



City of Albuquerque Information Technology Services Division Data Management

Bus Location and Direction Metadata Requirements

Contact Information

Who is the contact for this dataset? The contact will be the City employee who is accountable for the data provided in this dataset and can act as front-line support in the event of any questions about the data.

Name	Joe Saraphon
Department/Division	Transit
Phone	505-724-3113
Email	Saraphon@cabq.gov

What Does this Dataset Describe?

What is the name of this dataset? How should a user identify this dataset in any communication with contact above? Provide a shorter description of the Dataset that can act as a one-line summary of the dataset when dealing with stakeholders. Provide a longer description of the data that can be readily understood by non-technical users.

Dataset Title	RioMetro buses/vans and RailRunner location and direction
Short Description	Real-time RioMetro buses and vans and RailRunner location and direction
Full Non-Technical Description	
RailRunner.kml; Data is updated between 5 to 15 seconds with the vehicle #, speed, date, time, longitude and latitude of the RailRunner that was at that location. Update interval is subject to change.	
RioMetro; Data is updated between 30 to 180 seconds with the vehicle #, speed, date, time, longitude and latitude of the RailRunner that was at that location. Update interval is subject to change.	

How Should this Dataset be Cited?

How should external sources refer to this dataset in publications or documentation? Often this will simply be the URL and the date retrieved.

Each bus route will have the following files, the files will be found in different folders within;

<http://data.cabq.gov/transit/RioMetro/RailRunner.kml>
<http://data.cabq.gov/transit/RioMetro/RioMetro.kml>

Does the Dataset Reflect a Particular Time Period?

Provide any date restrictions that may affect the validity of the data. The table fields are defined as follows:

	<i>Definition</i>
<i>Start Date</i>	<i>Start date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>End Date</i>	<i>End date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Refresh Interval</i>	<i>Time period between Dataset refreshes. Format: “nn [seconds/minutes/hours/days/weeks/months/years]” or the word “Static” if never refreshed.</i>
<i>Data Expiration Date</i>	<i>Date after which the data must be considered stale and no longer of sufficient utility (fit-for-purpose). Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Review Date</i>	<i>Date after which this dataset will be reviewed by the City for utility (fit-for-purpose) and usage. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Comments</i>	<i>Specific comments related to any time-specific features of this dataset.</i>

Start Date	N/A
End Date	N/A
Dataset Refresh Interval	RailRunner.kml Between 5 to 15 seconds RioMetro.kml Between 30 to 180 seconds
Dataset Expiration Date	N/A
Dataset Review Date	N/A
Comments	

The files are in the KML format. The format is dictated by Google at <https://developers.google.com/transit/gtfs-realtime/>

Dataset Definition/Format

Provide a field-by-field breakdown and definition of each record. This section acts as the formal data dictionary for an individual record.

The format is dictated by Google at <https://developers.google.com/transit/gtfs-realtime/>.

Dataset Technical Description

Provide a technical description of the dataset. This should be a complete technical description aimed at developers and expert users who need to understand the scope, strengths and limitations of the dataset.

File name	File description
RailRunner.kml	This file contains the location for the RailRunner locomotives and cars. This file is currently in beta testing and is showing one coach car only.
RioMetro.kml	This file contains the location of Rio Metro's buses and vans. This file is currently in beta testing and is showing vehicles at random intervals.

Dataset Assumptions

What technical and business assumptions are implied in the creation of this dataset? Examples could include the way in which a salary figure was calculated or data that was omitted for a specific reason.

This is a beta file and should not be used for application development. This file may be moved at any time to a new location including a location outside of data.cabq.gov. This data does not belong to the City of Albuquerque. We are assisting MRCOG in testing the possibility of having open data available on the RailRunner and on their buses and vans.

The cabin represented by this file may not be in service at all times.

Who Produced the Dataset?

Which department in the City produced this dataset? Note that this might not always be the data owner. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.

The City of Albuquerque, Transit Department, IT Division is responsible for gathering the data sent from the Sierra Wireless modem aboard each bus, matching it with the Trapeze data from the

Trapeze OPS database, computing direction based on GPS data, and publishing it in the format required by Google. Transit IT is responsible for maintaining the data on data.cabq.gov/transit.

Who Owns the Data?

Where did this data originate? Who owns the data used in this dataset? Note that this might not always be the dataset producer. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.

The City of Albuquerque, Transit Department is the data owner.

Why was the Dataset Created?

All datasets should have an explicit reason for existence and should, somehow, have value to someone. What is the perceived value that this dataset will bring?

The dataset was created to allow applications to map the bus route, the location of the bus on the route and the direction the bus is headed.

How was the Dataset Created?

How was this dataset produced? Was it a manual process? An automated process? What were the main IT systems involved in producing this dataset?

This data is created by an application (written in-house by Joe Saraphon) that reads the Trapeze Transit database and the data being transmitted by the on-board modems and generates a file uploaded to data.cabq.gov/transit.

What Similar or Related Data Should the User be Aware of?

Are there any other datasets available that may contain related or similar information? Might there be situations in which these other datasets might be a better alternative?

Please see the metadata file in each of the dataset folders for information on how to use the information.

Transit Scheduled Route

<http://data.cabq.gov/transit/realtime/route/allroutes.kml>

<http://data.cabq.gov/transit/realtime/route/routeX.kml>

Where X is the route number between 1 and 9.

<http://data.cabq.gov/transit/realtime/route/routeXX.kml>

Where XX is the route number between 10 and 99

<http://data.cabq.gov/transit/realtime/route/routeXXX.kml>

Where XXX is the route number between 100 and 999.

<http://data.cabq.gov/transit/realtime/route/routeXXXX.kml>

Where XXXX is the route number between 1000 and 9999

The files are comma delimited with field names in the first row; the format is dictated by Google at <https://developers.google.com/transit/gtfs/reference>

How Reliable are the Data?

Are there any concerns about overall data reliability? Are there any data problems that the user needs to be aware of? Are there any constraints with data accuracy? What levels of confidence with this dataset could the user reasonably assume?

The data used in this dataset is derived from Transit production operational data located in Trapeze. The Trapeze database is filled nightly with the next day's assignment (bus 705 is going to drive route 66, block 4 on that day) using a manual process. If a bus is removed from a route or a change-out occurs, there is a risk that the Maintenance or Dispatch group may not immediately update Trapeze, a manual process. If this occurs there will be an outage (1 run on 1 route by 1 bus) of information. In the normal course of business, the data is always available.

Data on data.cabq.gov is accessed via the City network. If the City has an outage, then the data will not be available. The City has a 99.5% up-time requirement. The City has scheduled maintenance on Sunday mornings; however it is the goal of the City to not interrupted service on data.cabq.gov.

GPS data is usually accurate within approximately 10 meters. There are occasions when the GPS data is highly inaccurate. These include when the onboard devices are initially powered up and/or when the antenna is under or next to tall metal objects such as Rail Runner stations with canopies.

There are known cellular "dead" spots along the Rail Runner train tracks between Albuquerque and Santa Fe. Data is not available when the Rail Runner encounters these dead spots.

How Well Have the Observations Been Checked?

What quality assurance steps have been performed? Sometimes, a third-party verification/audit process may also be required. If so, provide the name of the third-party who performed the verification.

This data has been verified by the data owner, Transit IT.

Are there Legal Restrictions on the Access or Use of the Data?

Are there any specific legal or compliance restrictions for this data? How might this affect the way in which end users might access and use this data?

There are no legal restrictions on the access or use of the data.

Legal Disclaimer

The City's standard copyright, disclaimers and legal statements may be found at <http://www.cabq.gov/about/legal>. The City data policy governing data.cabq.gov may be found at <http://data.cabq.gov/policy/>.