

Understanding GTFS: An intro and overview

What is GTFS?

The General Transit Feed Specification (GTFS) is an Open Standard (https://www.interoperablemobility.org/definitions/#open_standard) used to distribute relevant information about transit systems to riders. It allows public transit agencies to publish their transit data in a format that can be consumed by a wide variety of software applications. Today, the GTFS data format is used by thousands of public transport providers.

GTFS is split into a schedule component that contains schedule, fare, and geographic transit information and a real-time component that contains arrival predictions, vehicle positions and service advisories. Information about routes, schedules, and fares are represented in basic text files, which means it can be created and maintained without the use of complicated or proprietary software.

GTFS is supported around the world and its use, importance, and scope has been increasing. It's likely that an agency you know already uses GTFS to represent routes, schedule, stop locations, and other information, and that riders are already consuming it via various applications.

Why Use GTFS?

GTFS is used by over 10,000 transit agencies in over 100 countries. Most transit agencies have heard of GTFS, and it has quickly become an industry standard. Some agencies produce this data themselves, while others employ a vendor to create and maintain data for them. And because it is a simple, text-based Open Standard, many transit technology vendors can already read and write to GTFS files. By better understanding GTFS, agencies can make better choices when it comes to data. The choices agencies make in how to maintain and distribute GTFS can have a huge impact on service quality.

Open Data means more opportunities and choices

GTFS is an Open Standard (<https://gtfs.org/>). This means that agencies can make information available using any of the many tools which already support GTFS (including simple text editing using a text editor or a spreadsheet). Open standards lead to the creation of data that can be easily shared. A feed is simply the collection of text files that describe a service, hosted online at a permalink that's publicly available. The same feed can be used by Google, Apple, Transit App, Open Trip Planner, and even apps created by riders. Anyone who wants to provide accurate and up-to-date transit information can use a GTFS feed to do so.

Some riders like to use different apps depending on their needs—having GTFS lets riders choose what trip planning app suits them best. Some apps may be more accessible or better at providing information for riders with disabilities, some may be simpler and easier to use, and sometimes riders just want the newest app.

GTFS can probably do more than you think

GTFS is most widely known for trip planning information, particularly in metro areas with fixed-route service. However, there are a variety of optional features above and beyond the basic GTFS Schedule specification that make GTFS more widely applicable, including Fares for showing fare costs and structures, Flex (in development) for demand-responsive transit options, such as dial-a-ride and paratransit services, and Pathways for displaying accessibility information that's vital to riders using mobility devices or needing additional accommodations. GTFS Realtime builds upon GTFS Schedule and on-vehicle GPS systems to provide real-time updates on vehicle locations.

GTFS is more than just trip planning

GTFS data is now being used by a variety of software applications for many different purposes, including data visualization and analysis tools for planning. Having up-to-date and high quality data provides accurate transit information not just to riders, but to planners and policymakers who are able to better understand how transit is being used in their communities. Beginning in 2023, the United States' Federal Transit Administration will require transit agencies there to submit valid GTFS data with their annual National Transit Database report.

(<https://www.federalregister.gov/documents/2023/03/03/2023-04379/national-transit-database-reporting-changes-and-clarifications>).

What is High Quality GTFS?

High quality GTFS is complete, accurate, and up-to-date. This means that it represents how services are currently operating and provides as much information as possible.

Complete Data

Quality GTFS includes important service details such as holiday and summer schedule changes, accurate stop locations, and names for routes and headsigns that match other rider-facing materials. Even if an agency works with a vendor to produce GTFS, it's ultimately up to the agency to ensure that the information presented in print, on board, and online is consistent.

For information on creating high quality data, see the California Transit Data Guidelines (<https://dot.ca.gov/cal-ntp/california-transit-data-guidelines>) and the GTFS Best Practices (<https://gtfs.org/schedule/best-practices/>)

Up to Date

Having out of date data is almost worse than no feed at all. It's not enough to simply publish information—it has to match what the rider sees and experiences. Some of the largest transit agencies update their GTFS weekly, or even daily, but most agencies will need to update their GTFS every few months, or a few times a year when service changes. This includes things like new routes or stops, timetable changes, or updates to fare structure.

Many agencies hire a vendor to create and manage their GTFS for them. Some vendors may be proactive in asking about service changes, but it's important that agencies communicate with vendors about upcoming service changes. It's possible to publish GTFS with service changes in advance, making sure the transition goes smoothly for everyone—agencies, vendors, trip planners, and riders!

Resources

The following resources are presented to help you find available information and tools. Inclusion in this list is not an endorsement of any specific tool or resource and it is up to your discretion to determine which options are most appropriate for your agency.

Technical details about GTFS, what it is, and how to create and maintain data:

- GTFS Schedule Overview (<https://gtfs.org/schedule/>)
- World Bank "Intro. to GTFS" online course (<https://olc.worldbank.org/content/introduction-general-transit-feed-specification-gtfs-and-informal-transit-system-mapping>)
- MBTA GTFS Onboarding (https://mybinder.org/v2/gh/mbta/gtfs_onboarding/main?urlpath=lab/tree/GTFS_Onboarding.ipynb)

View example feeds with various features:

- GTFS Mobility Database (<https://database.mobilitydata.org/>)
- Transitland (<https://www.transit.land/>)

For free tools and instructional materials:

- MobilityData GTFS Schedule Validator (<https://gtfs-validator.mobilitydata.org/>)
- NRTAP lessons and GTFS Builder (<https://www.nationalrtap.org/Technology-Tools/GTFS-Builder/Support>)
- Arcadis IBI Data Tools (<https://www.ibigroup.com/ibi-products/transit-data-tools/>)

For ideas on vendors who offer GTFS services:

- Center for Urban Transportation Research, University of South Florida List of GTFS Vendors (https://docs.google.com/spreadsheets/u/1/d/1Gc9mu4BIYC8ORpv2IbbVnT3q8VQ3xkeY7Hz068vT_GQ/pubhtml)