Transport for NSW

API Gateway

GTFS Release Notes

CONTENTS

AUTHOR:	TfNSW Data Services Team (DCIS)
DATE:	17 April 2018
VERSION:	4.0
DIVISION:	CED

Revision History

Version	Author	Issue Date	Changes
3.3	TfNSW API Gateway	29 Jun 2015	Public release
4.0, 4.1	TfNSW Data Services Team (DCIS)	17-20 April 2018	Amended routes.txt > route_type for extended route types. Amend Temporary Bus code

1. OVERVIEW

The TfNSW Transport Data Exchange program provides public transport data in the GTFS (General Transit Feed Specification) format to all program subscribers.

While GTFS is well documented, it provides several optional fields and some flexibility in how to populate some fields. This document provides some general information regarding how the NSW transport data is represented in GTFS.

This GTFS file contains data for:

- All Sydney Trains and NSW Trains services
- All ferry services in Greater Sydney and regional NSW
- All bus services in Greater Sydney and regional NSW
- All light rail services in the greater Sydney region

2. GENERAL INFORMATION

2.1   GTF Specification Compliance

With the exception of a documented Google Transit extension field, the data is compliant with the GTF specification,

(https://developers.google.com/transit/gtfs/reference) as specified at 2 February 2015.

2.2   Publication Schedule

The data is scheduled to be published on a weekly basis by 5pm each Friday. However there may be instances where publication is delayed due to scheduled maintenance or unforeseen circumstances.

2.3   GTFS currency versus TransportNSW.info

As Transport regularly receives timetable amendments, www.transportnsw.info may present slightly different data compared with the current published GTFS Feed. As there are frequent amendments, it is recommended that participants regularly update their data with the newest available GTFS Feed.

2.4   Data Scope

2.4.1 Geographic coverage

The data is a representation of the public transport data currently utilised within TransportNSW.info and will cover the same geographic region.
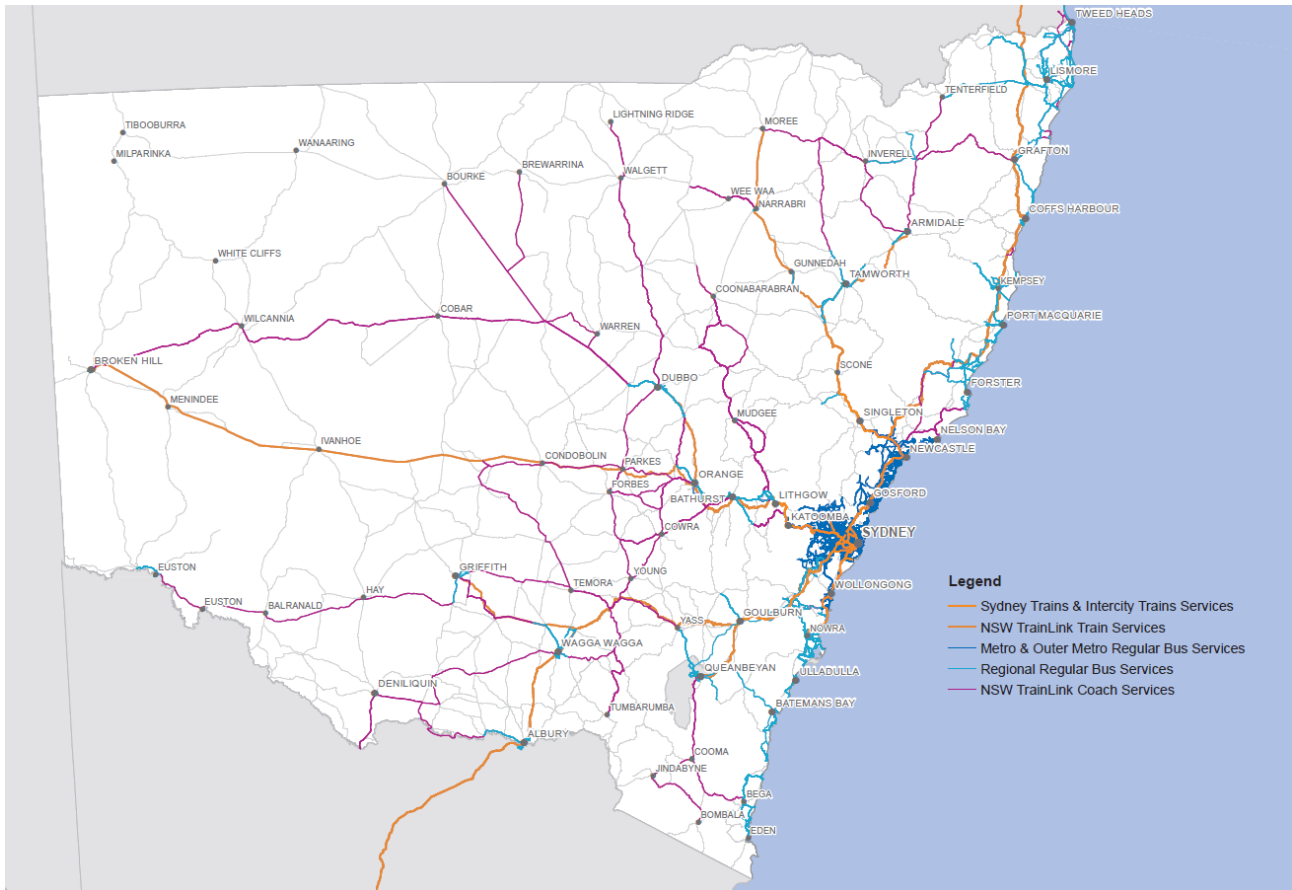
2.4.2   Validity Period (Date ranges)

Data will be exported with up to 90 days of forward data (from the date of export).

Some operators may have a lesser validity period as service information is not available at the time of export. As this information becomes available to TransportNSW.info it will also be published within the Transport Data Exchange exports.

2.5   Identifiers

Many of the identifiers used within the feed will include characters like plus (+), underscore (_), hyphen (-) or full stop (.). Care may need to be taken to consider this when using the data or developing applications. Identifiers may not be consistent across datasets – ie the same journey may be represented under a different trip_id in a different dataset.

2.6   Value Quoting

All values within the files are enclosed in double-quotes (e.g. ”X”).

2.7   Feed Size

The feed size will vary depending on the volume of data held by Transport at the time of export. TransportNSW.info typically utilises more data during special events, timetable changes or holidays periods - hence the data volume will typically increase around these events. This may alter the feed size considerably – compressed sizes typically range from 110Mb – 240Mb. Sizes may increase beyond these typical ranges and it is suggested to consider or plan for this scenario when utilising the data.

2.8   Transit Paths / Shapes

Included within the feed (shapes.txt) are the inter-stop paths that the vehicle is expected to follow – this is The Shapes.txt section of this document should be reviewed for considerations when utilising this information.

2.9   GTF Data Validation

The data contained within the feed has been validated using the freely available FeedValidator tool from the googletransitdatafeed project. At time of writing, FeedValidator 1.2.16 does not consider the several optional fields to be part of the primary specification, hence it is recommended to utilise the GoogleTransit extension module as part of any validation utilising FeedValidator.

3.   GTFS RELEASES – ARCHIVE FILE CONTENT

The GTFS data is provided via 8 of the specified files: agency.txt, calendar.txt, calendar_dates.txt, routes.txt, trips.txt, stops.txt, stop_times.txt & shapes.txt. Please see below for notes on individual files and fields.

4.   GTFS FILE: AGENCY.TXT

The following fields are supplied & populated within the agency.txt file:

agency_id,agency_name,agency_url,agency_timezone,agency_lang,agency_phone

4.1   Field notes - agency_url

This is always populated as http://transportnsw.info rather than individual transport operator sites.

4.2   Field notes - agency_timezone

This is currently populated at 'Australia/Sydney' for all operators included within the feed.

4.3   Field notes - agency_phone

This is always populated as 131500 rather than individual transport operator phone numbers

5.   GTFS FILE: CALENDAR.TXT

The following fields are supplied and populated within the Calendar.txt file:

service_id,monday,tuesday,wednesday,thursday,friday,saturday,sunday,start_date,end_date

5.1   Calendar & Calendar_dates

The calendar.txt file must be used in conjunction with calendar_dates.txt to obtain an accurate representation of the days that services operate. Depending on individual data suppliers, some calendars heavily feature date inclusions or exclusions. For example some calendars may have all Monday to Friday fields marked as active with a significant number of these activations overridden by exclusions in calendar_dates.txt. Note that any ex/inclusion within calendar_dates.txt overrides calendar.txt definitions for the specified date.

5.2   Start & End Dates

As per Data Scope validity period, the export is based on a 90 day period. Many Start & End dates will reflect this period by being valid for the entire period. However there will be calendars with shorter validity periods that start in the future or end earlier. In general these will relate to change in schedules (e.g. a timetable amendment).

6.   GTFS FILE: CALENDAR_DATES.TXT

The following fields are supplied and populated with the Calendar_dates.txt file:

service_id,date,exception_type

7. GTFS FILE: ROUTES.TXT

The following fields are supplied & populated within the routes.txt file:

route_id,agency_id,route_short_name,route_long_name,route_desc,route_type,route_color,
route_text_color.

7.1 Field notes - route_short_name

This field is populated differently depending on the transit mode:

- Bus: uses the route number (e.g. '400', '610X')
- Rail: not populated (null)
- Ferry: not populated (null)
- Light Rail: not populated (null)

7.2 Field notes - route_long_name

This field is populated differently depending on the transit mode:

- Bus: This is populated with a route description. Note that this is a directional description (e.g. 'Bondi to City')
- Train: This is populated with the line name (e.g. 'South Coast')
- Ferry: This is populated with the route name (e.g. 'Eastern Suburbs')
- Light Rail: This is populated with the line name (e.g. 'L1 Lilyfield line')

7.3 Field notes - route_desc

This field is populated with the descriptive Network name (e.g. 'Regional Trains and Coaches Network') applicable for the route. This is intended to assist passengers with identifying and utilising public transport.

route_desc	Line colour	Line Hex code	Logo
Sydney Trains Network	Train orange	#F6891F	T
Intercity Trains Network	Train orange	#F6891F	T
Regional Trains and Coaches Network	Train orange	#F6891F	T
Regional Trains and Coaches Network	Coach purple	#732A82	C
Blue Mountains Buses Network	Bus blue	#00B5EF	B
Sydney Buses Network	Bus blue	#00B5EF	B
Central Coast Buses Network	Bus blue	#00B5EF	B
Hunter Buses Network	Bus blue	#00B5EF	B
Illawarra Buses Network	Bus blue	#00B5EF	B
Temporary buses	Bus blue	#00B5EF	B
Private bus services	Bus blue	#00B5EF	🚌
Sydney Ferries Network	Ferry green	#5AB031	F
Newcastle Ferries	Ferry green	#5AB031	F
Private ferry and fast ferry services	Ferry green	#5AB031	⛴
Sydney Light Rail Network	Light Rail red	#EE343F	L

7.4 Field notes - route_type

The route_type field describes the type of transportation used on a route.

ℹ️ UPDATE (May 2018)

Prior to May-2018, standard GTFS route types (0 -> 7) were implemented.

From May-2018 onwards, extended GTFS route types are implemented for certain services which provides a greater level of granularity for data classification.

Reference: Extended GTFS Route Types
https://developers.google.com/transit/gtfs/reference/extended-route-types

---

Diva Transport Id	Route Description	GTFS Route Type (prior to May 2018)	Extended GTFS Route Type (from May 2018)
4	Blue Mountains Buses Network	3	700
5	Sydney Buses Network	3	700
6	Central Coast Buses Network	3	700

Diva Transport Id	Route Description	GTFS Route Type (prior to May 2018)	Extended GTFS Route Type (from May 2018)
9	Private bus services	3	700
15	Hunter Buses Network	3	700
16	Illawarra Buses Network	3	700
17	Private coach services	3	700
8	School Buses	3	712
1	Sydney Trains Network	2	2
2	Intercity Trains Network	2	2
3	Regional Trains and Coaches Network	2	106
19	Temporary trains	2	2
7	Regional Trains and Coaches Network	3	204
22	Temporary coaches	3	203
13	Sydney Light Rail Network	0	0
20	Temporary light rail	0	0
21	Newcastle Light Rail	0	0
23	On Demand	3	715
14	Temporary Buses	3	714
11	Newcastle Ferries	4	4
12	Private ferry and fast ferry services	4	4
98	Sydney Ferry	4	4
18	Temporary ferries	4	4
10	Sydney Ferries Network	4	4

7.5   Field notes - route_color

This field is populated for all routes with the relevant colours utilised on timetables and signboards.

7.6   Field notes - route_text_color

This field is populated for all routes and is generally either black or white, selected for contrast with the route_color.

8. GTFS FILE: TRIPS.TXT

The following fields are supplied & populated within the trips.txt file:

route_id,service_id,trip_id,shape_id,trip_headsign,direction_id,block_id,wheelchair_accessible

8.1 Field notes -Trip_headsign

The trip headsign is populated slightly differently depending on the transit mode and trip behaviour:

- Bus: In the majority of cases the trip_headsign will be the suburb within which the trip finishes (e.g. 'Chatswood'). However should the trip be a loop service (starting and ending in the same suburb), the relevant directional description for the route will be utilised (e.g. 'Chatswood Loop via Killara').
- Train: This is populated with the name of the final station on the Journey e.g. 'Central Station'. Please note additional information is populated within the stop_times.txt file (stop_headsign) when trips are identified as loops. Refer to GTF File: Stop_times.txt for further information.
- Ferry: In the majority of cases the trip_headsign will the name of the final Wharf on the trip (e.g. 'Circular Quay'). However should the trip be a loop service (starting and ending in the same suburb), the trip_headsign will show the next stop in the trip.
- Please note additional information is populated within the stop_times.txt file (stop_headsign) when trips are identified as loops.
- Light Rail: This is populated with the name of the final station on the Journey e.g. 'Central Station Light Rail'.

8.2 Field notes - block_id

Block information is populated for selected services – primarily Rail. This aids consumers in 'connecting' trips which are suitable for passengers to continue travel on without disembarking. It essentially links trips performed by the same vehicle.

8.3 Field notes – wheelchair_accessible

Accessibility information is populated for all services, with flag values of '1' (accessible) or '2' (non accessible) being used. This information is based upon the vehicle type being used for the service. Please note:

- this flag indicates whether an entire trip is accessible and does not indicate whether a trip is accessible at a particular stop along the journey, (see wheelchair_boarding in stops.txt) and
- this information is based upon planned schedules. Changed operational circumstances may mean that some journeys are operated by a different vehicle type and hence access level may be changed.

9. GTFS FILE: STOPS.TXT

The following fields are supplied & populated within the stops.txt:

stop_id,stop_code,stop_name,stop_lat,stop_lon,location_type,parent_station,wheelchair_boarding,platform_code

The optional Parent Station fields are utilised within this file to group Rail platforms together. These Parent Stations can be ignored if not relevant to individual applications.

The optional wheelchair_boarding field is used to indicate accessibility for certain stop types.

9.1   Field notes - stop_id

As per the notes regarding identifiers, the stop_id field may include special characters such as underscore (_).

There are two general stop_id 'types' within the data:

- Transit Stop Numbers (TSN's) – e.g. '200035'. These are used for Buses, Ferries, Light Rail and Rail Platforms
- Parent Station identifiers (e.g. 'PST1234') – used to group Rail Platforms

9.2   Field notes – stop_code

The stop_code field is populated for stops that utilise a Transit Stop Number (TSN), most commonly bus stops and ferry wharves. This number is used within many NSW transport systems to uniquely identify a stopping location for passengers.

Where the stop does not have a TSN, this field will not be populated.

9.3   Field notes - location_type

This field is only populated for stops which are Parent Stations – all other stops will be null.

9.4   Field notes - parent_station

This field will only be populated for Sydney Trains and NSW TrainLink platforms, where it will be related to the appropriate Parent Station. All other stops will be null.

9.5   Field notes – wheelchair_boarding

This field is populated differently depending on the transit mode that utilises the stop point
and stop type (parent or child).

- Bus stops are always '0' (unknown)

- Train Platforms are '1' (accessible) where TransportNSW.info considers the platform to be accessible and able to be reached from the street in an accessible manner (e.g. flat street entrance or lift is available). Where platforms are not specifically considered to be accessible, they are marked as '2' (inaccessible). Note that some platforms marked as inaccessible may be accessible with assistance or by specific services, but accessibility cannot be guaranteed.
- Ferry Wharves are '1' (accessible) where TransportNSW.info considers the wharf to be accessible.
- Where wharves are not considered to be accessible, they are marked as '2' (inaccessible).
- Parent Stations are marked dependent on their child stops:
  - Where all child stops are wheelchair accessible ('1'), the parent station will be marked accessible.
  - In all other cases the parent station will be marked as '0' (unknown) and consumers should evaluate the child stop values to determine the accessibility of a particular journey.

9.6   Field notes – platform_code

This field will be populated for Train platforms that have a designated platform 'number' (e.g. '15' if the stop is 'Central Station Platform 15'). It should be noted that this field is alphanumeric and may include words (e.g. 'Sprint' for 'Lidcombe Station Sprint Platform'). This field will be null for stops without platform designations, Bus stops and ferry wharves.

10. GTFS FILE: STOP_TIMES.TXT

The following fields are supplied & populated within the stop_times.txt file:

trip_id,arrival_time,departure_time,stop_id,stop_sequence,stop_headsign,pickup_type,drop_off_type, shape_dist_traveled

10.1   Field notes – arrival_time

As per the GTF specification, arrival_times may be outside the conventional 24 hour clock values (e.g. '25:14' being 1:14am the following day). Care needs to be taken to handle these values appropriately.

Note that in some cases (due to supplier constraints) this field will not be populated. This will occur where otherwise the current stop would have the same arrival_time as the previous stop on the trip.

As per the GTF specification, where the arrival_time field is not populated, consumers should estimate (interpolate) the value based on previous and subsequent populated values. This may be done by either simple time division or by making more complex estimates considering transit distance by leveraging the shape_dist_traveled field.

This will not occur for either the first or last stop on a trip.

10.2   Field notes – departure_time

As per the GTF specification, arrival_times may be outside the conventional 24 hour clock values (e.g. '25:14' being 1:14am the following day). Care needs to be taken to handle these values appropriately.

Note that in some cases (due to supplier constraints) this field will not be populated. This will occur where otherwise the current stop would have the same arrival_time as the previous stop on the trip.

As per the GTF specification, where the arrival_time field is not populated, consumers should estimate (interpolate) the value based on previous and subsequent populated values. This may be done by either simple time division or by making more complex estimates considering transit distance by leveraging the shape_dist_traveled field.

This will not occur for either the first or last stop on a trip.

10.3   Field notes – stop_headsign

The stop_headsign field is used to provide more accurate destination information where the trip behaviour is such that traveller confusion is likely. An example of this is a loop service where the origin and destination is the same.

The stop_headsign field will be populated when the following conditions are met:

- The route is operated by a train or ferry
- The trip is determined to be a loop service

For trains, a loop trip is considered to occur where the origin and destination is at the same parent station (not necessarily the same platform).

For ferries, a loop trip is considered to occur where the origin and destination stop are within the same TransportNSW.info locality.

Where the stop_headsign field is populated, the field will be populated with the general name of the following stop on the trip. The exception to this is the final stop on the trip, where the stop_headsign is blank. For example in a three stop train journey from Central to Parramatta via Strathfield, the stop_headsign values would be as follows:

Sequence	Stop	stop_headsign
1	Central Station	Strathfield Station
2	Strathfield Station	Parramatta Station

3	Parramatta Station	<null>

11. GTFS FILE: SHAPES.TXT

The following fields are supplied & populated within the shapes.txt file:

shape_id,shape_pt_lat,shape_pt_lon,shape_pt_sequence,shape_dist_traveled

11.1   Spatial differences due to geographic data

The shapes portrayed via the shapes.txt file are a representation of the path taken by vehicles on their journeys. These paths are based around the geographic data utilised by Transport. In some cases the paths supplied may not align with other geographic data or imagery.

In cases where there is insufficient geographic data, there may be instances where shapes are not available.

11.2   Path accuracy

The path information provided is generated based on a mix of automatic algorithm and manual pathing. There may be inaccuracies with the paths and hence the paths may not always match on-road operations.