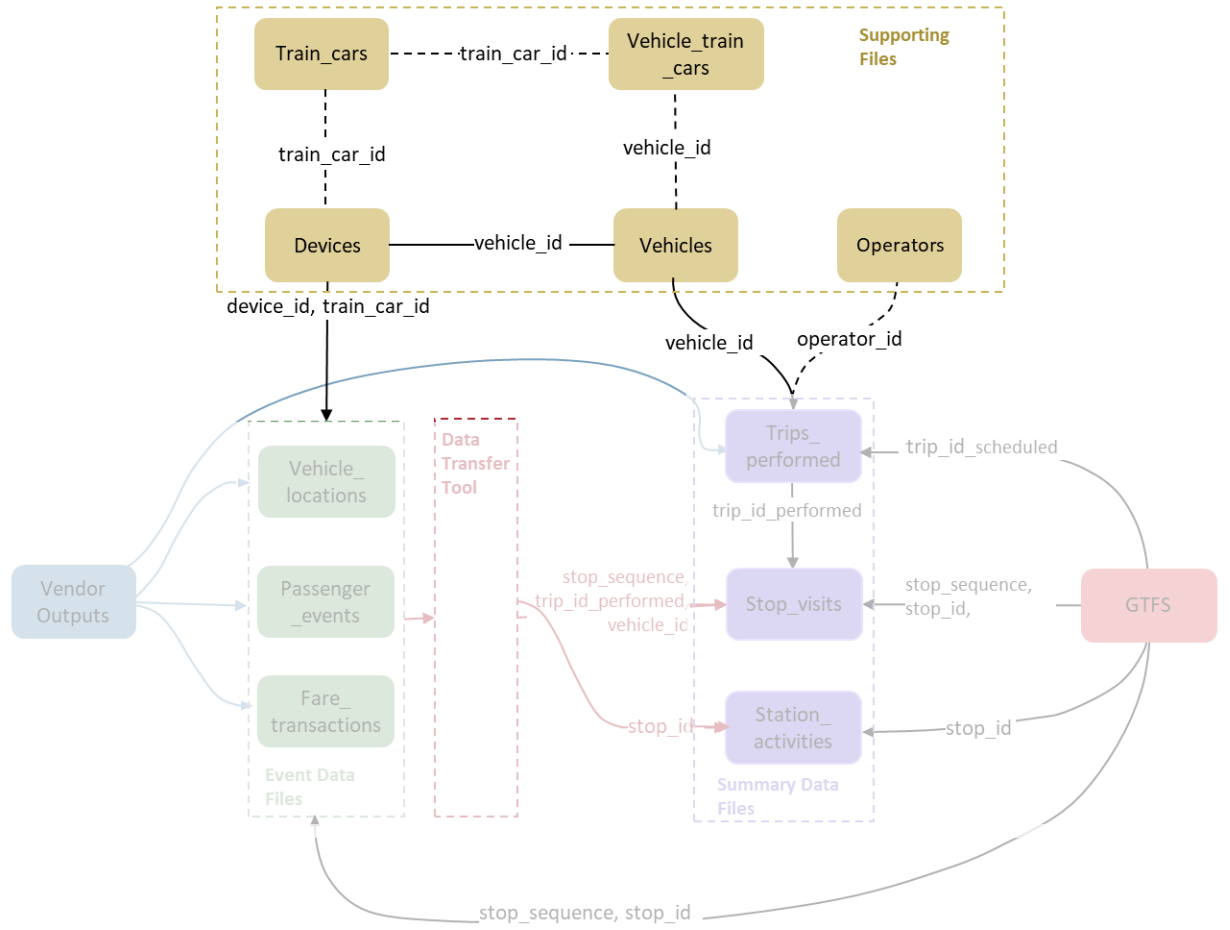
Data Structure



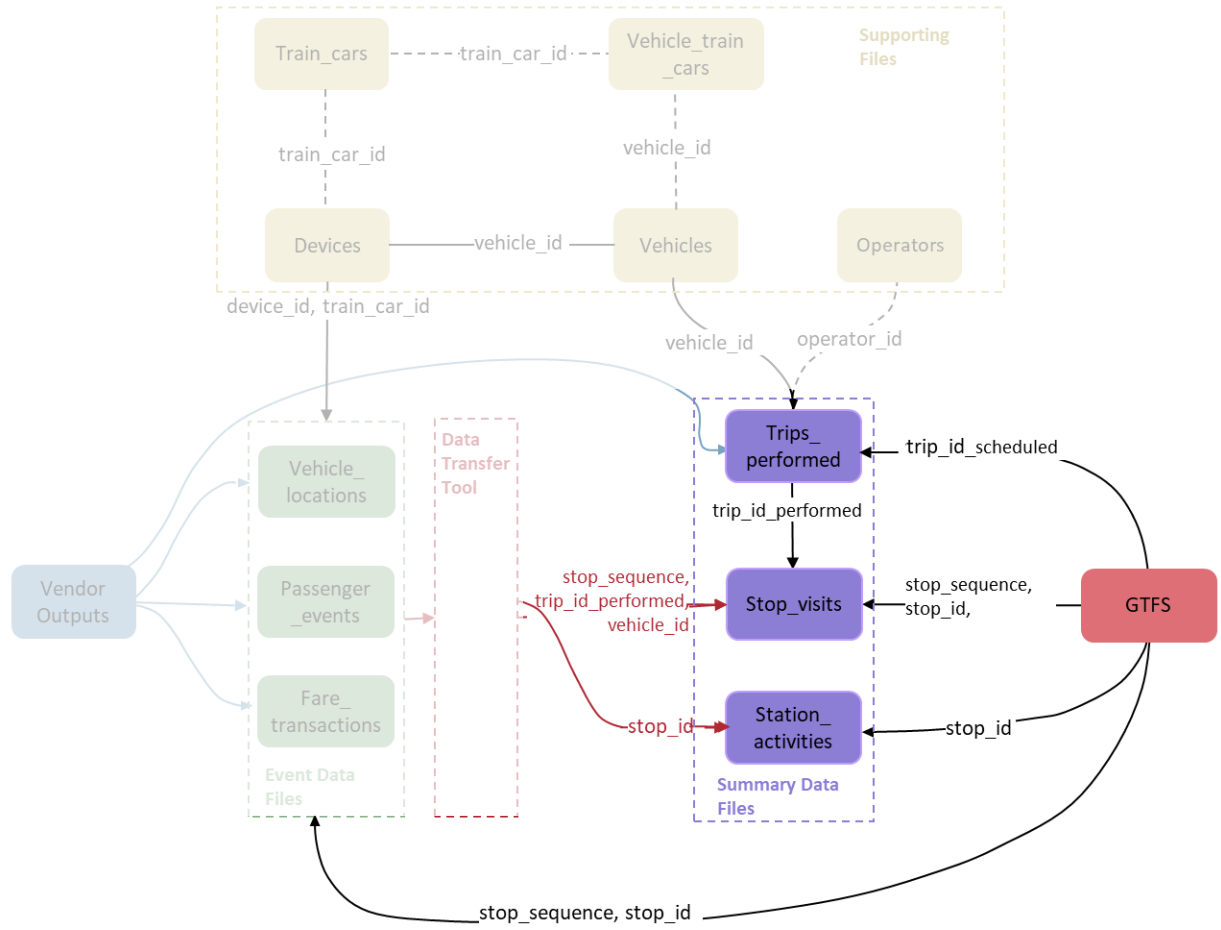
Data Structure



Data Structure



Data Structure



Best Practices

The report provides guidance for using the data structure in a series of best practices that can grow and evolve with adoption of the data structure.

- Versioning data
- Uniting data from multiple ITS sources
- Integrating the data structure with GTFS
- Cleaning data
- Maintaining information on different types of trips
- Generating and maintaining unique identifiers
- Using specific fields
- Using the data structure in special cases
- Communicating and using results

Adoption and Governance

Adoption can be supported through:

- **Tool development** – enables transit agencies to implement the data structure and gain the **benefits** of adoption, such as smoother processes for calculating KPIs.
- **Pilot efforts** – support transit agencies in converting their data to the structure and develop more detailed implementation guidance to reduce the **costs** of adoption for transit agencies

These outcomes will be facilitated by:

- **Community of practice** – building a group of individuals and organizations committed to the success of the effort
- **Clear governance** – establishing agreed upon processes for ongoing communication, promotion, management, and evolution of the structure



Thank You!




Research Into Practice:

Transit ITS Data Exchange Specification (**TIDES**)

... the open data format for historical transit operations data

What is TIDES

- TIDES is an open data format for historical transit operations data
 - Based on results of TCRP Research Report 235
Improving Access and Management of Public Transit ITS Data
 - Build a community of transit agencies, ITS system vendors, software developers, researchers, and others
 - Turn proposed tables and fields into fully developed data format
 - Turn best practices into full documentation of the data format
- 

TIDES Project GitHub


The screenshot shows the GitHub interface for the TIDES-transit/TIDES repository. At the top, there's a search bar and navigation links for Pull requests, Issues, Marketplace, and Explore. Below that, the repository name 'TIDES-transit / TIDES' is displayed with a 'Public' badge. A secondary navigation bar includes links for Code, Issues (14), Pull requests (3), Discussions, Actions, Projects, Wiki, Security, Insights, and Settings. The main content area shows the 'main' branch with 5 branches and 0 tags. A commit by 'e-lo' titled 'Fix cases of redirects (#61)' is highlighted, showing a file list with columns for file name, description, and time. The file list includes .github, docs, spec, .editorconfig, .flake8, .gltignore, and .markdownlint.yaml. On the right, the 'About' section provides metadata for the repository, including its description, license (Apache-2.0), and statistics like 5 stars and 12 watchers.

File	Description	Time
.github	add and cache docs/requirements.txt	4 days ago
docs	documentation enhancement omnibus PR of 2022 (#55)	4 days ago
spec	documentation enhancement omnibus PR of 2022 (#55)	4 days ago
.editorconfig	Add .editorconfig (#57)	3 days ago
.flake8	Add pre-commit and GH Action to validate schemas	11 days ago
.gltignore	Create Documentation site (#38)	11 days ago
.markdownlint.yaml	Add pre-commit and GH Action to validate schemas	11 days ago


<https://github.com/TIDES-transit/TIDES>

- Focal point for community to build and manage the specification
- Establish a “home” to store the specification and documentation
- Create processes to develop the specification, bring forward and resolve issues, and manage revisions

Where are we today


- Working towards Version 1.0 of the data specification
 - Encouraging community to review and contribute to spec
 - Transforming agency data sets into TIDES structure
- 

Where are we going

- Continue to develop and expand the data specification
 - Address need for long-term governance and resources to manage the continued development of the specification and tools
 - Development of tools
 - Data transfer and aggregation
 - Data quality
 - Data analysis and reporting
 - Vendor/system data in TIDES format
 - RFP and contract requirements
 - Vendor adoption
- 

TIDES Resources

...explore more or become a contributor

- TIDES documentation
<https://tides-transit.github.io/TIDES/main/>
 - GitHub Repository – Home for specification and data samples
<https://github.com/TIDES-transit/TIDES>
 - TCRP Research Report 235 – Background on needs and approach
<https://nap.nationalacademies.org/catalog/26674>
 - Google Group – Announcements, past communications
<https://groups.google.com/g/tidesproject>
- 



Thank you!

John Levin
Metro Transit
john.levin@metrotransit.org

Today's presenters



John Levin
Metro Transit
John.Levin@metrotransit.org



Polly Okunieff
GO Systems and Solutions LLC
okunieff@gmail.com



GO Systems and Solution LLC



Cecilia Viggiano,
EBP
cecilia.viggiano@ebp-us.com



Sandy Brennan,
Foursquare ITP
sbrennan@foursquareitp.com



Upcoming events for you

July 8, 2023

TRB's National Conference on
Transportation Asset Management



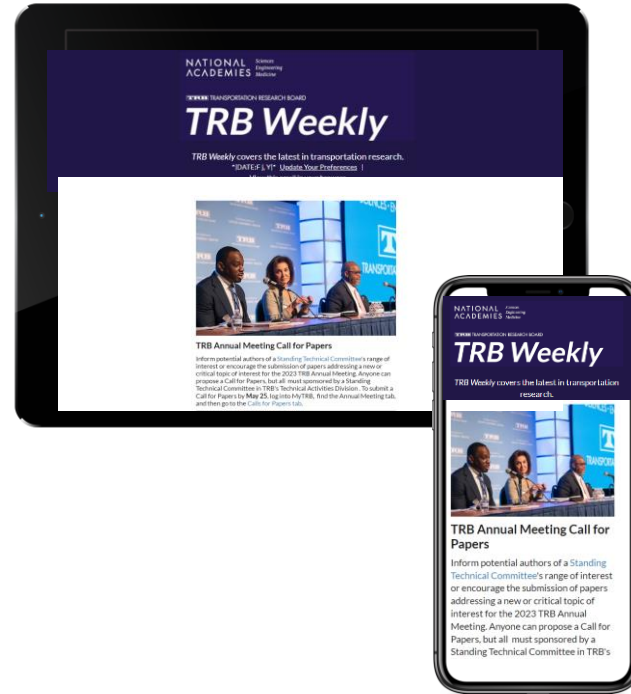
[https://www.nationalacademies.org/trb/
events](https://www.nationalacademies.org/trb/events)

Subscribe to *TRB Weekly*

If your agency, university, or organization perform transportation research, you and your colleagues need the *TRB Weekly* newsletter in your inboxes!

Each Tuesday, we announce the latest:

- RFPs
- TRB's many industry-focused webinars and events
- 3-5 new TRB reports each week
- Top research across the industry



Spread the word and subscribe!

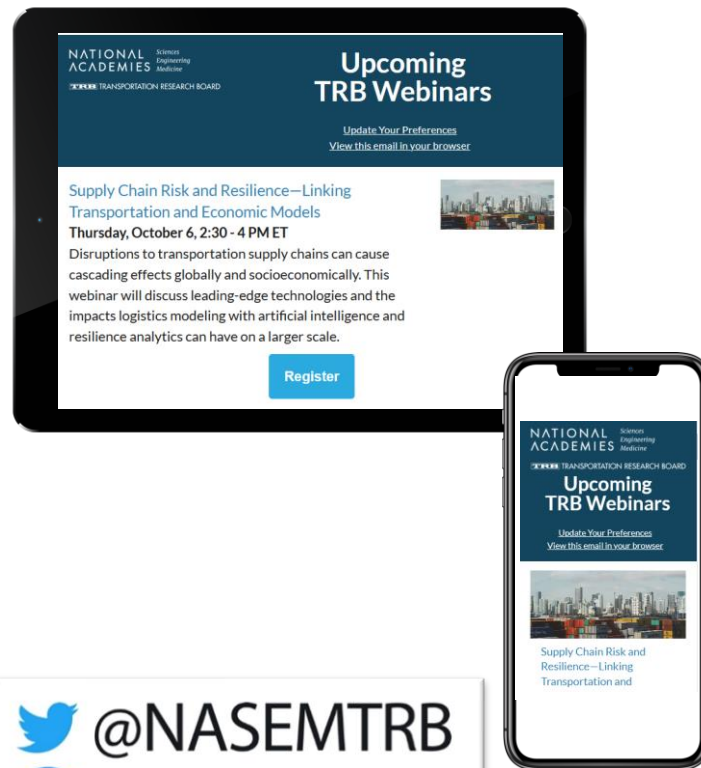
<https://bit.ly/ResubscribeTRBWeekly>

Discover new TRB Webinars weekly

Set your preferred topics to get the latest listed webinars and those coming up soon every Wednesday, curated especially for you!

<https://mailchi.mp/nas.edu/trbwebinars>

And follow #TRBwebinar on social media



Get involved

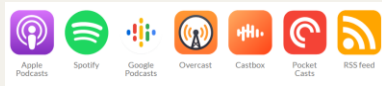
<https://www.nationalacademies.org/trb/get-involved>

- **Become a Friend of a Standing Technical Committee**

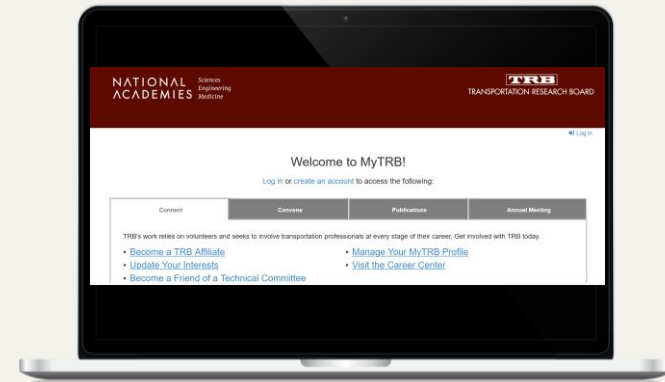
Network and pursue a path to Standing Committee membership

- **Work with a CRP**

- **Listen to our podcast**



<https://www.nationalacademies.org/podcasts/trb>



We want to hear from you

- Take our survey
- Tell us how you use TRB Webinars in your work at trbwebinar@nas.edu

