

Transport Open Data Policy

Applicable to:

This Policy applies to permanent, temporary and casual staff; staff seconded from another organisation; and contingent workers including labour hire, professional services contractors and consultants for the following agencies:

- **Transport for NSW**
- **Roads and Maritime Services**
- **Sydney Trains**
- **NSW Trains**
- **State Transit Authority**
- **Sydney Metro**

Division:	Customer Strategy and Technology
Date of Publication:	14 August 2019
Effective Date:	14 August 2019
Review Date:	14 August 2020
Document Approver:	Secretary
Document Owner:	Executive Director, Innovation Research and Reform
Superseded Documents:	CP16005.1 Transport Open Data Policy

1. Purpose

Transport for NSW is committed to providing transport-related data to the community, in line with the NSW Government's Open Data Policy. The purpose of this policy is to:

- Simplify and facilitate the release of appropriate data;
- Raise awareness of the NSW Government's commitment to open data and open government;
- Create a practical policy framework that supports the release of high-value datasets to the community;
- Help operating agencies understand community and industry priorities for open data; and
- Support the *Government Information (Public Access) Act 2009* (NSW) (GIPAA) and promote simple and efficient compliance with the requirements set out in the act.

This policy helps facilitate the implementation of best practice open data principles across the Transport cluster.

2. Introduction

2.1 What is Open Data?

Data and/or information is open to the extent that its characteristics, management, and release meet the principles of openness outlined in this policy. Open data should be both technically and legally open: technically open in that is available and accessible to the community, and legally open in so far as there are appropriate licensing frameworks in place to facilitate its release, use and reuse by the community.

The NSW Government Information and Communications Technology Strategy identifies Open Data as a priority, and the NSW Government Open Data Policy (2013) helps facilitate the implementation of Open Data Principles across the NSW public sector. This policy reflects the principles and requirements of the NSW Government policy and is designed to provide specific guidance in the implementation of open data principles in the Transport environment.

2.2 Benefits of Open Data

Embedding open data principles at all levels of the transport cluster can lead to smarter, more responsive service delivery. This policy is designed to drive benefits to the community because Open Data:

- Promotes open, transparent and accountable government;
- Contributes to the digital economy in NSW, and promotes the development of new business and industries that can make use of government data;
- Leads to better public services;

- Advances community engagement with the government and with the work of government;
- Enables data sharing between government agencies in NSW and across jurisdictions;
- Facilitates greater understanding of agencies own data and the potential of that data; and
- Supports evidence-based policy making and policy research.

2.3 Open Data and the NSW Government Information Management Framework

The [Information Management Framework](#) is a set of standards, policies, guidelines and procedures which enable data and information to be managed in a secure, structured and consistent manner. It defines the whole information life-cycle from data gathering, retention and storage, treatment and analysis through to retirement and disposal. It will ensure that data and information can be appropriately shared or re-used by agencies, individual public sector staff, the community or industry for better services, improved performance management and a more productive public sector.

Many of the elements of best-practice information management are common to, or support, the open data principles or define how they should be implemented. Open data principles should be embedded throughout the information management cycle, so that data can be released to the community without additional re-work.

2.4 NSW Government Data and Information Custodianship Policy

The [Data and Information Custodianship Policy](#) is a key part of the Information Management Framework.

It defines a set of principles for the management and maintenance of the State's core data and information assets. It defines the roles of those responsible for managing information and data within the organisation, and defines and formalises their responsibilities - including decision rights over whether these datasets can be proactively released in keeping with good open data practice.

Clarity around custodianship helps reduce unnecessary effort in the acquisition, maintenance and sharing and release of data and information. It also facilitates homogenisation and harmonisation of data as a specific output, aimed at meeting the needs of government and the community.

3. Open Data Principles

3.1 Open by default

Data collected by transport agencies should be open by default, in accordance with NSW Government Policy.

Consideration should be given to open data principles in planning for data collection so that data can be gathered and stored using methods and formats that support its later release without significant rework.

Openness should be the default position in relation to all datasets, unless there is a specific, overriding reason for data not to be released (refer to section 3.2 for the grounds upon which data should be not be released). Even where a dataset is incomplete, it should be released, with any limitations or gaps highlighted for users in the associated metadata.

It is acknowledged that resourcing and investment may be required to prepare data for release, particularly for existing and real-time datasets. These costs may limit the ability to release some datasets, particularly while agencies establish arrangements for Open Data implementation. High-value datasets should therefore be identified and prioritised for release (see section 3.3 Prioritised).

Where a determination has been made not to release a specific dataset (in part or full) the rationale for not releasing data should be recorded in the dataset's metadata.

3.2 Protected where required

Data should not be released, or not released in full, where any of the following considerations preclude its release:

- Privacy
- Security
- Confidentiality
- Legal privilege
- Public interest

It may be necessary to remove personal or identifying information from datasets in line with GIPAA, the *Privacy and Personal Information Protection Act 1998* (NSW) or other applicable legislation.

For further information on the identification and protection of private data and information, refer to the [Transport for NSW Privacy Policy](#).

For detailed information on the identification and protection of confidential data and information, refer to the Transport for NSW Confidential Information Policy.

Caution should be exercised where transport datasets, although de-identified, could be linked or combined with other government or non-government datasets to re-identify individuals or otherwise breach relevant privacy legislation or policy. If unsure, technical advice should be sought from the Bureau Statistics and Analytics and Group IT to ensure sufficient protections are in place to prevent re-identification of individuals by linking or combining datasets.

3.3 Prioritised

High-value datasets should be identified and prioritised for release. In some cases, public or industry demand for data or consultation with stakeholders will identify datasets as high-value. In other cases, it is the potential to inform the enhancement of services or service delivery or its value to research and evidence-based policy making that will distinguish a high-value dataset.

Even datasets considered low-value internally or for which there is no current demand externally may reveal hidden potential when shared more broadly and should be released wherever possible. Simply making data available and discoverable can enable individuals or organisations to identify novel and innovative uses for it.

3.4 Discoverable

Data must be easily discoverable. High quality metadata enhances the discoverability of datasets. Consolidating the release of transport data via the DataNSW portal (data.nsw.gov.au) greatly improves discoverability. Datasets themselves may reside or be hosted elsewhere, but links to transport data should be provided via DataNSW.

3.5 Usable

Data should be released to the community in a format that makes it easy to access, use, transform and reuse. Data should be made available in machine-readable, non-proprietary formats. The provision of standardised, quality metadata increases the usability of datasets.

3.6 Primary

Wherever possible, data should be released as collected at source, at the most granular level of detail possible, and not in aggregate or modified forms.

For further information on ensuring that data is high quality, usable and sufficiently granular, refer to the NSW Government Data Quality Standard (currently in work). For further information on creating quality datasets and metadata, refer to the [NSW Government Data Interoperability Standards](#).

3.7 Timely

Data should be current and, where practicable, real-time feeds provided where appropriate and where this enhances the utility of the dataset. Publishing dates or timestamps should be included as part of metadata for all datasets to allow users to assess the currency of the data. As data is superseded, agencies must aim to make new datasets available to users in a timely manner.

3.8 Well managed, trusted and authoritative

Users must have confidence that data is trustworthy and authoritative. Any limitations, gaps or qualifications with regards to data accuracy or completeness must be highlighted in the metadata associated with the dataset.

Data must be well managed to ensure its accuracy, integrity and currency and protected from unauthorised or accidental alteration or loss. Managing data and information in accordance with an Information Management Framework assists in producing high quality data that can be made available to the community with little a minimum of additional effort.

3.9 Free where appropriate

Data should be provided free of charge. Where it is not possible to provide data free of charge, it must be provided at a reasonable price. Restrictions on the use or reuse of data should be avoided. The use of the Creative Commons Attribution licence (CC BY AU) is the preferred licensing agreement for Australian Government data, as it lets others distribute, remix, tweak, and build upon the government's work, even commercially, as long as they credit the materials' original creator. This is the most accommodating of licenses offered and is recommended for maximum dissemination and use of licensed materials.

For further information and advice on the application of Creative Commons licenses, refer to the [AusGOAL website](#).

3.10 Subject to public input

Open data facilitates engagement and collaboration between the public sector and the community, and supports the development of new online and mobile applications that make use of those datasets.

Mechanisms for users to engage with the data provider should be developed together with datasets and data delivery channels, to allow users to provide feedback on data provision, quality, usefulness and enhancement, or to request additional datasets. This will allow continuous improvement to be targeted to meet community needs, and provide intelligence on the use of government data by the community.

4. Open Data Maturity

The NSW Open Data policy recommends the use of the [5-star Linked Open Data Model](#), developed by Sir Tim Burner-Lee (2006) to gauge the level of data openness across the organisation. This model defines five levels of openness and usability on a cumulative scale. This measure should be applied to each dataset and the result captured in the dataset's metadata.

5-star Linked Open Data Model

Level	Description
★	Data is available on the web via an open licence
★★	Data is available in a machine-processable, structured form
★★★	Non-proprietary formats are used
★★★★	Universal Resource Identifiers (URIs) and open standards are used to describe data
★★★★★	Data is linked to other datasets to provide context

5. Release of Data

Agencies should develop documented procedures for the release of data, and these should form part of the Information Management Framework. These procedures should document how a data owner or data custodian seeks approval and authorisation for the release of data. Data owners should authorise the release of data where the principles of Open Data have been applied and unless there are specific reasons for data to be protected (see section 3.2 Protected where required). Procedures may be required to define separate processes for the publication of static data and the ongoing provision of real-time, or near real-time, data.

6. Mandatory requirements

Transport for NSW, operating agencies and other transport cluster entities must apply the principles of Open Data in so far as practicable for all data they collect. As per the Open Data principles, data should be made Open by default, but must be protected where required (refer to section 3 for detail).

Resourcing implications can affect an agency's ability to fully implement every principle for each dataset. High-priority datasets must be identified and prioritised for release.

Datasets should be defined individually and documented in the Transport Cluster Data Asset Register in line with the Information Management Framework.

Metadata associated with each dataset should be captured in the Transport Cluster Master Data Asset Register.

Where the data custodian determines that release of a particular dataset is not appropriate, this decision and the grounds upon which it was made must be recorded against the dataset in the Transport Cluster Master Data Asset Register.

7. Accountabilities

The Information Management Framework defines specific roles and responsibilities for the management of data and information across the organisation. Details can be found in the [NSW Government Data and Information Custodianship Policy](#).

Managers are responsible for the implementation of the Information Management Framework, including the identification of appropriate data owners and data custodians for each dataset created.

Managers are responsible for seeing that their staff understand their responsibilities to comply with GIPA legislation as outlined in the agency's GIPA procedures.

8. Document history

Date & Policy No	Approved by	Amendment Notes
11 April 2016 CP16005	Secretary	Approved
1 July 2018 CP16005.1	Secretary	Inclusion of Sydney Metro as agency to which this policy applies from 1 July 2018.
14 August 2019 CP16005.2	Secretary	Updates to obsolete position titles, Division name and template changes to front page.

9. Attachments/related links

The following documents relating to this Policy can be found online at the following locations:

- Legislative Framework:** The NSW Government ICT Strategy Internet Page provides a single point to access [key laws which regulate information management](#) across NSW Government.
- Information Management Framework:** Further guidance on information management can be found in the [NSW Government Information and Communication Strategy site](#).
- Roles and Responsibilities:** Further guidance on accountabilities for information management can be found in the [NSW Government Data and Information Custodianship Policy](#).
- Private Information:** The [Transport for NSW Privacy Policy](#) can be found on the TfNSW Corporate site, and further guidance is available in the [Privacy Plan](#).
- Confidential Information:** Further Information on dealing with confidential information can be found in the Transport for NSW Confidential Information Policy.
- Data Quality:** Information on developing and maintaining high quality data can be found in the NSW Government Data Quality Standard (currently in development) and the [NSW Government Data Interoperability Standards](#).
- Data Licensing:** The [AusGOAL website](#) contains information on licensing government datasets under the creative commons open licensing regime.