Open Opal Data – Frequently Asked Questions

What Opal data has been released?

Opal data was released in March 2017. The Opal dataset includes 6 Comma-Separated Values (CSV) files (spreadsheets) containing two non-consecutive weeks (14 days) of Opal data across the four public transport modes (train, bus, ferry and light rail).

What information is being released in the Opal tap on and tap off data?

The CSV files (one for each of the two weeks of Opal data) released in the Opal dataset include:

- Tap on and off time (single attribute data set)
- Tap on and off location (single attribute data set)
- Tap on and off time and location (double attribute data set)

The two weeks of Opal data are:

- 25 July to 31 July 2016 (before paper ticket retirement – paper ticket data is not included in the dataset)
- 8 August to 14 August 2016 (after paper ticket retirement).

It is important to note that any information that can be linked to an individual or to an individual's card has not been released in the dataset. For more information, see *How is Opal customers' privacy being protected?*

What can the Opal data be used for?

Opal data is useful to anyone who has an interest in public transport trips in Greater Sydney. The main users of the dataset will be researchers who can use the data for insights, analytics and informed decision making.

Where can I access the Opal data?

Opal data can be directly downloaded from the Open Data Hub (opendata.transport.nsw.gov.au). The dataset can be found in the Data Catalogue. You need to register and log in before downloading the data.

How is Opal customers' privacy being protected?

In releasing Opal data, customers' privacy is the utmost priority. Algorithms were developed to privacy-protect the Opal dataset. The algorithms used raw Opal data to produce a privacy-protected dataset that ensures no information that can identify an individual was released in the dataset. There is no possible way identify an individual or their personal information using the Opal dataset.

The information protected in the dataset includes:

- Data that is, or can be, linked to an individual's Opal card

- Geographic information that could be used to identify travel patterns of individuals on low volume transport services, such as isolated bus stops used by a small number of passengers.

Is privacy-protected data still valuable?

Privacy-protected data is still valuable as it is proportionally representative of the raw data, even though it has been significantly refined.

Can I stop my Opal data from being released in the future?

Released Opal data is completely anonymised and privacy-protected, so there is no way to identify yourself or anyone else within the data. Transport for NSW cannot identify individuals within the raw data, so your Opal data cannot be removed from future datasets.

You can review the Opal Terms of Use on the Opal website (www.opal.com.au)

Why can't I find my trip in the Opal dataset?

Opal data is privacy-protected using algorithms that ensure no information that can identify an individual was released in the dataset. It is not possible to identify your own trips or anyone else's trips within the data. The data is representative of trips made and does not contain actual trips.

What disclaimers will be included with the Opal data?

The Opal data may differ from expectations due to the train trackwork. The usual case during trackwork is that neither tap on or tap off data is collected for the period for the train stations effected. Customers may still tap on, however it is not necessary as the buses which replace train services during trackwork do not require customers to tap on or off (and may not have an enabled opal card reader).

What other Opal data is available?

Opal data containing information about how full buses are was released in August 2016. The data tells users whether a bus has seating, has standing room or is full. Unlike the Opal dataset, this Opal data updates in real-time and is featured in some mobile applications to give customers the choice of either jumping on or waiting for the next service.

Opal trip data, providing official utilisation figures for all four public transport modes, was released 17 January 2017. The five downloadable datasets, can be found on the Open Data Hub and Developer Portal.

How is the Opal data affected by the retirement of paper tickets?

From 1 August 2016, paper tickets were no longer sold or accepted. Opal cards and Opal single trip tickets replaced the paper tickets for all public transport in Greater Sydney.

The Opal data **does not include** trips made using paper tickets. It only includes Opal card and Opal single trip ticket use. It is important to note that in the first week of Opal data paper tickets were still in use and customers could use a paper ticket or Opal card. In the second week of Opal data only Opal cards or single trip tickets could be used as paper tickets were retired.

Will more weeks of Opal data be available?

We're constantly working to find new data to release and the Open Data Program team are investigating whether further Opal data can be released at a later date.

Where can I get more information about Opal?

Visit the Opal website Opal.com.au or call 13 67 25 (13 OPAL).

What is the Open Data Hub?

The Open Data Hub is where you can find all Transport for NSW open data. It provides access to all APIs and datasets, as well as developer information to get you started and a forum to connect to the Open Data Program staff.