

GTFS and GTFS-RT case study

A case study of General Transit Feed Specification (GTFS) and the General Transit Feed Specification Realtime (GTFS-RT) extension to support organisations developing or thinking of developing open standards for data.

Our standards case studies provide a summary of the history, development process, governance and key actors involved in the development of an open standard for data.

Contents

Contents	3
About	3
Key facts	4
Design process	5
About GTFS and GTFS-RT	6
Data overview	6
Licensing	6
Contributions	6
History	7
Community	8
Governance	8
Actors	9
Process	10
Initiation	10
Development	11
Adoption	11
Monitoring	12
Revision	12
Lessons learned	14
Involve a wider community	14
Develop under wraps, maintain in the open	14
Have a clear aim (and check demand)	14
Governance can be short and sweet	14
Build on previous work	14
Fragmentation can lead to confusion	14
Resources	15
Tech infrastructure	15
Tools	15

About

This report is part of a research programme funded by Innovate UK. The project aims to make it quicker and easier for organisations to create open standards for data, and to ensure that the standards they create achieve better adoption and wider impact.

The report has been researched and produced by the Open Data Institute, and published in February 2018. If you want to share feedback by email or would like to get in touch, contact the open standards project lead Leigh Dodds at leigh.dodds@theodi.org.

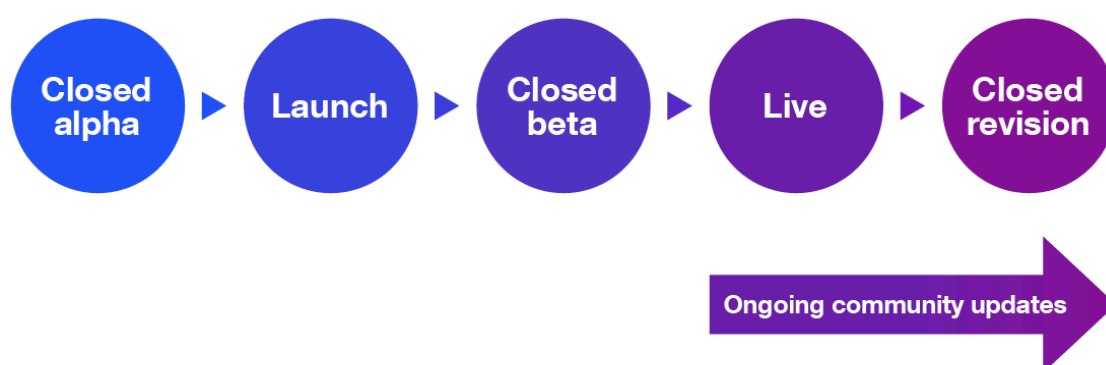


How can it be improved? We welcome suggestions from the community in the comments.

Key facts

Quick description	<p>What if it was as easy to find public transport information as it is to get driving directions?</p> <p>This question is at the heart of GTFS, the standard for transit data. Every time you use CityMapper, Google Maps or even ride-sharing apps like Park.io, you're relying on GTFS</p>
Self described as	Specification
Domain	Transit
Owner	Google
Community	Transit developers (consumers), transit agencies (producers), contributors to standards changes (advocates) and other stakeholders.
Development	Closed (initiation) to fully open (maintenance) Google retains ultimate control over governance and extensions
Launch	GTFS : December 2005 GTFS-RT : V1 in 2011, V2 September 2017 ¹
Adoption	700+ transit agencies worldwide as of 2013, and thousands more in 2017. Accurate adoption figures are difficult to obtain – the public feeds wiki is not up to date.

Design process



¹ Barbeau, S. (2017), 'What's new in GTFS Realtime v2.0', <https://medium.com/@sjbarbeau/whats-new-in-gtfs-realtime-v2-0-cd45e6a861e9>

About GTFS and GTFS-RT



What if it was as easy to find public transport information as it is to get driving directions? –
Bibiana McHugh

This question is at the heart of the [General Transit Feed Specification](#) (GTFS), the standard for transit data. Every time you use [CityMapper](#), [Google Maps](#) or even ride-sharing apps like [Park.io](#), you're relying on GTFS.

The [General Transit Feed Specification](#) (GTFS) and [General Transit Feed Specification Realtime](#) (GTFS-RT) extension are open and common formats for sharing public transport information such as transit schedules with associated geographic information (GTFS) and real-time transit updates (GTFS-RT). GTFS-RT is an extension of GTFS – it isn't possible to share real-time updates without a transit feed.

Data overview

GTFS supports sharing agency transit data including: the agency, routing (stops, routes, trips and maps), scheduling (schedules, stopping times, dates), and other supporting data.

The feed files contain information about the feed, for example the publisher (where it's different from the agency), as well as other useful information for data users, such as when the feed expires. The attributes reflect a simplified feed specification. Transit schedules are built using a calendar, the stop times – with start and end times – and optional information to handle irregular scheduling. The specification is based on a relational data model, shown in the [GTFS feed diagram](#), [reference](#) and [examples](#).

GTFS-RT supplements GTFS with trip updates (when something has changed), service alerts (when something affects stops and the network), or vehicle position (including updates on congestion). Agencies can provide one or more updates, also known as [field entities](#). GTFS-RT supports updating stop or end times and sharing uncertain or approximate delays.

Licensing

GTFS and GTFS-RT are openly licensed under [Apache 2.0](#), while the website content is openly licensed under the [Creative Commons 3.0 Attribution](#) license. [Google developer site policies](#) outline what can and can't be reused; all content and code is openly licensed, except for Google trademarks and other brand features. Google retain intellectual property rights on both standards.

Contributions

Individuals and corporations retain rights to their contributions to GTFS and GTFS-RT but must grant a licence to Google and recipients of software distributed by Google. Individuals sign the [Google Individual Contributor License Agreement](#) (CLA), while corporations must sign the [Software Grant and Corporate Contributor License Agreement](#) before their code can be included on the GTFS codebase. Contributors are known as advocates.

History

GTFS was developed in secret (using non-disclosure agreements) by Google and Portland TriMet. In 2005, [Bibiana McHugh](#) of Portland TriMet reached out to industry leaders Google, Yahoo and Mapquest about making transit information as accessible as driving directions².

GTFS was renamed Google Transit in 2005), then Google Transit Feed Specification in 2009, and finally General Transit Feed Specification, to make the standard more acceptable to a wider audience and reflect its *de facto* status. GTFS is also known as 'GTFS static' or 'static transit'³. The standard is now stable (with a second version released in 2017), openly licensed and managed.

GTFS adoption has been rapid. Google Transit launched in December 2005 with the TriMet schedule, expanding to six cities by September 2006 and 14 by December 2006. By 2013, GTFS was adopted by over 700 transit agencies worldwide and to date, thousands more. Since external monitoring by the [GTFS Data Exchange](#) ended, up-to-the-minute figures have been difficult to come by, with newer monitoring platform [TransitFeeds](#) struggling to keep up with adoption rates.

GTFS-RT, also known as GTFS Realtime, was developed in 2011 to allow real-time updates and as an extension to GTFS. [Version 2.0](#) was released in September 2017. Development began with six agencies as opposed to GTFS' single agency. The standard was developed in secret before being released to the community under an open licence and open management.

The extension is mature, with a robust technical infrastructure to support the demands of real-time data sharing. Data is exchanged in binary format, while small units of code called 'repeaters' or 'adapters' are used to make other feeds interoperable with GTFS-RT.

Repeaters were first created by [OneBusAway](#), an open-source transit app developed to convert King County Metro's legacy MyBus system to GTFS. The OneBusAway repeater also supports converting Transport for Greater Manchester's ATCO-CIF, [NextBus API](#) and [legacy Automatic Vehicle Location \(AVL\) formats](#) to GTFS-RT. Repeaters are developed and run by third parties.

GTFS-RT adoption has been rapid accessing accurate data remains difficult. Adoption is also boosted by the use of repeaters.

Some uses of GTFS and GTFS-RT include:

² McHugh, S. (2013), 'Beyond Transparency: Pioneering Open Data Standards: The GTFS Story', <http://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story/>

³ Google Transit APIs, 'GTFS Static Overview', <https://developers.google.com/transit/gtfs/>

- [CityMapper](#) adopted both GTFS and GTFS-RT for routes, stop locations, schedules and real-time arrival predictions for [Smartbus](#) routes, run in coordination with [Transport for London](#).
- [OpenTripPlanner](#) uses GTFS and GTFS-RT in its open source platform to provide passenger information and transportation network analysis, allowing transit, pedestrian, bike, and car journeys to be combined into itineraries.
- [Google Transit project](#) – which includes Google Maps – is the largest user of GTFS and GTFS-RT feeds. The project incorporates feeds into multi-platform, multi-lingual trip-planning using Google Maps. It also supports [Hotel Finder](#), allowing users of Google Maps to find hotels along a transit route. The same service model applies to petrol stations and other conveniences.
- [Bing Maps](#) also support GTFS.

Other users include ride-sharing apps like [Park.io](#) and [Digital Matutus](#) (a Nairobi civic tech project); mobile apps such as [Nokia Transport for Windows](#); bus alert apps like [Route Shout](#); accessibility travel apps like [Sendero Group BrailleNote GPS](#), and more. GTFS has also been adopted for planning and analysis tools, Interactive Voice Response (IVR) – which gives voice-enabled travel directions – and academic research.

Community

Both standards are aimed at transit agencies and transit developers, known as producers and consumers. The resources are aimed at users with good domain and technical knowledge. All resources are hosted on either Google platforms or [GitHub](#):

- The website and documentation is hosted on the Google developers platform
- The GTFS and GTFS-RT code are on Github
- The GTFS community discussion groups are on Google groups

Governance

Both standards operate under a set of [guiding principles](#) to preserve the original aims of the standard and keep them on track. Governance revolves around ease of use and publication, backwards compatibility and ensuring usefulness.

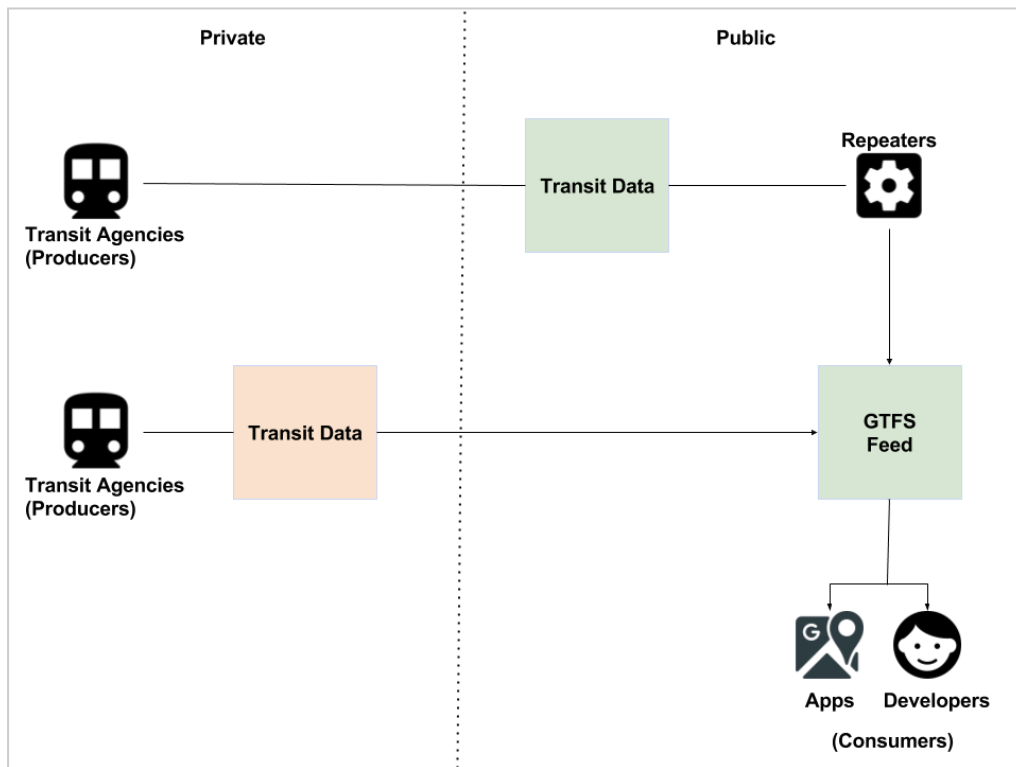
The principles are:

1. Feeds should be easy to create and edit
2. Feeds should be easy to parse
3. Changes to the spec should be backwards-compatible
4. Speculative features are discouraged

Changes to GTFS are driven by the community. There are no explicit hierarchies or working groups. [Google project members](#) make sure the [specification amendment process](#) and [guiding principles](#) are followed. However, the project members are likely to be more involved in the community.

Anyone in the community of producers and consumers of GTFS can propose a change in English using GitHub. Contributors or advocates can grant rights to Google and its customers to use their code. Advocates must submit changes as pull requests which others can then comment and vote on. A change is accepted once a consensus has been reached.

Actors



- **Consumers:** Transit developers, i.e. people or organisations that develop applications using transit information
- **Producers:** Transit agencies, i.e. organisations that run or manage transit services
- **Other stakeholders:** Non-transit organisations and individuals involved in the GTFS community, including researchers and journalists

Process

Each of the following sections summarises a key stage in the development of a standard:

- Initiation – the beginning of the standardisation process
- Development – first development of a new standard
- Adoption – implementation of the standard
- Monitoring – understanding adoption changes in use or requirements
- Revision – adapting the standard, leading to changes in adoption, and so on

Initiation

How did the development of the standard begin?

GTFS⁴

Aims: Make transit information as accessible and easy to use as driving directions

Conceived By: Bibiana McHugh (Portland TriMet) with Chris Harrelson (Google Maps)

Development Time:

- Alpha with TriMet: Five months
- Beta with multiple transit agencies: Nine months

Development process:

Alpha with TriMet: Closed (protected by non-disclosure agreements). Initial script to export data from centralised enterprise database to public CSV. TriMet agency and leadership buy-in backed by existing open source and open data policies, including an open source-friendly procurement policy.

Beta with multiple agencies: Workshop with representatives to address concerns over open publication and to highlight benefits to agencies, including positive press. The feed was updated to include new features, for example, use of csv data format to make it easier for transit agencies to publish transit information.

Team:

- Alpha: Google and Portland TriMet
- Beta: Multiple agencies from Seattle, Chicago, New York, and others

Initial resources:

- Alpha: CSV export from Portland TriMet centralised enterprise database
- Beta: Feed from Portland TriMet, domain expertise from multiple agencies

Current tools:

⁴ McHugh, S. (2013), 'Beyond transparency: Pioneering open data standards: The GTFS story', <http://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story/>

- Website: [GTFS](#)
- Examples: [GTFS examples](#),
- Reference guides: [GTFS reference](#)
- Validation tools: [Validation tools](#)
- Community resources: [GTFS community page](#)

GTFS-RT⁵

Aims: Up-to-date fleet information

Conceived By: Google

Development time: One year

Development process: Closed and protected by non-disclosure agreements

Team: Google, transit developers and transit agencies in Boston, Portland, San Diego, San Francisco, Madrid, and Turin⁶

Initial resources: Broad stakeholder group, [guiding principles](#) and development experience

Current tools:

- Website: [GTFS-RT](#)
- Resources: [GTFS Realtime discussion group](#), Repeaters to convert existing data to GTFS-RT for interoperability

Development

How was the standard developed?

GTFS

Pre-development needs:

Partners with domain knowledge, partners with technical acumen, people to connect the two, staff buy-in (TriMet), agency leadership, and an open data policy that allowed anyone to reuse the data without charge and for any purpose

Development aims: To generalise the initial standard developed with TriMet

Development turnaround: Under five months – Conversations began in July 2005, with first release in December 2005. The initial data schema was generated in approximately one hour.

Development process: Workshops with stakeholders to agree on specification, format and content of files

Development team: TriMet (Bibiana McHugh, agency staff, agency leadership – Carolyn Young), Google (Chris Harrelson), connecting party (Jeremy Faludi)

Contract or agreement: Non-disclosure agreement signed by participants

⁵ Rychev, V. (2011), 'Introducing GTFS Realtime to exchange real-time transit updates', <https://developers.googleblog.com/2011/08/introducing-gtfs-realtime-to-exchange.html>

⁶ Reed, L.T. (2013), 'Real-time transit passenger information: A case study in standards development', <https://smartech.gatech.edu/bitstream/handle/1853/50218/REED-THESIS-2013.pdf>

Resources developed: Unknown

Tools (guidelines, checklists, technology) used: Unknown

Adoption

How did the new standard get adopted?

Initial adoption: GTFS focused on Portland's TriMet agency. Initial interest was seen in Europe overnight at launch and in the US the next day. Bibiana McHugh was interviewed extensively, which helped spread the message. Workshops for the beta were organised with other transit agencies to boost adoption (Seattle, Chicago, and New York). Format simplicity (csv) and media coverage also aided adoption.

Adoption aims: Google is both the owner and largest consumer of both standards. Its aims are to improve products like Google Maps without creating silos that exclude other transit developers.

Increasing adoption: GTFS and GTFS-RT are needs-driven standards, i.e. they focus on satisfying the highest needs by working with producers in the development process and prioritising needs over speculative features. By encouraging open data publication, adoption has surpassed both competing and proprietary standards. Google still controls the standards to a large extent by controlling the initiation stage choices, including the technology infrastructure.

Adoption promoted by: [White House](#), [World Bank](#)⁷, Google, and others.

Monitoring

How is adoption being measured?

It's not clear what monitoring is available for GTFS and GTFS-RT or how it's used by Google and the transit community. Several feeds are available but none provide a complete list of published feeds. [GTFS Data Exchange](#), setup in 2008 (externally) to monitor GTFS, has now shut down, citing growth in adoption. [TransitFeeds](#) is a newer attempt to monitor GTFS and GTFS-RT feeds, however they acknowledge the difficulty in keeping up with new feeds, which appear daily.

Revision

How has the standard been revised and updated?

Initiation: Anyone can initiate a change request for either standard as long as they follow the [guiding principles](#):

- Proposals are submitted on the discussion list for feedback by the community
- Proposers implement and test the proposal with at least one consumer and one publisher
- Results and specification are submitted as a Request For Change (RFC)
- Adoption follows if there are no objections or change requests
- Changes are listed on the changes site

⁷ World Bank (2013), 'Applying the General Transit Feed Specification (GTFS) to the global south: Experiences in Mexico City and beyond', <http://siteresources.worldbank.org/INTURBANTRANSPORT/Resources/340136-1395424136020/GTFS-Mexico.pdf>

Revision approaches: Both standards use a single revision approach.

Revision aims: Changes must be widely useful, tested and receive community approval. See the [guiding principles](#).

Revision turnaround: Turnarounds are fairly rapid, at around seven days.

Versioning: Google maintains control of launching new versions. [Version](#) two of GTFS-RT was released in September 2017.

Managing extensions: Only GTFS-RT has emerged as an extension, hinting that Google retain control of official extensions to GTFS.

Revision stakeholders: GTFS and GTFS-RT community, including transit agencies and transit developers.

Resources developed: Specification, RFC, test plan & results.

Revision tools: [Revision history](#), [changes site](#), [GitHub repo](#)

Lessons learned

Involve a wider community

GTFS was developed with a single transit authority, resulting in immediate change requests on launch. Changes had a big impact and included adding support for fare information and shapefiles, normalising the model to remove dates in every file, expanding route types, allowing multiple, non-overlapping feed files, and fixing bugs (like requiring column headers in every file and making certain attributes optional). GTFS-RT included more transit authorities to encourage more useful and generalisable solutions, with wider international scope. This may have had a positive impact on stability and versioning. GTFS-RT was updated to version two in September 2017, six years after its initial launch.

Develop under wraps, maintain in the open

GTFS and GTFS-RT were both developed under non-disclosure agreements with small groups with the assertion that consensus with a wider audience can be a barrier to rapid development. The size of the group was increased in lessons learned from the initial launch that only featured one transit agency but the process of closed development remained. Once launched, an [open and transparent policy](#) was adopted for maintenance, allowing anyone in the community to suggest improvements. The principle of 'develop under wraps, maintain in the open' was also applied to the new version of GTFS-RT.

Have a clear aim (and check demand)

Google are a primary consumer of data that is published using both standards, however, to ensure wider adoption, both the standards and data produced by transit agencies are open. The aim is to efficiently and effectively share transit information with a wide audience.

The tools and resources reflect the aim and highlight two important applications: trip planning and timetable management. The change management process ensures changes are subject to demand.

Governance can be short and sweet

The initial aims are protected by the four [guiding principles](#), which keep changes on track, but do introduce barriers to extension. Extensions appear to be tightly controlled by Google.

Build on previous work

The GTFS-RT website is more clear and focused than the GTFS website, reflecting the lessons learned from implementation of GTFS. The GTFS-RT website clearly answers questions on data formats, technical infrastructure and copyright, under their own headings.

Fragmentation can lead to confusion

Despite improvements made on the GTFS-RT website, information on important considerations like contributor agreements are found in the [GTFS Github repo](#) and not linked from the main website.

Resources

Resources associated with or used to support the standard

Website: [GTFS](#)

Examples: [GTFS examples](#)

Reference guides: [GTFS reference](#)

Validation tools: [Validation tools](#)

Community resources: [GTFS community page](#), [GTFS Realtime discussion group](#)

Revision tools: [Revision history](#), [changes site](#), [GitHub repo](#)

Tech infrastructure

Both standards benefit from a distributed structure, i.e. producers host their own files and announce availability by registering their feed. There is some support for hosted information using the [Transit Partner Dashboard](#), but it isn't a recommended option.

- GTFS data exchange format is based on feeds (zipped text files for stops, routes, trips and other schedules)
- GTFS-RT data exchange format is based on [protocol buffers](#) (a lightweight format based on XML that can be used with any programming language on any platform and easily extended with new features)

Tools

Google have actively produced a number of open source tools for GTFS and GTFS-RT that support [validation](#), [testing](#), [merging](#), [viewing KML files and more](#). Contributions are welcome under an Apache 2.0 license. Several tools have also been produced by the community to support the production of feeds, convert of existing feed and to overcome technical barriers to using the specifications.

- [Awesome transit](#) is a useful community resource of transit APIs, research, datasets and software
- [General Transit Feed Specification](#) is a useful wiki for getting started. Last updated in 2016
- The [GTFS slack channel](#) is another resource for tools and help with GTFS and GTFS-RT