Contact

datansw@finance.nsw.gov.au
Department of Finance & Services
Level 15, McKell Building
2-24 Rawson Place
SYDNEY NSW 2000
## DOCUMENT CONTROL

### Document history

<table>
<thead>
<tr>
<th>Date</th>
<th>Version No.</th>
<th>Description</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2013</td>
<td>1.0</td>
<td>Final</td>
<td>Department of Finance and Services</td>
</tr>
</tbody>
</table>

### Document custodian

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Finance &amp; Services, Strategic Policy</td>
<td></td>
</tr>
</tbody>
</table>
Standards, best practice ....................................................................................................................... 15
Principles, policies, frameworks ........................................................................................................... 15
Legislation ............................................................................................................................................. 16
1. PURPOSE

The NSW Government ICT Strategy identifies open data as a priority initiative, and this Policy aims to assist agencies across the NSW Government in embedding open data principles in their operations and in releasing high-value datasets. The open data initiative seeks to make appropriate government data available to industry and the community.

The purpose of the NSW Government Open Data Policy is to:

- Simplify and facilitate the release of appropriate data by NSW Government agencies;
- Make explicit the NSW Government’s commitment to open data and open government;
- Create a practical policy framework that enables high-value datasets to be released to the public;
- Help agencies in understanding 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 that Act.

This Policy helps facilitate implementation of best practice open data principles across the NSW public sector.

2. INTRODUCTION

2.1 Information Management Framework

A key initiative of the NSW Government ICT Strategy is the development of an Information Management Framework to support the way government administers and uses data and information.

The Framework is a set of standards, policies, guidelines and procedures which are implemented either manually or, where possible, automated through technology. This Framework will enable data and information to be managed in a secure, structured and consistent manner.

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. This Policy is part of the Framework.

2.2 Open data principles

Open data principles can lead to more responsive and smarter government, and better service delivery. More specifically, this Policy drives benefit for the people of NSW because it:

- Promotes open, transparent and accountable government;
- Contributes to the digital economy in NSW, and promotes the development of new businesses and industries that can make use of government data;
- Leads to better public services;
- Advances citizen engagement with government and with the work of government;
- Enables data sharing between government agencies, in NSW and across jurisdictions;
- Embeds open data principles across government operations;
• Facilitates greater understanding amongst agencies of their own data and the potential of that data; and
• Supports evidence-based policy making and policy research.

Establishing a culture of open data and open government, and applying the right policy settings from when data is first generated or collected, has been shown to deliver benefits to communities and economies, both in Australia and overseas.

2.3 What is open data?
Data or information is open to the extent that its management, release and characteristics meet principles of openness outlined in this Policy. Open data should be both technically and legally open – in the sense that it is technically available and usable, and appropriate licensing frameworks are in place to facilitate its release and use.

2.4 Open data and open government
Open data supports the open government principles of transparency, participation, collaboration and innovation that are identified as priorities in the NSW Government ICT Strategy. It is acknowledged that open data and open government principles have much in common, and in some jurisdictions the concepts are applied interchangeably.

3. SCOPE
This Policy applies to all NSW Government agencies including cluster agencies, Departments, Statutory Bodies and Shared Service Providers. In accordance with Premier’s Memorandum M1999-19 Applicability of Memoranda and Circulars to State Owned Corporations, the Policy is not mandatory for State Owned Corporations, although it is recommended for adoption.

This Policy’s scope accords with guidance provided to NSW cluster agencies on open government – Guidance for NSW Government cluster agencies on preparing an Open Government Plan. It also takes direction from Premier’s Memorandum M2012-10 - Open Government.

Target audiences for this Policy include government agencies, non-government organisations (NGOs), academia, industry (including ICT developers) and members of the public who are interested in or have a specific use for government data.

4. TERMS
While this Policy focuses on raw unprocessed data, the Policy can also apply to information.

Release of high-value datasets has significant benefit for a range of audiences, and agencies are best placed to determine which of their datasets are regarded as having high-value. In their considerations, agencies should take account of the perspectives of potential data users.

A definition of high-value datasets, as well as details of other terms, are provided at Appendix A.
5. DATA QUALITY

Complying with best practice data quality standards helps facilitate data openness, and many open data principles draw on data quality precepts. Ensuring the quality of data, from when it is first generated or collected, can result in a lower resource impact when related datasets are published. Data quality considerations should include the views of stakeholders who have an interest in the data.

Data should be made available with a statement regarding its quality, as this will allow potential users to determine whether the data is suitable for their requirements. Data quality issues should not unduly delay the publishing of datasets.

A recommended approach is outlined in the ABS Data Quality Framework (2009), which was developed for measuring the quality of statistical data. It identifies seven dimensions in measuring data quality: institutional environment, relevance, timeliness, accuracy, coherence, interoperability and accessibility.

6. PRINCIPLES

The following principles describe what open data looks like when it meets best practice requirements. These principles will not all be fully implemented for every dataset – instead they provide guidance to agencies seeking to implement open data practice.

Resourcing implications can affect an agency’s ability to fully implement every principle with respect to all datasets, and so high-value datasets should be identified as the priority for standardisation and release first.

6.1 Open by default

Data is open by default. Agencies start from a position of data openness, with the prerogative in favour of data release, unless there is a specific, overriding reason for data not to be released in accordance with GIPAA (such as in relation to the public interest test).

Even where a dataset is incomplete, releasing the dataset is favoured – provided sufficient context and metadata is included to notify the user of any limitations or gaps in the dataset.

6.2 Protected where required

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

- privacy – where personal information is involved;
- security – because of the nature of the data or information;
- confidentiality – arising because of the nature of the data or information itself or because a contractual promise has been made in relation to the data or information; or
- legal privilege – attaches to certain legal advice.
• **public interest** – if there are public interest considerations against disclosure and, on balance, those considerations outweigh the public interest considerations in favour of disclosure.

Information labels and security classifications can help indicate whether data is protected.

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

Caution, and in some cases further investigation, may be required where disparate datasets, individually de-identified, could potentially be linked or combined with other government or non-government datasets, to re-identify individuals, or breach relevant privacy legislation or policies.

### 6.3 Prioritised

High-value datasets should be prioritised for release, in line with demand from the public and industry, as a result of stakeholder consultation, or where the release of the datasets will contribute to better service delivery in NSW.

In some cases, even data perceived as low-value or for which there is no current demand, may reveal hidden potential. Simply making data available for discovery enables individuals or organisations to identify new and innovative uses for it.

### 6.4 Discoverable

Data must be easily discoverable and searchable, with good metadata, and to further support this, must be published through a single, easy-to-use data portal: [data.nsw.gov.au](http://data.nsw.gov.au). Data will be made available, as far as possible, at permanent web addresses, so it may be more easily referred or linked to.

Agencies must make reasonable efforts to ensure high-value datasets are accessible and open, so far as this disclosure is consistent with the principles outlined in this Policy, and with applicable standards and legislation. Data should be available on a non-discriminatory basis to anyone where practicable and appropriate.

### 6.5 Usable

Data should be in a format that makes it easy to use, transform and reuse. Commonly accepted open data standards, such as those proposed by the [World Wide Web Consortium (W3C)](http://www.w3.org), will be implemented as far as possible to enhance data usability. Characteristics that support data usability include:

- machine-processable formats;
- non-proprietary formats;
- completeness; and
- clear, high quality metadata.
Standard formats will support inter-agency data sharing, data linking and federation of data portals across Australian and potentially international jurisdictions. Datasets will be designed so they allow connection and linking between multiple public sector agencies, so far as this is practical.

Standardised metadata provides descriptive and contextual information that makes the dataset understandable and more useful. This is particularly important when publishing raw data, which may be difficult to use or interpret without explanation.

### 6.6 Primary

Data should be published as collected at the source, with a high level of granularity, and not in aggregate or modified forms.

### 6.7 Timely

Data should be current, and if practicable, live – with real-time feeds provided as appropriate and where this enhances the utility of the dataset.

Datasets will include timestamps or other information for users to identify the currency of the data, and where newer replacement data becomes available agencies must aim to make this available to users in a timely manner.

### 6.8 Well managed, trusted and authoritative

Data must be well managed to help ensure its ongoing integrity and efficacy for users, e.g. protected from unauthorised alteration or deletion in accordance with agency policies. Appropriate data governance arrangements must be established, and data must be maintained in accordance with NSW Government’s information management principles. Applicable policies include the *NSW Data and Information Custodianship Policy* within the broader NSW Information Management Framework.

It is important for users to have confidence that the data they are accessing is trustworthy and authoritative, and where there are limitations in regard to a dataset, or a dataset is incomplete, users will be alerted to the limitations (see section 5, Data quality).

### 6.9 Free where appropriate

Data should be provided free of charge where appropriate, to encourage its widespread use for innovation, achieve the maximum value from the data for the people of NSW, and to enhance transparency of government.

Where it is not possible to provide data free of charge, or this is not suitable in the circumstances, data will nonetheless be provided at a reasonable price. Factors affecting pricing decisions include the public interest and the effect of other policy or legislative provisions, as determined by the agency.

Restrictions on the use of the data should be avoided, and using the *Creative Commons* (CC BY AU) licensing framework will support data openness and facilitate data usage. Clear licensing
frameworks should help to prevent the misuse of datasets, while also providing a high degree of flexibility for the licensee in their use of the data.

6.10 Subject to public input

Open data facilitates two-way public, industry and stakeholder engagement and collaboration with government, and the potential for new online and mobile applications to be developed to make further use of those datasets.

Data portals should be subject to public input, and have mechanisms by which users can engage with the data provider, and with the broader community around the dataset. A contact person should be available to respond to user queries.

Feedback on data could relate to, for example, data format, quality, granularity, usefulness, accuracy and requests for additional or related datasets. Data users could reciprocate, sharing data generated through their use of government data. This can help develop intelligence about the use of government services, and contribute to enhancements in service delivery.

7. DATA.NSW

NSW Government data will be disseminated through the data.nsw portal. This portal will be a central dataset catalogue, and provide links to datasets, along with some contextual information.

For each dataset agencies must provide adequate metadata, and the details of a contact person in their agency in relation to each dataset. Data.nsw may be contacted through the contact form on the site, or by emailing datansw@finance.nsw.gov.au

8. OPEN DATA MATURITY

Agencies can gauge their level of open data maturity by reviewing their compliance with the principles outlined in this Policy. However, a recommended way to measure data openness is with the 5 Star Linked Open Data Model, developed by Sir Tim Berners-Lee (2006), and adopted in a number of jurisdictions.

The model puts forward five levels of data openness, which are cumulative, in that 2-stars encompasses what is required at the 1-star level, and so on:

* Data is available on the web with an open licence.

** Data is available in machine-processable, structured form (e.g. Excel format instead of an image scan of a table).

*** Non-proprietary formats are used (e.g. CSV, XML).

**** Use Universal Resource Identifiers (URIs) and open standards to describe data (e.g. RDF, SPARQL), so people can point at your data.

***** Data is linked to other people’s data to provide context.
9. LINKED OPEN DATA

This Policy accords with best practice linked open data implementation. Following the principles outlined in the Policy will enhance data discoverability and metadata, and encourage a standardised approach to data quality – which will make it easier for agencies to link their datasets to other open datasets.

In addition, this Policy is designed to help maximise the opportunities of big data, by making larger datasets available for analysis by researchers, industry, and government agencies. Intelligence gained can be used to support service delivery and evidence-based policymaking.

10. IMPLEMENTATION AND REVIEW

A separate implementation plan will be prepared for this Policy to assist agencies in achieving open data objectives.

This Policy will be maintained by the Strategic Policy Branch at DFS, which will report to the NSW ICT Board on its progress. This Policy is subject to review at least every 2 years, or as appropriate.

11. RELATED GUIDANCE

This Policy should be read in conjunction with relevant guidance on the NSW Information Management Framework, in particular:

- The *NSW Data and Information Custodianship Policy*, which addresses the responsibilities of NSW government agencies with respect to specific datasets or information assets; and
- The *Guidelines for Agencies using Data NSW*, which establishes conditions for the appropriate use and re-use of government data.

This Policy operates consistently with all existing applicable NSW laws, policies and standards. Reference documents are set out at Appendix B.
## APPENDIX A – GLOSSARY

<table>
<thead>
<tr>
<th>TERM</th>
<th>DEFINITION</th>
</tr>
</thead>
</table>
| Data    | The representation of facts, concepts or instructions in a formalised (consistent and agreed) manner suitable for communication, interpretation or processing by human or automatic means.  

Typically comprised of numbers, words or images. The format and presentation of data may vary with the context in which it is used.  

Data is not *Information* until it is utilised in a particular context for a particular purpose (Office of the Australian Information Commissioner (OAIC), 2013).  

Data is typically considered to be conceptually at the lowest level of abstraction. |
| Dataset | An identifiable collection of *Data*.  

Most commonly a dataset corresponds to the contents of a single database table, or a single statistical data matrix. The term can also be used to refer to the data in a collection of closely related tables.  

A dataset may comprise a smaller grouping (or subset) of data which, though limited by some constraint or feature type, is located physically within a larger dataset. |
<table>
<thead>
<tr>
<th>TERM</th>
<th>DEFINITION</th>
</tr>
</thead>
</table>
| High-value dataset          | In practice, high-value datasets will be identified by the agency responsible for managing the Dataset (the ‘custodian’).  
High-value datasets often have one or more of the following characteristics (drawn from US and NZ models), in that their release:  
• Responds to a need and/or demand identified through public or stakeholder engagement, or supports positive social outcomes;  
• Has the potential to enhance services or service delivery;  
• Furthers the core mission of the agency;  
• Increases agency accountability and responsiveness;  
• Increases government transparency;  
• Will create economic opportunity, generate efficiencies, or reduce costs; and/or  
• Will support evidence-based policy-making or research. |
| Information                 | Includes government Information as defined in GIPAA.  
Any collection of Data that is processed, analysed, interpreted, classified or communicated in order to serve a useful purpose, present fact(s) or represent knowledge in any medium or form.  
This includes presentation in: electronic (digital), print, audio, video, image, graphical, cartographic, physical sample, textual or numerical form (Office of the Australian Information Commissioner (OAIC), 2013).  
*Information* is typically considered to be at a higher level of abstraction than *Data*. |
| Metadata                    | Data that defines or describes content, quality, format or structure of a Dataset or Information asset; data that defines or describes the system, location and context in which the dataset or information was produced, collected, processed or stored.  
Metadata allows datasets or information assets to be found, understood, controlled and managed. |
| Shared Service Provider     | Provider of a standardised service to multiple clients across a single business stream using a shared platform.  
The list of NSW Shared Service Providers, as amended from time to time, is determined by NSW Government. |
<table>
<thead>
<tr>
<th>TERM</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>End consumer of a <em>Data or Information</em> resource, or those who use data or information for reference, or as input to solve problems and/or make decisions.</td>
</tr>
</tbody>
</table>
APPENDIX B – REFERENCE

The following are sources of related guidance – legislative, policy and standards-based – for the implementation and ongoing management of the open data initiative.

Standards, best practice

This Policy aligns with accepted best practice standards, including:

- 5-Star Open Data Model (Sir Tim Berners-Lee, 2006)
- Australian Standard ISO 15489 – Records management
- ISO 16175 Principles and Functional Requirements for Records in Electronic Office Environments
- NSW Intellectual Property Management Framework for the NSW Public Sector Guidelines
- Standard on Digital Recordkeeping, Standard on Full and Accurate Records, Standard on Managing a Records Management Program, and all related standards issued by State Records NSW

Principles, policies, frameworks

This Policy aligns with information management principles, policies and frameworks in effect in NSW, including:

- Principles on Open Public Sector Information (Cth) (OAIC, May 2011)
- Publishing Public Sector Information Web Guide (Cth) (DFD, Feb 2012)
- Confidentiality Information Series (Cth) (NSS)
- ABS Data Quality Framework Cth (2009)
- Information Security Management Guidelines – Australian Government Security Classification System (Cth) (July 2011)
- NSW Data and Information Custodianship Policy
- Guidelines for Agencies using Data NSW
- NSW Digital Information Security Policy
- NSW Government ICT Strategy

Memoranda, circulars

This Policy aligns with the following memoranda:

- M1999-19 Applicability of Memoranda and Circulars to State Owned Corporations
- DFS Guidance for NSW Government cluster agencies on preparing an Open Government Plan
- M2012-10 Open Government
- DFS C2013-1 OpenGov NSW
- C2002-69 Labelling Sensitive Information
Legislation

Legislative instruments relating to this Policy include:

- *State Records Act 1998 (NSW)*
- *Government Information (Public Access) Act 2009 (NSW) (GIPAA)*
- *Privacy and Personal Information Protection Act 1998 (NSW)*
- *Health Records and Information Privacy Act 2002 (NSW)*
- *Copyright Act 1968 (Cth)*