Overview

GTFS data can be utilized in many ways, and publicly sharing your agency's GTFS data provides many benefits to your riders and your agency as a whole. These include:

- Integrating your feed into mobile and web-based trip planning applications, allowing riders to plan trips on your system
- Submitting your feed to a GTFS aggregator such as Transitland, which provides a wider audience and more visibility for your agency
- Using tools that allow GTFS data to be visualized and analyzed in Geographic Information Systems (GIS) and other map-based analysis programs

This playbook discusses the many benefits to sharing your GTFS data publicly, as well as best practices for doing so, and some potential drawbacks and considerations to keep in mind.

Benefits to Sharing your GTFS

Trip Planning Apps

When your agency's GTFS data is publicly shared, it can be utilized by a variety of trip planning applications in addition to Google Maps. These can include other navigation apps such as Bing Maps + Apple Maps, as well as transit-specific platforms like Transit, Moovit, and Citymapper. Additionally, access to open feed data enables developers to make apps that are geared toward particular regions or can have features that aren't included in the standard trip planners. Two examples include:

- **Vamos** (San Joaquin & Stanislaus Counties, California) - Vamos is a mobile app that supports trip planning and fare payments for transit agencies in San Joaquin and Stanislaus Counties in California, including Altamont Commuter Express, San Joaquin RTD, and Stanislaus County RTA. Within the app, users can plan multi-agency trips and also buy passes from participating agencies. Coming updates to the app include fare payment for all agencies in the trip planner, as well as support for GTFS extensions such as Fares V2, which will allow the app to support and display more complex fare structures, particularly on inter-agency trips.
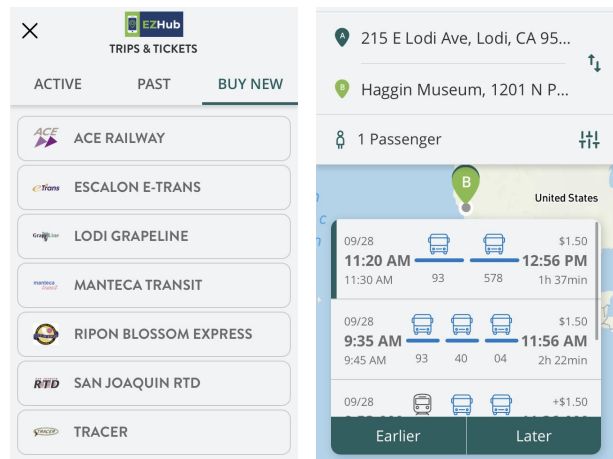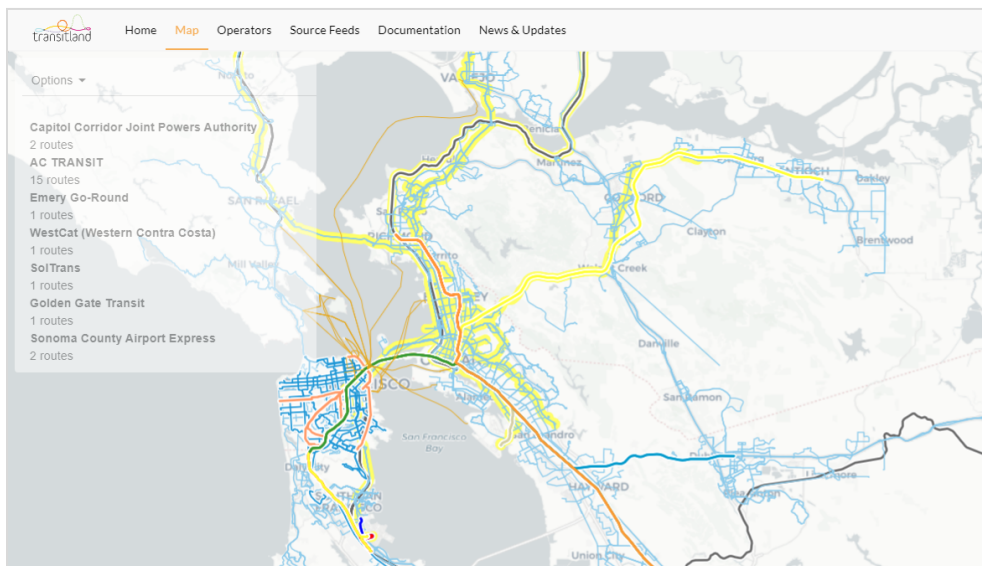
- **MetroHero** (Washington DC Area) - MetroHero is a mobile app that uses WMATA's GTFS-static and realtime feeds to allow you to visualize your commute in multiple ways. Users can view real-time vehicle locations, as well as station-level arrival times, service alerts, and elevator or escalator outages. The app also allows for extensive analysis of current and historical agency data, even allowing users to carry out a "Live Performance Audit" of WMATA headway adherence at the stop level.
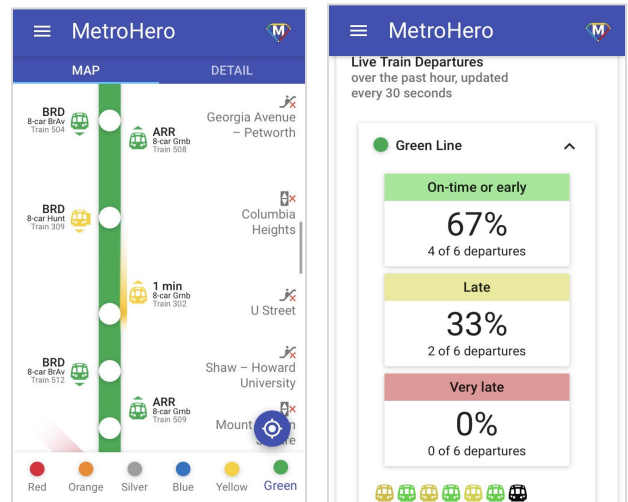
Feed Aggregators

Sharing your GTFS data also allows it to be indexed by GTFS feed aggregation platforms, which can include state- or region-level directories of GTFS feeds, as well as international feed aggregators like Transitland. Being indexed on a feed aggregator increases your agency's visibility and allows developers, researchers, and other interested parties to easily access your agency's data for a variety of purposes, including the development of new applications.

Transitland map visualizing GTFS static feeds in the San Francisco Bay Area

Detail from home page of vermont-gtfs.org, a state-level GTFS feed repository

Integration with GIS, Analysis and other Platforms + Tools

GTFS data can also be imported into and used in a variety of geospatial analysis platforms. Geographic Information Systems (GIS) programs like Esri's ArcGIS, as well as the open-source QGIS have their own plugins and extensions that can import and visualize GTFS stop and route data:

- Esri has a [wide variety of tools and plugins](#) that use GTFS data, including visualizing schedule data
- In QGIS, [GTFS-GO](#) and [GTFS Loader](#) allow you to visualize routes + stops within the platform

Other platforms allow you to visualize and analyze GTFS data in unique ways:

- [Conveyal](#) is an open-source program that allows users to import GTFS data to visualize schedules, routes, and patterns, and analyze impacts of potential service changes. Users can also import and work with demographic data to carry out analyses on, for example, how different routes or schedules would affect access to jobs in a particular urban area.
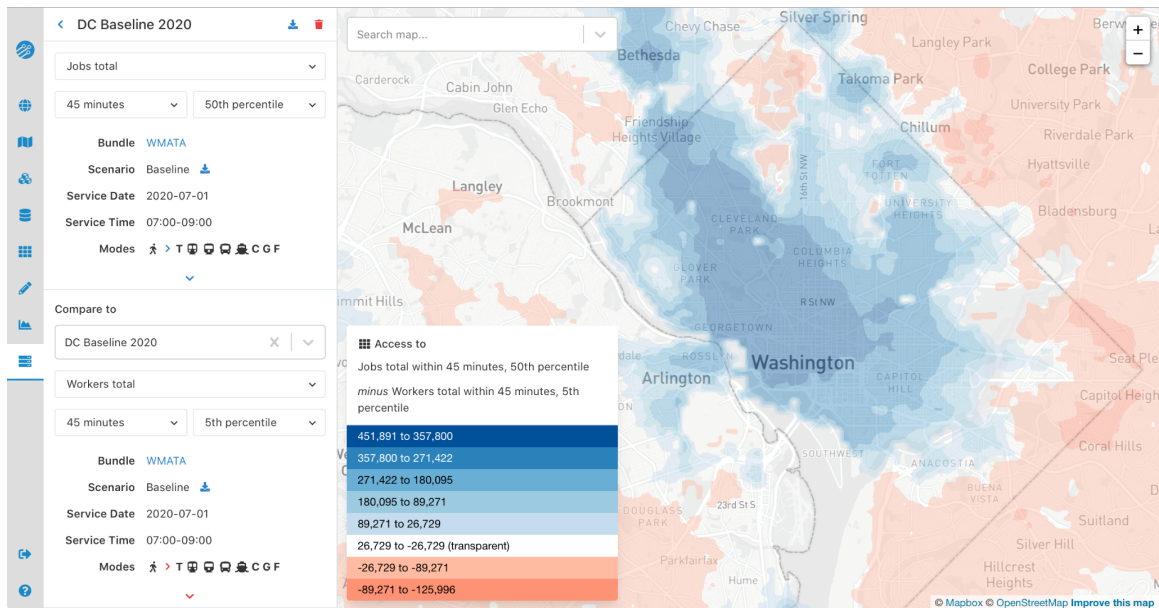
- [GTFS to HTML](#) is an open-source tool that enables the conversion of GTFS schedule data into HTML timetables. It allows agencies to automatically publish and update their schedules to their

website in a format that can also be read by screen readers, making the data accessible to visually impaired people.

Sharing Your Data - Tips & Best Practices

Create and Maintain a Permanent Fetch Link

A fetch link is a permanent URL at which your agency's static GTFS files are stored. Typically, it is hosted either on your agency's website or by your vendor, if you contract with one for GTFS production. This is the way that trip planning apps like Google Maps access your data. Ideally, your static GTFS files should be downloadable directly from this URL without requiring a login. However, if this is not feasible due to licensing restrictions, your agency can control access to the data by using and issuing API keys to users of the data. If your agency does not currently host its static GTFS files on its website, contact your webmaster or IT specialist - they should be able to set up a simple hosting URL for you.

URL and File Names

Consistent fetch link and GTFS file names are crucial to ensuring access to your feed data. If your agency does not use consistent URLs and file names for its data, it means that trip planning apps, feed aggregators, and other users will not get the most accurate and up-to-date data, which will cause problems in the long run.

Once you have set a URL for your permanent fetch link, **DO NOT CHANGE IT.** This means that the URL name should stay consistent, even if the files themselves are updated. As such, keep your URLs as simple and generic as possible, and avoid using URLs that have dates or version numbers in them.
- **GOOD:** http://www.bart.gov/dev/schedules/google_transit.zip,
- **AVOID:** http://www.bart.gov/dev/schedules/google_transit_Fall_2021.zip

Likewise, keep the name of the ZIP folder containing the GTFS static files consistent, even if you make any updates to the feed itself. For example, when you update a feed, you should **not** add any sort of date or version number to the ZIP folder name. If you would like to include data on the feed version or feed start and end dates in the feed, you can include it in the feed_info.txt file.
- **GOOD:** "YourAgency_gtfs.zip", "google_transit.zip", "gtfs.zip"
- **AVOID:** "YourAgency_gtfs_092921.zip", "YourAgency_Fall2021.zip"

File Configuration + Integrity

Your GTFS is a zip file containing several interconnected text files (.txt). In order to ensure the format is correct, always do the following:

1) When opening a text file, make sure you keep the values as **text**. There are a lot of fields in a GTFS that applications like Excel will read incorrectly, or abbreviate.
2) The files are **comma delimited**, not tab delimited. This means each file contains records in rows, and the different fields are separated by commas.
3) Extra rows or columns will cause errors on consuming applications, so make sure there are no empty rows or columns when saving the file.
4) Do not discard any of the files in your GTFS - they work together and any missing files may cause errors with consuming applications.
5) When re-zipping the files, make sure to zip up **the files**, not a containing folder. You can be sure you have done this correctly by unzipping your file and immediately seeing the files in that folder, not another folder containing the files.

Additional Considerations

GTFS is an open data specification, and the best practice is to ensure that when you organize your agency's information into GTFS format, that it is made available for anyone wishing to use it. This means that it is also possible your agency's data could be included in a different GTFS than the one you are maintaining.

The best way to ensure that **your** GTFS is the one that gets used by an application is to make that data publicly available (see *Sharing Your Data - Tips & Best Practices*). This ensures that if a conflict arises about the source of a GTFS, you can definitively point to your GTFS as the best one to use.

There are cases where an application, such as Google Maps, must select data from one GTFS over another. One common example is a stop's name and code - many agencies may share a single stop, each with their own name and code. Because only one name and code is shown on this application, the best way to ensure it matches what you wish riders to see is to **coordinate** your GTFS with other overlapping agencies. If all affected agencies agree upon the name and code, then the likelihood of conflicts between multiple GTFS datasets is minimized.

In the event you have multiple GTFS datasets available to you - usually the result on one being produced for public applications like Transit App, and another being produced for internal operational CAD/AVL systems, you may need to decide which one will be *the* GTFS. It is recommended you choose to promote the GTFS that contains the most **rider-facing information**. GTFS that contain internal/jargony names for routes or stops, lacks fare information, or contains stop locations placed to intersect with the path of travel of the bus (common for CAD/AVL systems) should not be considered for public use. Whenever possible, seek to have your GTFS datasets match (same ids for things like stops and trips) so that internal ones do not conflict with public ones, and integrating other feeds like GTFS-rt feeds is possible.

Resources & Further Reading

- [Cal-ITP - California Transit Data Guidelines](#)
- [GTFS.org - Best Practices for Publishing GTFS Data](#)
- [GTFS.org - Applications and Uses for GTFS Data](#)
- [Principles of Open Data](#)