**GTFS-ride: THE FIRST STANDARD FOR TRANSIT RIDERSHIP DATA​**

**High-Level Summary of Key Challenges**

**First Consortium Meeting**

**Wednesday, June 13, 2018**

This document summarizes the salient points gleaned from the discussion and feedback recorded during the GTFS-ride Consortium Group webinar on June 13, 2018. Most of the feedback was in response to direct questions prompted by the project team, but there is also some discussion that developed from interactions among consortium group members during the conference call. As reference, the main questions prompted by the project team are restated below:

1. Why are you interested in ridership data standards and GTFS-ride?
2. If you have reviewed the GTFS-ride data format, do you see any fundamental problems with its approach?
3. How do we avoid an endless conversation where software vendors say, "if a significant portion of our customers and prospective customers have GTFS-ride data we are willing to support the standard", and transit agencies say "if creating GTFS-ride data will give us access to new/better functionality we will create GTFS-ride data"?
4. How do we move from a data format to a data standard?

Based on the comments recorded during the conference call, the project team has developed a high-level summary and categorization of these comments into eight main topics:

* Standard Adoption Challenges
* Analysis and Planning Opportunities
* Commercial Interests
* Data Quality Challenges
* Interoperability and Standardization
* Logistical and Technical Challenges
* Privacy
* Sharing and Reporting

These topic categories and the individual challenges listed within these categories will serve as the starting point to continue the Consortium Group discussion and steer the GTFS-ride project development.

**STANDARD ADOPTION CHALLENGES**

* Transit agencies, especially those with limited staff, will need help in implementing process changes to use GTFS-ride.
* GTFS-ride could simplify the process of reporting to the National Transit Database (NTD), but there is a need to develop software tools to make this happen.
* Having a “wow” factor will motivate transit agencies to create GTFS-ride datasets. An example of a “wow” factor is MBTA’s dashboard (<http://mbtabackontrack.com>).
* Creating standard operating practices to create a stronger link between ridership data and GTFS trips and stop IDs is very important.

**ANALYSIS AND PLANNING OPPORTUNITIES**

* GTFS-ride and software tools built around could help in linking ridership data and service levels, which would benefit not only individual transit agencies but also would enable data sharing across transit agencies.
* GTFS-ride compliant ridership data could be used in many ways, including analyzing and visualizing ridership and route performance, understanding loading and passenger flows, understanding passenger use, and calibrating and validating modeling software.
* Transit agencies would like software tools that facilitate the conversion to GTFS-ride compliant ridership data and a concrete value proposition to justify the conversion.

**COMMERCIAL INTERESTS**

* There is a consensus by software vendors that ridership data in a standardized format will increase the utility and interoperability of their existing products and/or provide opportunity for the development of new products and services.
* It would help vendors to hear interest in GTFS-ride from their clients to make the case for a value proposition and to drive the adoption and development.
* Some vendors already see the value in the GTFS-ride data standard and are prepared to initiate wider adoption.

**DATA QUALITY CHALLENGES**

* Internally, transit agencies have concerns about the quality of the ridership data they currently collect via automated passenger counter (APC) technology.
* Current data quality issues could make transit agencies hesitant to share ridership data openly.
* GTFS-ride may facilitate quality assurance (QA) and quality control (QC) of ridership data.
* Explore having GTFS-ride accept different quality levels of data, e.g., raw data, corrected but incomplete data, and complete data with inference to fill gaps.

**INTEROPERABILITY AND STANDARDIZATION**

* Both transit agencies/organizations and software vendors recognize the need for standardization to facilitate sharing/reporting/exchanging/analyzing of ridership data and to standardize the exports from and imports to transit software.
* This standardization creates commercial opportunities through the interoperability of data between various vendors and agencies.
* Suggested to focus initially on having GTFS-ride perform well for known specific (perhaps unique) needs of organizations who want to receive ridership data.

**LOGISTICAL AND TECHNICAL CHALLENGES**

* GTFS-ride does not currently provide fields to specifically support average or sample data.
* Metadata is not currently a part of GTFS-ride, and its inclusion would support understanding about the specific representation of the data.
* Many current systems used by transit agencies are not representing trips/stops in a GTFS-compliant manner and it will require system-specific processes to resolve the current disconnect.
* Defining GTFS practices that facilitate the use of GTFS-ride could be beneficial.

**PRIVACY**

* Both the field *rider\_id* and patterns of activity within the GTFS-ride file *rider\_trip.txt* could be used to expose identity of a specific rider. With rural stops that only serve limited residences, stop-level data may cause similar exposure.
* Different GTFS-ride feeds could be compiled for release or internal use with varying levels of obfuscation to protect rider privacy. Alternatively, an agency might elect not to release *rider\_trip.txt* to the public.

**SHARING AND REPORTING**

* Some transit agencies are willing to share ridership analysis tools they have developed internally.
* File size when preparing GTFS-ride compliant data, especially at the stop-level, could be a challenge.
* Open/Public vs. Private Data.