Navigating the Transit Data Landscape

Nina Kin
Who am I?
Nina Kin

---

Tech Lead
Digital Experience Team
Customer Experience Department

➢ Software engineer
➢ Public transit rider
➢ Urban explorer
➢ Civic tech nerd
➢ Multi-disciplinary crafter and musician
➢ Generally inquisitive person
Nina Kin

- MaptimeLA - Community for map enthusiasts interested in learning to use open source mapping tools. Contribute to open mapping projects like OpenSidewalks. [https://www.meetup.com/maptimela/](https://www.meetup.com/maptimela/)

- Data + Donuts LA - Monthly breakfast lecture series highlighting local government tech and data projects. [https://datadonuts.la/](https://datadonuts.la/)
Nina Kin

---

➢ **LA Arts Datathon** – Brings together members of the arts and tech communities. Coming soon!  https://www.lacountyarts.org/learning/arts-datathon

➢ **International Humanitarian Mapathon** – Annual event teaching people how to contribute to humanitarian mapping efforts. April 23-25  https://mapathon.la/en/

➢ **Board of Directors, MobilityData**  https://mobilitydata.org/
Nina Kin

2006
Berkeley
EECS

13 years
County of Los Angeles

4 years
Metro
Nina Kin

What I do:

➢ Advocate for rider information needs inside the agency.
➢ Improve the way digital information reaches riders.
➢ Advance data standards in the transit industry.
➢ Pilot more sustainable tech and data practices.

... with varying degrees of success! 😊
Nina Kin

---

I am...

➢ Not involved in operations!
➢ Not involved with the hardware or data infrastructure!
➢ Not involved with service or planning!

I DO like to ask questions, learn, and share!
“Transit Data”? 
## Schedule Data

**Metro Local**

**Start**

**Hollywood/Western Station**

**End**

**Chesterfield Square**

**Metro**

**Subject to change without notice**

### Eastbound (Approximate Times)

<table>
<thead>
<tr>
<th>DOWNTOWN</th>
<th>LOS ANGELES</th>
<th>LINCOLN HEIGHTS</th>
<th>ALHAMBRA</th>
<th>ROSEMEAD</th>
<th>EL MONTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5:23a</td>
<td>5:44a</td>
<td>5:59a</td>
<td>6:08a</td>
<td>6:37a</td>
<td>6:58a</td>
</tr>
<tr>
<td>5:48</td>
<td>6:08a</td>
<td>6:37a</td>
<td>6:58a</td>
<td>7:14a</td>
<td>7:31a</td>
</tr>
<tr>
<td>5:55a</td>
<td>6:16a</td>
<td>6:37a</td>
<td>6:58a</td>
<td>7:14a</td>
<td>7:31a</td>
</tr>
<tr>
<td>5:55a</td>
<td>6:16a</td>
<td>6:37a</td>
<td>6:58a</td>
<td>7:14a</td>
<td>7:31a</td>
</tr>
</tbody>
</table>

### Westbound (Approximate Times)

<table>
<thead>
<tr>
<th>EL MONTE</th>
<th>ROSEMEAD</th>
<th>ALHAMBRA</th>
<th>LINCOLN HEIGHTS</th>
<th>DOWNTOWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:30a</td>
<td>4:44a</td>
<td>4:54a</td>
<td>5:12a</td>
<td>5:23a</td>
</tr>
<tr>
<td>4:56a</td>
<td>5:15a</td>
<td>5:26a</td>
<td>5:43a</td>
<td>5:55a</td>
</tr>
<tr>
<td>5:26a</td>
<td>5:37a</td>
<td>5:48a</td>
<td>6:06a</td>
<td>6:19a</td>
</tr>
<tr>
<td>5:44a</td>
<td>5:55a</td>
<td>6:07a</td>
<td>6:26a</td>
<td>6:37a</td>
</tr>
<tr>
<td>5:57a</td>
<td>6:10a</td>
<td>6:22a</td>
<td>6:42a</td>
<td>6:58a</td>
</tr>
</tbody>
</table>
Realtime Data
Data In Your Hands

Google Maps

Transit App
This!

What happens here?
Lots... LOTS! There’s so much.
First Stop: Generating the Data
Service Planners + Schedulers

Lots of factors go into planning & scheduling:

➢ Operator Safety and Experience – very high priority!
  ○ Stops are in safe places
  ○ Minimize the possibility of accidents
  ○ Trip length
  ○ Route familiarity
  ○ Access to essential amenities (bathrooms!)
Service Planners + Schedulers

Community
- Feedback
- Changing traffic patterns
- Commercial/residential developments
- How much we expect the route to be used.
- How it connects to other existing routes (including from other agencies!)

Physical Limitations
- Street width
- Turning angles at intersections
- Length of buses used on a route
Service Planners + Schedulers

➤ Efficiency

- Optimizing for cost/benefit
- Efficiently allocating resources – optimize time spent delivering service versus moving buses and staff around.
- Run the most service for the least cost
- Nearside vs Farside stops

➤ Jurisdiction

- Cities have to agree to route, stop, and layover locations!
- Layover locations and times need to be negotiated with other transit agencies that use them!
Scheduling System

**HASTUS** handles:

- Scheduling
- Operator Assignments
- Service Cancellations
- Cost Optimization
- Operator Payroll
- Long Term Planning

Exported as **GTFS** data
What is GTFS?
What is GTFS?

It is a...

➢ **Data Specification** – a set of rules defining what a dataset should look like for producers and consumers.

➢ **Representation of Transit Service** – a transit agency’s service in a machine-readable format.
Why does GTFS exist?

➢ Without a data standard, everyone would publish their data differently and you’d have chaos.

➢ Consumers of the data would have to create custom code to account for every single transit agency.

➢ LA County alone has over 50 public transit service providers.
Detour: Origins of GTFS
Google Maps in 2005

July 2005

TriMet & Google Partnership

“...make it just as easy to get transit directions as it is to get driving directions from anywhere in the world.”
- Bibiana McHugh, TriMet
Technological Barriers:

- **Machine-readable data already existed**
  - TriMet’s service data already existed in their scheduling system as a set of tables.

- **Data was already accessible as a file export**
  - TriMet’s scheduling system allowed for the data to be exported in CSV format.
An Opportune Moment

---

Business Barriers:

- **Agency already supported Open Source and Open Data**
  - TriMet already had a long history of supporting open source and open data.
  - TriMet was already publishing an open data feed used by outside developers.

- **Procurement policies already existed for Open Source**
  - TriMet already had an open source-friendly procurement policy in place for a decade.

- **Willing partner in private industry**
  - Google Maps was already interested in expanding to include transit.
Google Transit Launches in 2005

Google Transit Launches with TriMet Data

“As Google Transit went live for the first time… the number of hits to the site increased exponentially… the counts were reaching staggering numbers, even by Google standards.”

- Bibiana McHugh, TriMet

GTFS is born! 🐱
GTFS
Google Transit Feed Specification
Google Transit Expands

2007

Google Transit Expands

2009

4G Launches

2008

First Android Launches

First iPhone Launches

More transit agencies recognize the benefits as they adopt and publish GTFS. Google Transit is now available in more than 10 cities and Japan.

Google Transit Expands

2008

First iPhone Launches

2009

4G Launches

Google Transit Expands

2007

Google Transit Expands

2008

First Android Launches

First iPhone Launches

More transit agencies recognize the benefits as they adopt and publish GTFS. Google Transit is now available in more than 10 cities and Japan.
Key Early Decisions

➢ Keep the **CSV format** published by the scheduling systems.
  ○ Keeps GTFS as simple as possible to keep the barrier to participation low for smaller, less-resourced agencies.

➢ Keep the data specification **open and free**.

➢ Make the resulting data **free and publicly accessible**.
Name Changed from “Google” to “General”

GTFS is widely popular!

But concerns are raised so name is changed to “General”, reflecting GTFS usage outside Google products.

Development starting to splinter as unique flavors of GTFS pop up to account for different needs.
GTFS-realtime Extension Released

August 2011

GTFS-realtime extension created by Google and a consortium of transit agencies.

June 2015

Two-Day Workshop on Transit Data Interoperability

Hosted in San Francisco by the Rocky Mountain Institute (RMI), transit data interoperability is seen as a critical foundation for the future of transit.

Results in a recommendation to create Best Practices to grow the space.
GTFS Best Practices Established

GTFS Best Practices Published

February 2017

RMI convenes a working group to identify issues and improve GTFS.

The working group publishes the GTFS Data Best Practices.

October 2017

Interoperable Transit Data Workshop

RMI and TransitCenter host a two-day workshop in New York to identify challenges and opportunities to improve transit data practices.
Grown out of the original working group, MobilityData is created as a non-profit to oversee the continued growth of GTFS.

GTFS is used by over 10,000 agencies in over 100 countries.

MobilityData’s work now covers:

- Public Transit data standards (GTFS)
- GTFS Data Platforms & Tools
- Shared Mobility data (GBFS)
Next Stop: GTFS-schedule
GTFS-schedule

---

For: **Planned Service**

Metro releases Bus & Rail as separate GTFS datasets!

- **Rail** - [https://gitlab.com/LACMTA/gtfs_rail](https://gitlab.com/LACMTA/gtfs_rail)
- **Bus** - [https://gitlab.com/LACMTA/gtfs_bus](https://gitlab.com/LACMTA/gtfs_bus)


GTFS Update Frequency

➢ **Rail** - updated every weekday night!

➢ **Bus** - Updated during **Shakeups**. Usually twice a year (June, December). Minor updates rolled out weekly.

**Shakeups:**

➢ 2x a year in June, December

➢ Major service changes - major changes to routes and stops

➢ Bus operators’ chance to bid on routes, which is why major changes don’t happen outside of Shakeups
What does GTFS-schedule data look like?

---

Zip File

gtfs.zip
Inside is a series of `.txt` files:

- Agency.txt
- Calendar.txt
  - and/or -
  Calendar_dates.txt
- Routes.txt
- Stop_times.txt
- Stops.txt
- Trips.txt

And more...

- Shapes.txt
- Fare_attributes.txt
- Fare_rules.txt
- Timeframes.txt
- Pathways.txt
- Transfers.txt
- Frequencies.txt
- Levels.txt
- Feed_info.txt
- Translations.txt
- Stop_areas.txt
- Areas.txt
Metro releases Bus and Rail GTFS separately!

Primary key: `agency_id` (not required in other files if only 1 agency)

Tells you what agencies are in this GTFS dataset

<table>
<thead>
<tr>
<th>agency_id</th>
<th>agency_name</th>
<th>agency_url</th>
<th>agency_timezone</th>
<th>agency_lang</th>
<th>agency_phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>LACMTA</td>
<td>&quot;Metro - Los Angeles&quot;</td>
<td><a href="https://www.metro.net">https://www.metro.net</a></td>
<td>America/Los_Angeles</td>
<td>en</td>
<td>(323)466-3876</td>
</tr>
<tr>
<td>route_id</td>
<td>route_short_name</td>
<td>route_long_name</td>
<td>route_desc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-13172</td>
<td>2</td>
<td>Metro Local Line</td>
<td>Westwood - Exposition Park via Sunset-Alvarado</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-13172</td>
<td>4</td>
<td>Metro Local Line</td>
<td>Downtown LA - Santa Monica via Santa Monica BL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-13172</td>
<td>10/48</td>
<td>Metro Local Line</td>
<td>W Hollywood-Dtwn LA - Avalon Sta via Melrose-Avalon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-13172</td>
<td>14/37</td>
<td>Metro Local Line</td>
<td>Cedars Sinai-Dtwn LA-Wash/Fairfax via Beverly-Adams</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-13172</td>
<td>16</td>
<td>Metro Local Line</td>
<td>Downtown LA - West Hollywood via West 3rd St</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-13172</td>
<td>18</td>
<td>Metro Local Line</td>
<td>Wilshire Western Sta-Montebello via 6th - Whittier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-13172</td>
<td>20</td>
<td>Metro Local Line</td>
<td>Downtown LA - Santa Monica via Wilshire BL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28-13172</td>
<td>28</td>
<td>Metro Local Line</td>
<td>Century City-Dtw LA-Via Olympic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-13172</td>
<td>30</td>
<td>Metro Local Line</td>
<td>Pico Rimpau - Dtwn LA - Little Tokyo Sta via Pico</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33-13172</td>
<td>33</td>
<td>Metro Local Line</td>
<td>Downtown LA - Santa Monica via Venice BL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-13172</td>
<td>40</td>
<td>Metro Local Line</td>
<td>Downtown LA-Sbay Galleria Via King Bl-Hawthorne Bl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45-13172</td>
<td>45</td>
<td>Metro Local Line</td>
<td>Lincoln Heights-Dtwn LA-Harbr Fwy Sta via Broadway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-13172</td>
<td>51</td>
<td>Metro Local Line</td>
<td>Westlake/Kcaret Bl Sta-Dtwn LA Csw Bl Via Avalon</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tells you what routes are in this GTFS dataset.
**Stops.txt**

Primary key: **stop_id**

This list doesn’t represent all possible stops, just the active ones!

<table>
<thead>
<tr>
<th>stop_id</th>
<th>stop_code</th>
<th>stop_name</th>
<th>stop_desc</th>
<th>stop_lat</th>
<th>stop_lon</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Paramount / Slauson</td>
<td></td>
<td>33.973248,33.973248</td>
<td>-118.113113, -118.113113</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Jefferson / 10th</td>
<td></td>
<td>34.025471,34.025471</td>
<td>-118.328402, -118.328402</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>120th / Augustus F Hawkins</td>
<td></td>
<td>33.924696,33.924696</td>
<td>-118.242222, -118.242222</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>120th / Martin Luther King Hospital</td>
<td></td>
<td>33.924505,33.924505</td>
<td>-118.240369, -118.240369</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>15054 Sherman Way</td>
<td></td>
<td>34.201075,34.201075</td>
<td>-118.461953, -118.461953</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>Riverside / Los Feliz</td>
<td></td>
<td>34.116932,34.116932</td>
<td>-118.271486, -118.271486</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>1st / Eastern</td>
<td></td>
<td>34.036881,34.036881</td>
<td>-118.173232, -118.173232</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>Indiana / Olympic</td>
<td></td>
<td>34.018999,34.018999</td>
<td>-118.192361, -118.192361</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
<td>226th / Norwalk</td>
<td></td>
<td>33.823987,33.823987</td>
<td>-118.072255, -118.072255</td>
</tr>
<tr>
<td>29</td>
<td>29</td>
<td>2nd / Santa Monica</td>
<td></td>
<td>34.015028,34.015028</td>
<td>-118.497346, -118.497346</td>
</tr>
<tr>
<td>31</td>
<td>31</td>
<td>3rd / Gardner</td>
<td></td>
<td>34.070240,34.070240</td>
<td>-118.353264, -118.353264</td>
</tr>
<tr>
<td>33</td>
<td>33</td>
<td>Balboa / Foothill West Jog</td>
<td></td>
<td>34.322084,34.322084</td>
<td>-118.495243, -118.495243</td>
</tr>
<tr>
<td>38</td>
<td>38</td>
<td>.5th / Colorado</td>
<td></td>
<td>34.013884,34.013884</td>
<td>-118.490199, -118.490199</td>
</tr>
</tbody>
</table>

Tells you what stops are in this GTFS dataset
### Calendar.txt

**Primary key:** service_id

<table>
<thead>
<tr>
<th>service_id</th>
<th>monday</th>
<th>tuesday</th>
<th>wednesday</th>
<th>thursday</th>
<th>friday</th>
<th>saturday</th>
<th>sunday</th>
<th>start_date</th>
<th>end_date</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC23-D01CAR-3_Sunday</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>20231210</td>
</tr>
<tr>
<td>DEC23-D01CAR-1_Weekday</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D01CAR-2_Saturday</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>20231216</td>
</tr>
<tr>
<td>DEC23-D02CAR-3_Sunday</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20231210</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-KINGSA1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-KINGMB0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-KINGHA0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-SANHA0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-SANMB0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-1_Weekday-SANSA1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D02CAR-2_Saturday</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>20231216</td>
</tr>
<tr>
<td>DEC23-D03CAR-3_Sunday</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20231210</td>
</tr>
<tr>
<td>DEC23-D03CAR-1_Weekday</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
<tr>
<td>DEC23-D03CAR-1_Weekday-LINCHS1-1011100</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20231211</td>
</tr>
</tbody>
</table>

**Tells you the dates that service runs**

**Metro separates service into Weekday, Saturday, and Sunday!**
Trips.txt

Tells you the trips that run on each route

Primary key: trip_id

A trip = 2+ stops during a specific time period, traveled by a single vehicle. The operator switch out for mid-trip relief.
### Stop_Times.txt

This file is over 200MB unzipped for Metro’s bus GTFS!

<table>
<thead>
<tr>
<th>trip_id</th>
<th>arrival_time</th>
<th>departure_time</th>
<th>stop_id</th>
<th>stop_sequence</th>
<th>stop_headsign</th>
<th>pickup_type</th>
<th>drop_off_type</th>
<th>route_code</th>
<th>destination_code</th>
<th>timepoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>10010007510530-DEC23</td>
<td>06:01:00</td>
<td>06:01:00</td>
<td>11643</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:02:00</td>
<td>06:02:00</td>
<td>11646</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:03:00</td>
<td>06:03:00</td>
<td>4800006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:04:00</td>
<td>06:04:00</td>
<td>4800001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:06:00</td>
<td>06:06:00</td>
<td>4800002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:07:00</td>
<td>06:07:00</td>
<td>3121</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:08:00</td>
<td>06:08:00</td>
<td>5279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:10:00</td>
<td>06:10:00</td>
<td>1131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:11:00</td>
<td>06:11:00</td>
<td>5292</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:13:00</td>
<td>06:13:00</td>
<td>15778</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
<tr>
<td>10010007510530-DEC23</td>
<td>06:13:00</td>
<td>06:13:00</td>
<td>15788</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>Melrose - Vine</td>
<td>0</td>
</tr>
</tbody>
</table>

### Description

Stop_Times.txt tells you the times a vehicle arrives at and departs from, for each stop on each trip.
Relationships, Primary Keys

calendar

agency

routes

stops

trips

stop_times

agency_id

route_id

stop_id

trip_id

trip_id, stop_sequence
A few things to be aware of...
### GTFS Gotchas

---

**In Bus stop_times.txt:**

<table>
<thead>
<tr>
<th>pickup_type</th>
<th>Enum</th>
<th>Optional</th>
<th>Indicates pickup method. Valid options are:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 or empty - Regularly scheduled pickup.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 - No pickup available.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 - Must phone agency to arrange pickup.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 - Must coordinate with driver to arrange pickup.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>drop_off_type</th>
<th>Enum</th>
<th>Optional</th>
<th>Indicates drop off method. Valid options are:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0 or empty - Regularly scheduled drop off.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 - No drop off available.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 - Must phone agency to arrange drop off.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 - Must coordinate with driver to arrange drop off.</td>
</tr>
</tbody>
</table>
GTFS Gotchas

---

In Rail `stops.txt`:

<table>
<thead>
<tr>
<th>stop_id</th>
<th>stop_code</th>
<th>stop_name</th>
<th>location_type</th>
<th>parent_station</th>
</tr>
</thead>
<tbody>
<tr>
<td>80122</td>
<td>80122</td>
<td>7th Street / Metro Center Station - Metro A &amp; E Lines</td>
<td>0</td>
<td>80122S</td>
</tr>
<tr>
<td>80122S</td>
<td>80122S</td>
<td>7th Street / Metro Center Station</td>
<td>0</td>
<td>80122S</td>
</tr>
<tr>
<td>80122A</td>
<td>80122A</td>
<td>7th Street / Metro Center Station - 7th &amp; Figueroa Elevator</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80122B</td>
<td>80122B</td>
<td>7th Street / Metro Center Station - 7th &amp; Figueroa Entrance</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80122C</td>
<td>80122C</td>
<td>7th Street / Metro Center Station - 7th &amp; Flower Elevator</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80122D</td>
<td>80122D</td>
<td>7th Street / Metro Center Station - 7th &amp; Flower Entrance</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80122F</td>
<td>80122F</td>
<td>7th Street / Metro Center Station - Elevator at the Bloc</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80122G</td>
<td>80122G</td>
<td>7th Street / Metro Center Station - Entrance to the Bloc</td>
<td>2</td>
<td>80122S</td>
</tr>
<tr>
<td>80211</td>
<td>80211</td>
<td>7th Street / Metro Center Station - Metro B &amp; D Lines</td>
<td>0</td>
<td>80122S</td>
</tr>
</tbody>
</table>

Metro’s data currently does not have separate points for each side of the platform.
Metro’s Bus GTFS - IDs Change

- - -

➢ route_id and trip_id are not persistent between Shakeups!
Metro’s Bus GTFS - IDs Change

---

In `routes.txt`:

<table>
<thead>
<tr>
<th></th>
<th>route_id</th>
<th>route_short_name</th>
<th>route_long_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2-13172</td>
<td>2</td>
<td>Metro Local Line</td>
</tr>
<tr>
<td>2</td>
<td>4-13172</td>
<td>4</td>
<td>Metro Local Line</td>
</tr>
<tr>
<td>3</td>
<td>10-13172</td>
<td>10/48</td>
<td>Metro Local Line</td>
</tr>
<tr>
<td>4</td>
<td>14-13172</td>
<td>14/37</td>
<td>Metro Local Line</td>
</tr>
</tbody>
</table>

In `trips.txt`:

<table>
<thead>
<tr>
<th></th>
<th>trip_id</th>
<th>arrival_time</th>
<th>departure_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10010007510530-DEC23</td>
<td>06:01:00</td>
<td>06:01:00</td>
</tr>
<tr>
<td>2</td>
<td>10010007510530-DEC23</td>
<td>06:02:00</td>
<td>06:02:00</td>
</tr>
<tr>
<td>3</td>
<td>10010007510530-DEC23</td>
<td>06:03:00</td>
<td>06:03:00</td>
</tr>
<tr>
<td>4</td>
<td>10010007510530-DEC23</td>
<td>06:03:00</td>
<td>06:03:00</td>
</tr>
</tbody>
</table>
Metro’s Bus GTFS - **IDs Change**

---

**Permanent trip_id**

100100007510530-DEC23

- **Day Type**: 1
  - Weekday = 1
  - Saturday = 6
  - Sunday = 7
- **Pattern**: 00751
  - 5 Digits with Leading Zeros
- **Line**: 0010
  - 4 Digits with Leading Zeros
- **Start Time**: 0530
  - HHMM, 36-hour clock
- **Shakeup**: DEC23
  - Month + Year
Some lines are “parent-child” routes and may not have their own route_id.
Metro’s Bus GTFS - Parent-Child Routes

➢ This is why we use `stop_headsign` in `stop_times.txt` instead of `trip_headsign` in `trips.txt`:
Metro’s Bus GTFS - Parent-Child Routes

---

In `stop_times.txt`:

➢ Non-standard `route_code`, `destination_code` fields.

<table>
<thead>
<tr>
<th>trip_id</th>
<th>arrival_time</th>
<th>departure_time</th>
<th>stop_id</th>
<th>stop_sequence</th>
<th>stop_headsign</th>
<th>pickup_type</th>
<th>drop_off_type</th>
<th>route_code</th>
<th>destination_code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1010007510530-DEC23</td>
<td>06:01:00</td>
<td>06:01:00</td>
<td>11643</td>
<td>39</td>
<td>48 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:02:00</td>
<td>06:02:00</td>
<td>11646</td>
<td>40</td>
<td>48 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:03:00</td>
<td>06:03:00</td>
<td>4800006</td>
<td>41</td>
<td>48 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:04:00</td>
<td>06:04:00</td>
<td>4800001</td>
<td>42</td>
<td>48 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:06:00</td>
<td>06:06:00</td>
<td>4800002</td>
<td>43</td>
<td>48 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:07:00</td>
<td>06:07:00</td>
<td>3121</td>
<td>44</td>
<td>10 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:08:00</td>
<td>06:08:00</td>
<td>5279</td>
<td>45</td>
<td>10 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:10:00</td>
<td>06:10:00</td>
<td>1131</td>
<td>46</td>
<td>10 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:11:00</td>
<td>06:11:00</td>
<td>5292</td>
<td>47</td>
<td>10 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Melrose - Vine</td>
</tr>
<tr>
<td>1010007510530-DEC23</td>
<td>06:13:00</td>
<td>06:13:00</td>
<td>15778</td>
<td>48</td>
<td>10 - Melrose - Vine</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Melrose - Vine</td>
</tr>
</tbody>
</table>
Metro’s Bus GTFS - Parent-Child Routes

---

Why do we do this?

➢ Some trips change route number mid-trip to avoid ending the trip where the route changes.

Why?

➢ At the end of a trip, the bus operator needs to be able to layover and access a bathroom (contractually, legally, and biologically!).
  ○ Some cities won’t agree to let us layover on their streets.
  ○ We negotiate contractual agreements with businesses, but sometimes we can’t find one that will allow our bus operators to use their bathroom.
  ○ Sometimes we have to build our own bathroom!
Next Stop: GTFS-realtime
GTFS-realtime

- VehicleLocation
  Live vehicle locations

- TripUpdates
  Arrival predictions, delays, cancellations

- Alerts
  Service disruption descriptions

Some apps show you when their predictions use realtime vs schedule data!
Detour!
Realtime Tracking - Bus

---

Two systems:

➢ **ATMS** - Advanced Transportation Management System
  - AVL - Automatic Vehicle Location
  - Transmits data via radio
  - Updates every 2–3 minutes
  - Fully deployed 2006

➢ **WiFi Routers**
  - Transmits data via cell towers
  - Updates every 2–5 seconds
  - Fully deployed 2021
Realtime Tracking - Rail

➢ Track Circuits
Realtime Tracking - Rail

➢ Supervisory Control and Data Acquisition (SCADA) system
  ○ Controls and monitors processes like train speed, switches, signals, etc.
GTFS-realtime

Format:

➢ **Protobuf** (Protocol Buffer) – a binary format created by Google to be more efficient when transferring lots of data.

➢ GTFS-realtime often provided as **JSON** as well.
No trip assigned?

➢ Our Wifi Routers send location as long as the bus is powered.
➢ Vehicle may be traveling or waiting between trips.
➢ Vehicle could be replacing another vehicle without the trip being reassigned.
➢ ATMS might be down.

Remember: Trip tells us what times a vehicle arrives and departs what sequence of stops! This is essential for predictions.
End of the Line: Where do we go from here?
GTFS is Constantly Growing!

---

Specification amendment process for:

- **Bus**
  
  [https://gtfs.org/schedule/process/](https://gtfs.org/schedule/process/)

- **Rail**
  
  [https://gtfs.org/realtime/process/](https://gtfs.org/realtime/process/)
GTFS Governance

---

Broadly speaking, new extensions require:

➢ 1 data producer
➢ 1 data consumer
➢ A production implementation
➢ A series of public votes on GitHub
GTFS Extensions - TripModifications

Detour shapes in GTFS-realtime! Voting closed 3/7!

https://github.com/google/transit/pull/403

Following the proposal of [GTFS Trip-Modifications in February](https://github.com/google/transit/pull/403), we are now ready for the pull request stage!

As a reminder, Trip-Modifications are modifications done to a trip to modify its shape, remove stops that are not served anymore, and potentially add temporary stops. Trip-Modifications is mainly used in a detour use case.

Today, we’re [launching the detour feature](https://github.com/google/transit/pull/403) in Transit for more agencies with Trip Modifications produced by Swiftly! Now that there is one producer and one consumer, the current PR is not an experimental field and actually removes the experimental tag for `shape`, which has been in the spec for 2 years.

Feel free to ask questions, or give feedback!
GTFS Extensions - Flex

Non-fixed route, on-demand, “flexible” transit.

September 2023 discussion reached consensus on using GeoJSON for zones!

https://github.com/google/transit/pull/433
GTFS Extensions - Fares v2

Base implementation of v2 passed vote in May 2022!
https://github.com/google/transit/pull/286

Development continues!
MobilityData
MobilityData is Active!

➢ Hiring
https://careers.mobilitydata.org/

➢ As of March 4, also home to:
  ○ Mobility Data Interoperability Principles (MDIP)
  ○ Transit ITS Data Exchange Specification (TIDES)
  ○ Transit Operational Data Specification (TODS)

➢ 2nd International Mobility Data Summit
https://mobilitydata.org/international-mobility-data-summit/
Follow what’s going on via:

- **Newsletter**
  [https://mobilitydata.org/contact-us/](https://mobilitydata.org/contact-us/)

- **GitHub**
  [https://github.com/mobilitydata](https://github.com/mobilitydata)

- **Slack**
Cal-ITP (Caltrans)

https://www.calitp.org/

Cal-ITP has 4 goals for their work:

1. invites travelers toward more efficient and sustainable modes of travel,
2. reduces operating costs for transit agencies,
3. simplifies every traveler’s life, and
4. simultaneously welcomes every single person in this country to a seamless, user-friendly system—making paying for a transit ride just as easy as buying coffee.
What does that translate to?

➢ **Contactless Payments**
  Making it easier for transit agencies to implement contactless payments so riders can tap with the debit card, credit card, or smartphone already in their pocket!

➢ **Benefits**
  Automating customer discounts so riders instantly qualify for and receive discounts

➢ **GTFS**
  Standardizing information for easy trip planning by expanding and beefing up the GTFS ecosystem
Mobility Data Interoperability Principles (MDIP)
MDIP

---

https://www.interoperablemobility.org/

➢ Effort started by Cal-ITP.
➢ Grown into an international coalition.
➢ A declaration that the industry needs interoperable data to advance.
➢ That means transit agency and transit vendor buy-in.
➢ That means creating a plan and providing resources.
➢ Recently rehomed to MobilityData to continue the work.
Metro
What is Metro doing?

Cal-ITP (CalTrans)
➢ Metro is a partner!

Mobility Data Interoperability Principles (MDIP)
➢ Metro is a co-author!
What is Metro doing?

---

**GTFS Extensions**

- Looking for partnership opportunities, actively voting on the new extensions being developed.

**Mobile App Consolidation Working Group**

- Metro is pushing for GTFS-schedule and GTFS-realtime data to exist regionally
- Metro is advocating for expanding data standards in the industry.
More Resources for your GTFS Journey
Creating and Using GTFS

➢ Intro to GTFS Video
   https://www.youtube.com/watch?v=SDz2460AjNo

➢ GTFS Best Practices
   https://gtfs.org/schedule/best-practices/
   https://gtfs.org/realtime/best-practices/

➢ California Transit Data Guidelines
GTFS Catalogs

- Cal-ITP – Monthly GTFS Quality Reports (California)
  https://reports.calitp.org/
- MobilityData – Mobility Database
  https://mobilitydatabase.org/
- Interline – Transitland
  https://www.transit.land/
GTFS Tools

➢ **GTFS Studio**
https://editor.gtfs.studio/

➢ **GTFS-schedule Validator**
https://gtfs-validator.mobilitydata.org/

➢ **Software libraries in different languages**
https://gtfs.org/resources/gtfs/
Open Source Projects/Apps that use GTFS data

➢ OneBusAway
   https://onebusaway.org/

➢ OpenTripPlanner
   https://www.opentripplanner.org/

➢ TheTransitClock
   https://thetransitclock.github.io/

➢ Shoutout to local project: Catenary Maps
   https://catenarymaps.org/
Metro’s Data

➢ Metro Open Data
   https://developer.metro.net/

➢ Metro Bus GTFS
   https://gitlab.com/LACMTA/gtfs_bus

➢ Metro Rail GTFS
   https://gitlab.com/LACMTA/gtfs_rail
Thank you to...

---

➢ **Rollin Baker**  
  Director - System Projects (HASTUS)  
  LA Metro, Operations

➢ **Monica Waggoner**  
  Principal Transportation Planner - Gateway Cities  
  LA Metro, Operations

➢ **Isabelle de Robert**  
  Director of Product  
  MobilityData
Reach out!

LinkedIn - www.linkedin.com/in/ninakin
Email - kinn@metro.net
References

- https://thecityfix.com/blog/introducing-google-transit/
- https://azavea.gitbooks.io/open-data-standards/content/standards/domain_specific_standards/gtfs_realtime_gtfsrt.html
References

➢ https://trilliumtransit.com/2015/12/10/gtfs-today-and-tomorrow/
➢ https://groups.google.com/g/gtfs-changes/c/ob_7MIOvOxU/m/zEScjv6VLMJ?pli=1
➢ https://rmi.org/data-interoperability-enables-mobility-service/
➢ https://rmi.org/transit-directions-just-got-little-bit-better/
Attributions

➢ “Electricity Energy” icon by Kharisma Mulyana from thenounproject.com CC BY 3.0
➢ “Wifi” icon by Adrien Coquet from thenounproject.com CC BY 3.0
➢ “Occupyed track circuit” diagram / Users: Mangoe, Completefailure / Wikimedia Commons / CC BY-SA 3.0
➢ “Smartphone wifi connection” icon by Royyan Wijaya from thenounproject.com CC BY 3.0